Chemistry Education Research and Practice

PAPER

View Article Online
View Journal | View Issue



Cite this: *Chem. Educ. Res. Pract.*, 2019, **20**, 570

# An explanative basis for the differential performance of students with low math aptitude in general chemistry

Vanessa R. Ralph and Scott E. Lewis*

Students who score within the bottom quartile on cognitive measures of math aptitude have been identified as at-risk for low performance in chemistry courses, with less attention as to *why* such differential performance persists. At-risk students struggle most differentially on assessment items related to the mole concept and stoichiometry. An exploration as to the nature of the differential performance observed became of great interest as the assessment of these topics rarely progresses beyond multiplication or division, and at-risk students who achieved proficiency with the mole concept and stoichiometry had no noticeable gaps in academic chemistry performance when compared to students scoring in the top three quartiles of math aptitude. Thus, students in first-semester general chemistry were surveyed to describe their solution processes toward assessment items involving the mole concept and stoichiometry. Three hundred and forty-eight students responded to all survey prompts with 101 identified as at-risk. Findings suggest that while all students were observed to struggle in the conceptualization of the algorithms by which they execute solution processes, not-at-risk chemistry students were more likely to achieve correct answers *via* chemically implausible solution pathways. Rather than suggest the removal of assessment practices involving algorithmic, multiple-choice assessment on these topics, the implications include practical suggestions and opportunities for further research toward improving the equitability of measures used to assess proficiency with stoichiometry.

Received 6th March 2019,
Accepted 23rd May 2019

DOI: 10.1039/c9rp00068b

rsc.li/cerp



## Introduction

### Prediction for the sake of intervention

Cognitive measures of math aptitude are heavily used in the research literature to identify students at-risk for unfavorable outcomes in college-level chemistry (Coley, 1973; Pickering, 1975; Andrews and Andrews, 1979; Ozsogomonyan and Loftus, 1979; Rixse and Pickering, 1985; Spencer, 1996; Wagner *et al.*, 2002; Lewis and Lewis, 2007; Hall *et al.*, 2014; Ye *et al.*, 2016). Each study provided a means by which instructors could identify at-risk chemistry students *via* incoming preparation in mathematics. More recent work identifies at-risk students as those scoring at or below the bottom quartile of math aptitude scores for each semester's cohort, which will be used in the current study (Lewis and Lewis, 2007; Ye *et al.*, 2016; Ralph and Lewis, 2018). While each of these studies presents evidence for the utility of math aptitude in predicting students' assessment performance in chemistry, the work falls short in developing an understanding as to *why* students of low math aptitude perform poorly in general chemistry.

Many of the cited works provide practical implications for their findings. Implications include assigning students to remedial instruction (Coley, 1973; Pickering, 1975; Wagner *et al.*, 2002), procedures at the institutional level to apply cut-offs for math aptitude scores necessary for enrollment (Coley, 1973), the assignment of remedial coursework or instruction with an emphasis in problem-solving (Andrews and Andrews, 1979), and the promotion of productive study habits by assigning practice problems for the students to complete before attending lectures (Ye *et al.*, 2016). Each set of implications relies on the reasonable assumption that students of low math aptitude scores struggle with the mathematical components of chemistry. However, in 2012, Scott analyzed student response processes in the form of handwritten solutions to analogous mathematics and chemistry problems and identified no measurable change in students' chemistry performance after practicing analogous mathematics problems. In 2009, Donovan and Wheland hypothesized that the connection between mathematics and chemistry is unclear and may not be dependent on "actual mathematics knowledge" but rather the development of a higher-order, cognitive skill set required in science. The article concludes: "However, if success in chemistry stems from mastery of specific mathematics skills and knowledge, then those specific skills must be identified and built upon for students who will take college-level chemistry classes."

*Department of Chemistry, University of South Florida, USA. E-mail: slewis@usf.edu*

Additionally, limited or no success has been observed as to the influence of prerequisite coursework on performance gaps (Pickering, 1975; Donovan and Wheland, 2009; Hailikari and Nevgi, 2010). From the research literature, it is clear that efforts to improve the success of students with low math aptitude are widespread and would benefit by furthering the evidence base as to why differentials persist in academic chemistry performance.

Prior work sought to identify the topics on which at-risk chemistry students most differentially perform (Ralph and Lewis, 2018). Students scoring in the bottom quartile of math aptitude were observed to perform most differentially, or experience the largest performance gaps, on items belonging to the mole concept and stoichiometry. Assessment of these topics rarely progressed beyond multiplication or division. Additionally, this previous study observed at-risk chemistry students who attained a proficiency of 65% or higher on each semester's collective mole concept or stoichiometry assessment items to outperform both their peers in the at-risk group and those students not-at-risk. Despite the identification of the topics on which at-risk chemistry students most differentially perform and the success of students following proficiency on these topics, the study was unable to address *why* at-risk chemistry students differentially performed on these topics. More specifically, amongst the assessment items within these topics, what features of the assessment items posed differential challenges to at-risk chemistry students? The current study, thereby, seeks to understand why the mole concept and stoichiometry present a disproportionate challenge to at-risk chemistry students and evaluate features of assessment design that result in the observed differential performance.

## Challenges chemistry students face with the mole concept and stoichiometry

No prior work was identified to describe challenges with the mole concept or stoichiometry amongst chemistry students with low math aptitude scores. However, other works (described below) have outlined the challenges of low-performing chemistry students on these topics. The use of symbolism to communicate chemical equations has often been observed in the instruction and assessment of stoichiometry and has been described as difficult for students to interpret (Davidowitz *et al.*, 2010). Representational competence comprises the skill sets by which elemental symbols and chemical formulae or reactions are interpreted (Kozma and Russell, 2005). To provide an example of representational competence in the context of problem-solving, suppose a stoichiometry prompt requests the mass of a product given the mass of a reactant. One would then use the provided nomenclature, chemical formulae, or a chemical reaction to interpret coefficients and subscripts and discern the number of atoms comprising each substance and the stoichiometric ratio between reacting chemicals. In the context of stoichiometry, low-performing students were observed to conflate the roles of coefficients and subscripts in balanced chemical reactions (Potgieter and Davidowitz, 2011). These challenges with the interