Eye tracking student strategies for solving stoichiometry problems involving particulate nature of matter diagrams

John Y. Baluyut¹ / Thomas A. Holme²

¹ Math and Science Division, University of Providence, Great Falls, MT 59405, USA
² Department of Chemistry, Iowa State University, Ames, IA 50011-1021, USA, E-mail: taholme@iastate.edu.

Abstract:
This study compared how students obtained and used information from particulate nature of matter (PNOM) diagrams as well as balanced chemical equations when asked questions about stoichiometry concepts such as limiting and excess reagents, and yield. The comparisons were made in terms of visual behaviors by examining eye fixations while students responded to a 30-item online instrument. Statistically significant differences were found between visual behaviors of high- and low-performers on seven items, mostly those that dealt with correctly identifying the limiting reagent for each diagram. High-performing students were found to have spent more time examining PNOM diagrams and transitioned more frequently between parts of the diagrams and other areas of interest (AOIs) than low-performing students did. This study gives an example of how underlying strategies students used to respond to conceptual stoichiometry questions may be triangulated by the use of both quantitative and qualitative techniques in relation to eye tracking visual behaviors.

Keywords: chemical education research, eyetracking, first-year undergraduate general chemistry, PNOM diagrams, stoichiometry

DOI: 10.1515/cti-2018-0003

Introduction

The use of both symbolic and microscopic representations is an important skill for students to learn in chemistry. Microscopic representations, such as particulate nature of matter (PNOM) diagrams can be used to teach many chemical phenomena. In the introductory 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 show that while a majority of students in a first-year chemistry course can write balanced chemical equations accurately, only slightly more than one out of five can translate chemical reactions represented by PNOM diagrams into the corresponding chemical equation (Davidowitz, Chittleborough, & Murray, 2010). These findings have been expanded upon over the years including more recent studies that have noted challenges students have drawing accurate molecular connectivity (Kern, Wood, Roehrig, & Nyachwaya, 2010), and depictions of equilibrium (Akaygun & Jones, 2014), in addition to the depiction of reactions. The role of teaching pedagogies in the laboratory (Dickson, Thompson, & O’Toole, 2017), or with active methods such as POGIL (Prilliman, 2014), have also been the focus of studies on the role of PNOM diagrams in enhancing student understanding of chemistry at the atomic/molecular scale. Despite teaching interventions that use PNOM diagrams, many students still find it challenging to understand chemical stoichiometry at the microscopic level (Ben-Zvi et al., 1987; Sanger, 2005).

When considering how people learn, dual coding theory suggests that information from text and diagrams are coded in different cognitive systems due to their different physical forms (Paivio, 1990). Thus, when information is displayed with both text and diagrams learners need more cognitive resources than when solely using text or pictures (Hegarty, Carpenter, & Just, 1991). The routine use of PNOM diagrams requires students to obtain visual information after reading sections of text. They must then try to build coherent visual and
verbal mental models from each representation and then integrate both mental models using their own prior knowledge to generate learning.

Given these cognitive challenges, students who have different problem-solving abilities usually have visual behaviors that show differences in eye movement patterns (Rosengrant, 2010; Tang, Topczewski, Topczewski, & Pienta, 2012; Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009). The strategy difference between students with high and low prior knowledge most likely comes from long-term memory (Ericsson & Kintsch, 1995). Students with greater proficiency likely have more domain-specific schemas and, thus, may sometimes bypass working memory capacity limits because more of their schemas have become automated (Kalyuga, Chandler, & Sweller, 1998; Kalyuga, Ayres, Chandler, & Sweller, 2003). These proficient students often possess a greater ability to attend to domain-relevant information, faster processing of such information, and show overall improvement of performance (Liu, Gale, & Song, 2007; Weber & Brewer, 2003). Eye tracking has shown how students of different abilities also differ in how they allocate their attention to different parts of visual stimuli and that students’ reading abilities, for example, may be inferred based on the different lengths of time they spend comprehending text materials (Kozma, 2003; Schmidt-Weigand, Kohnert, & Glowalla, 2010). In chemistry, experts have been found to coordinate information within and across representations quite easily while most students have 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 effort learners spend to integrate information (Mason, Pluchino, Tornatora, & Ariasi, 2013; Schwonke, Berthold, & Renkl, 2009).

Chemistry education is an especially important domain to understand the coordination of representations. Reasoning in chemistry often deals with unobservable concepts and processes, which is why the use of visualizations to relate information is important. 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 durations, fixation counts, and transitions between areas of interest?

2. How do students with different levels of proficiency integrate information from symbolic and microscopic representations between different areas of visual interest?

Methods

Research participants

Participants were recruited from two different chemistry courses. One group was recruited from those who were registered in the second-semester course of a two-semester general chemistry sequence. This course covers solution properties, kinetics, thermodynamics, electrochemistry, chemical equilibrium, and nuclear chemistry and will be referred to as Chem A. Students from this course were chosen from among those who took the first semester of the same course, during the previous fall semester. A second group of students was recruited from a one-semester survey of general chemical principles. Topics discussed in this course 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 and will be referred to as Chem B. Participants from Chem B were recruited during the 2 weeks immediately following their examination on stoichiometry. It was originally hypothesized that because students who came from Chem A were taking a course with more detailed content coverage they would exhibit visual behaviors that are different from those students coming from the survey-style Chem B course.

A total of 15 students from Chem A and 14 students from Chem B participated in this study. Data from one student from each group 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.

Apparatus

An SMI Red eye tracker with BeGaze 3.3 software was used to collect all experimental data. The eye tracker was placed directly below a 23-inch LCD monitor, on which the questions for the study were displayed. All visual
stimuli were maximized in size to occupy the full area of the monitor. Each student was seated about 60–70 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 participant. This calibration and validation routine was repeated until eye tracking resolution came to within 0.5° along both dimensions of the monitor.

**Procedure**

A 30-item instrument covering concepts of stoichiometry such as excess and limiting reagents, as well as yield, was developed based on preliminary interviews with 18 students from two different general chemistry courses at a large state university: a one-semester advanced general chemistry class, designed specifically for first-year students either in the honors classes or with advanced high school preparation in chemistry; and the first semester of a 1-year general chemistry course. The interviews were conducted no more than 2 weeks after students took their examination covering stoichiometry concepts mentioned above. During these interviews, subjects were given two stoichiometry problems of different chemical contexts. Among the tasks assigned, students were asked to draw diagrams of the problem when none were provided, explain the meanings of representations used in their diagrams and how they would change if chemical contexts were changed, and to numerically rate their perception of the difficulty of the chemical problems. Students were asked to think aloud as they went through solving each problem. The responses given during the interviews were used as basis for developing the instrument. Initial, large scale validation of the instrument occurred twice, via online administrations among a much larger group of students from several general chemistry courses during the semester prior to this study. This administration occurred once during the 2 weeks immediately following an exam covering stoichiometry, and then again between 2 and 4 weeks before the end of the same semester to determine test-retest validity. The same instrument was used for this eye tracking study with participants seeing 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 displayed on the monitor. An example of the eye tracking item arrangement is shown in Figure 1. Additionally, PNOM diagrams only for the subsequent sets of questions are provided in Figure 2.

![Diagram](image)

**Figure 1:** Example problem from the eye tracking study. Several questions were asked for each PNOM diagram pair.
Students from Chem A were interviewed first in this project. In the process of initially evaluating data collected from these 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. To remedy this artifact, students from Chem B 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. Thus, participants from Chem A saw five online pages for each chemical context given in the instrument, while those from Chem B saw six. No time limit was imposed on students in responding to the instrument, although it was observed that participants took between 5 and 25 min to go through all items. This time range suggests our sample covers a broad level of comfort with interpreting PNOM diagrams among the participants in the study.

The instrument was designed so that items were split among three different chemical contexts (Figure 1 and Figure 2): complete combustion of methane involving relatively small numbers of molecules; synthesis of ammonia from the elements at 50 % completion also with a small number of molecules in the PNOM diagram; and dissubstitution of carbon tetrachloride with hydrogen fluoride at 75 % completion using a much larger number of molecules. The order of the chemical contexts was designed to slowly build into the instrument increases in both the conceptual and visual complexities of the problems. For each chemical context there were two sets of PNOM diagrams with the only difference being what was pictured in the products box. At most, only one of the two product diagrams shown was correct. Both diagrams for each context came with the same set of five items noted in Box 1:

**Box 1:**

| a. The diagram shows the reactants being allowed to react completely with each other. (True, False, I don’t know) |
| b. The correct number of unreacted molecules for the reaction is drawn in the diagram. (True, False, I don’t know) |
| c. The diagram shows the correct ratio between reactants consumed and products formed by the reaction. (True, False, I don’t know) |
| d. The number of molecules of product is based on the correct limiting reagent. (True, False, I don’t know) |
| e. Based on your choices from the previous questions, is the given diagram correct or not? (Correct, Incorrect) |
These first four items in Box 1 were randomly ordered for each diagram to eliminate any item-order effects. The final question on the overall correctness of the diagram was asked as the fifth question of each set.

After completing the instrument, each student was shown a playback of their gaze video, which showed them exactly what they were looking at as they responded to each question. Students were asked to describe out loud the specific steps they went through mentally as they watched their video. They were told to use the gaze video to remind themselves of their thought processes as they responded to each item. These retrospective think-aloud (Holmqvist et al., 2011) sessions lasted between 35 and 50 min each.

Data analysis

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 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). AOIs for this study were defined around each side of the balanced chemical equation (labeled U and V, respectively), each side of the PNOM diagram given in each item (X and Y), and on the question stem and answers for each item (Z). Strings consisting of characters denoting the different AOIs were written out to represent the sequence with which subjects viewed each AOI. Temporal binning was incorporated into the AOI strings by repeating characters corresponding to every 25 ms of fixation on each AOI (Cristino, Mathôt, Theeuwes, & Gilchrist, 2010). This way scan paths generated from eye tracking data included location, sequence, and durations of eye fixations on visual stimuli. Figure 3 shows an example of what could be a segment of a subject’s fixations. Taking the tail of the blue arrow as the point of initial fixation and the head of the red arrow as the point of final fixation, and going on in a tail-to-head manner the scan path segment for this subject would be coded as “UXUVYUXXY.”

![Diagram of the coding procedure used to produce AOI strings from scan paths.](Figure 3: Schematic diagram of the coding procedure used to produce AOI strings from scan paths.)
Sequence alignment

AOI strings were compared with each other, pairwise, to identify similarities and differences among behaviors of subjects with respect to visual stimuli using the Needleman-Wunsch algorithm (Needleman & Wunsch, 1970). This algorithm has been used for decades in bioinformatics to analyze DNA or protein sequences. In the present study, it 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 (Wolfram Research Inc, 2014). The best alignment between two strings is determined iteratively to obtain a similarity score for these two strings.

Permutation test

Similarity scores of the AOI strings for each item were compared with each other using the permutation test (Feusner & Lukoff, 2008). The permutation test is a nonparametric test used when scores are associated with pairs of, rather than individual, subjects. Because a similarity score is computed between two different scan paths, it does not provide direct numerical measures of individual scan paths. The use of more common statistical methods such as a t-test or a Wilcoxon signed-rank test is, thus precluded. The null hypothesis of a permutation test is that members of the groups of participants used in a study are interchangeable, and that differences observed among them may not be due to the initial basis of the experimental groups. The permutation test as used was implemented using a Python 3.0 script. The permutation test in this study was optimized by varying the number of regrouping samples from 10 to 20,000 to make sure that reliable p values were obtained.

Results and discussion

The number of correct responses from participants enrolled in Chem A ranged from 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 B 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 α = 0.05.

Reclassification of participants based on performance

Despite these overall performance differences, when participants were grouped based on the chemistry course they took, the p values obtained from the permutation tests showed nearly no items with statistically significant differences in eye tracking scan path strings between students from Chem A and Chem B. The only exception was for item 30, the last item. This may be attributed to instrument fatigue on the part of one group of participants, although this hypothesis was not verified.

Thus, differences among participants grouped by course did not show any important influence on the student’s visual behavior in analyzing the diagrams given in the instrument. As a result, the participants were regrouped based on scores obtained on the instrument. Those who obtained at least 15 items correctly were grouped as high-performers and those who did not consisted the low-performing group regardless of which course they came from. Of the 27 participants, 18 were classified as high performers. All of the following discussion in this paper will be based on participant grouping based on performance.

The p values resulting from permutation tests based on this new classification revealed statistically significant differences between high- and low-performers on seven items. Of these seven items, four dealt specifically with the correct choice of the limiting reagent, two were about numbers of unreacted molecules, and one was on the ratio between numbers of reactant molecules used by the reaction and number of product molecules formed. The highest incidence of significant differences between the visual behaviors of high- and low-performers on items pertaining to identifying the correct limiting reagent is consistent with another study’s finding that a substantial fraction of general chemistry students could not identify limiting reagents correctly (Davidowitz et al., 2010) even with the use of PNOM diagrams.
Quantitative differences between high- and low-performers across all items

Table 1 lists mean number of AOI transitions high- and low-performers went through on items that yielded statistically significant differences between students based on overall performance in terms of their AOI string similarity scores using the permutation tests. Early in the study, high-performing participants generally moved across AOIs on these items more frequently than low-performers did. Only by the end of the interview (at item 29) did the transitions among AOIs from low-performers catch up to those of the high-performers. Thus, it appears that high-performers invest more cognitive resources early in the process which then notably improves their efficiency, while low performers change their gaze patterns less dramatically over the course of the interview.

| Item number | Mean AOI transitions |
|-------------|----------------------|
|             | High-performers      | Low-performers      |
| 3           | 20.6                 | 10.6                |
| 4           | 13.2                 | 12.3                |
| 13          | 11.9                 | 7.6                 |
| 17          | 8.5                  | 8.0                 |
| 18          | 8.8                  | 7.1                 |
| 28          | 7.7                  | 6.1                 |
| 29          | 5.8                  | 5.9                 |

Another key difference between high- and low-performers seen across all items is the way fixation durations on PNOM diagrams themselves consistently decreased for high-performers as participants went from one item to the next in the same chemical context. These observations that suggest early cognitive investment of high-performers in analyzing the PNOM diagrams is corroborated by the fixation durations that arise when visual complexities of the diagrams increase. Mean fixation durations of more than 9 and 10 s, respectively, on the reactant and product sides of the first diagram for the carbon tetrachloride reaction were observed for high-performers. Mean fixation durations on the diagrams among low-performers, on the other hand, remained closer to 2 or 3 s even as visual complexity increased. It appears that high-performers were more deliberate and attempted to integrate more information in analyzing changes between diagrams than low-performers did.

Beyond fixation durations, heat maps provide an important perspective in analyzing the fixation durations spent by different groups of subjects on the computer screen, including AOs. Heat maps use color gradients to show the variation of mean fixation durations on different AOs among participants from the same group. More intense colors (red or orange) indicate the longest mean fixations on an AO while less intense colors (blue or green) indicate shorter fixations.

Figure 4 shows heat maps across items 22 through 25 which is the first set of items for this visually complex context. The heat maps show high performing and low performing students with similar cognitive efforts on the question stem (at the bottom). By contrast, the visual processing of the PNOM diagrams is different, as low performing students needed to constantly attend to the diagram rather than 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-performers stop paying much attention to the diagram as they proceed to later items as indicated by the lack of color overlaying the diagrams lower on the series of heat maps.
Heat maps show the relative lengths of eye fixations on different AOs within a visual stimulus. Areas of interest that receive the longest average fixation durations are overlayed with orange to red clouds, while those that attract short fixations have green to blue clouds on them. The apparent decrease in fixations for high-performing students mostly corresponded with decreases in visual attention given to the PNOM diagrams for these items. On the other hand, low-performing students spent large fractions of their times analyzing the diagrams.

This observation can be quantified by looking at Table 2, which lists mean times spent on the PNOM diagrams (AOIs X and Y) as well as total time spent for each of items 22 through 25 by high- and low-performing groups. In this case, statistically significant differences in fixation counts and AOI transitions were analyzed using rank biserial correlation coefficients (Cureton, 1956), $r_{pb}$, which relate dichotomous and ordinal variables. Two trends are important. First, high-performing students consistently decrease time spent looking at the PNOM diagrams for later questions about them. Second, low-performing students both spend more time analyzing the diagrams and do not consistently lower time spent. This difference in attentiveness to the PNOM diagrams most likely results from greater ability to store information into working memory by high performers even as low performers partially catch up later in the series.

To further understand how different visual approaches may imply differences in student thinking about PNOM diagrams, we will focus on the items that asked participants if the diagram correctly depicted the limiting reagent for each context (see Table 3). The first thing to note from this data is that for most cases where significant difference arise, high-performing students are investing more effort in analyzing the PNOM diagram. The only time lower-performing students invest more attention to an AOI is to the text of the question in the first scenario. This evidence from fixation durations bolsters this interpretation as the high-performing students are using more fixations than the low-performing students on the PNOM diagrams. As the experiment proceeded, the correlations between fixations and performance were strongest for the product portion of the PNOM diagram (AOI – Y). This result suggests that the high performing students are taking their cues from observing the “unreacted” molecules significantly more than the low performing students. This observation
further suggests ways that instructional interventions may help students learn how to use PNOM diagrams, by noting that inclinations to start with reactants and then consider products. While sensible in terms of temporal progress of a reaction, moving from reactants to products may not be the most productive strategy for all questions that might be asked.

Table 2: Fixation and total time spent on visually complex PNOM items.

| Item number | Group | Fixation durations on AOIs X and Y | Total time spent on item |
|-------------|-------|-----------------------------------|-------------------------|
|             |       | Mean, s p-value                   | Mean, s p-value         |
| 22          | High  | 3.939 0.0381                      | 6.786 0.0100            |
|             | Low   | 8.956 15.50                       |                         |
| 23          | High  | 1.692 0.0452                      | 5.488 0.0336            |
|             | Low   | 4.488                             | 9.896                   |
| 24          | High  | 1.009 0.0446                      | 6.628 0.0567            |
|             | Low   | 4.857                             | 11.314                  |
| 25          | High  | 0.950 0.0597                      | 2.446 0.0474            |
|             | Low   | 2.312                             | 4.889                   |

Table 3: Means and correlation coefficients for fixation counts associated with areas of interest for items questioning the correct depiction of the limiting reagent (4, 13, 17, 28) \( r_{rb} \) = rank biserial correlation coefficient.

| Item number | AOI                  | Fixation count |   |       |
|-------------|----------------------|----------------|---|-------|
|             | (U) Eqn- reactants   | 2.2            | 2.2| 0.200 |
|             | (V) Eqn- products    | 3.4            | 3.4| -0.068|
|             | (X) PNOM -reactants  | 6.2            | 6.2| 0.368 |
|             | (Y) PNOM -products   | 9.2            | 9.2| 0.239 |
|             | (Z) question         | 9.5            | 9.5| -0.626|
| 13 (NH\textsubscript{3}) | (U) Eqn- reactants   | 2.1            | 2.1| 0.043 |
|             | (V) Eqn- products    | 0.6            | 0.6| 0.012 |
|             | (X) PNOM -reactants  | 12.5           | 12.5| 0.478 |
|             | (Y) PNOM -products   | 8.7            | 8.7| 0.447 |
|             | (Z) question         | 32.8           | 32.8| 0.012 |
| 17 (NH\textsubscript{3}) | (U) Eqn- reactants   | 1.2            | 1.2| 0.006 |
|             | (V) Eqn- products    | 0.1            | 0.1| -0.175|
|             | (X) PNOM -reactants  | 7.4            | 7.4| 0.259 |
|             | (Y) PNOM -products   | 13.9           | 13.9| 0.319 |
|             | (Z) question         | 24.2           | 24.2| 0.000 |
| 28 (CCl\textsubscript{4}) | (U) Eqn- reactants   | 0.6            | 0.6| 0.049 |
|             | (V) Eqn- products    | 0.4            | 0.4| 0.056 |
|             | (X) PNOM -reactants  | 5.3            | 5.3| 0.130 |
|             | (Y) PNOM -products   | 9.2            | 9.2| 0.375 |
|             | (Z) question         | 18.3           | 18.3| -0.207|

*Moderate correlation.

Retrospective think-alouds (RTAs)

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 they worked on 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 RTA was used in this study instead of the concurrent think-aloud (CTA), during which a participant explains their thought processes as items are responded to and eye movements are being recorded (Holmqvist et al., 2011). This choice minimizes complications associated with the think-aloud process itself using cognitive resources. RTAs were used in this study to
elucidate strategies used by participants in responding to the different items on the instrument. These strategies are discussed next.

**Atom-to-formula matching**

A main concern among participants when they look at the diagrams on the instrument was to match each of the different circles in the PNOM diagrams with the appropriate element as illustrated by diagonal lines in Figure 5.

![Diagram](image)

**Figure 5**: Eye movements associated with matching circles in the PNOM diagram with the correct chemical in the balanced equation (in the blue oval).

This type of engagement was observed even as the matches between colors and elements were all explicitly stated in each context. Participants also had to determine for themselves which groups of circles corresponded with which formula in the given balanced equation. Often, this involved vertical eye movements from one side of the equation to the same side of the diagram. An example of this type of strategy explanation from the RTA of student 18, a high performer, demonstrates this style of engagement with the PNOM diagram.

**Student 18**: Yeah. *I was just trying to match up like where it says N₂, where it was trying to say, okay, this is N₂ or this is H₂ and just counting the molecules and trying to compare that to what was stated in the diagram.*

**Interviewer**: When you say you match up N₂ with N₂, what does that mean?

**Student 18**: 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.

**Atom/molecule counts**

Counting atoms and molecules was also commonly observed among participants as they moved their eyes from one atom (or molecule) to the next within the same box of the PNOM diagram. These were generally characterized by roughly circular eye movements with brief pauses on atoms or molecules of the same kind within the same box of the PNOM diagram (see Figure 6).
Figure 6: Eye movements associated with counting atoms and/or molecules in a PNOM diagram (in the red circle).

Many participants also compared numbers of atoms on both sides of the diagram, starting with one box first then moving on to the other: “Counting the ones in the first box, so I can go to the second box and see if they’re the same number represented.”

Several participants reported counting atoms and molecules in terms of sets based on coefficients given in the balanced equation: “...then I looked back and again trying to mentally group them and see if I saw that grouping on the right side.” This is similar to students’ use of lines to group atoms and molecules together when they have PNOM diagrams printed on paper rather than displayed on a monitor (Davidowitz et al., 2010).

Atom-to-atom matching

This phrase refers to participants whose visual attention switched back and forth between both sides of the PNOM diagram (see Figure 7) to determine how each reactant atom ended up on the product side: “I looked at the diagram and then I’d be like CH₄, find it, and then I looked at the other side and I tried to find the same CH₄, 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.” Participants described how they tried to mentally break up reactant molecules into their component atoms and then regrouped them to form product molecules: “…then I looked back and again (tried) to mentally group them and see if I saw that grouping on the right side.” On several occasions, this strategy was used by participants to make sure that the same number of each atom was drawn on both sides of the diagram. Several participants stressed the importance of the law of conservation of mass in making sure the diagram was correctly drawn: “I guess the main thing would be that there is the correct number of molecules in the reactant side and the product side and that nothing is lost in the diagram.”
Summary and teaching implications

The previous sections illustrate how eye tracking can be used to determine differences in the visual behaviors on PNOM stoichiometry diagrams between high- and low-performers in terms of eye fixation sequences, durations, and counts on areas of interest. High-performers were generally observed to be better at keeping track of information they have obtained from PNOM diagrams as indicated by gradually decreasing fixation durations in going from one item to the next. High-performers generally chose to compare information obtained from the diagrams most frequently with those obtained from other AOIs. Low-performers, meanwhile, tended to focus on having a better understanding of the question being asked for each item.

It is clear from the RTAs that participants tended to focus on underlying strategies when asked to describe visual behaviors they went through as they responded to each item on the monitor. Participants generally showed good recollection of their thought processes as they each saw recordings of their visual behavior on their gaze videos. Most comments obtained from participants were focused on reasons behind the way participants looked at the different AOIs for each item.

Introductory chemistry instructors may very well benefit from explaining more explicitly to their students how to obtain information from PNOM diagrams. It is clear that even with the use of such diagrams, certain misconceptions remained quite resilient based on how even some of the high-performers responded to certain items on the instrument used for this study. This was observed even on items that revealed significant differences between the visual behavior of high- and low-performing participants. It might help to illustrate how atoms and molecules present in PNOM diagrams may be counted in groups based on coefficients given in the balanced equation for the specific chemical context being described.

Acknowledgments

This work was supported by the American Chemical Society Examinations Institute.

References

Akaygun, S., & Jones, L. L. (2014). Words or Pictures: A comparison of written and pictorial explanations of physical and chemical equilibria. *International Journal of Science Education, 36*(5), 783–807.

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

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.
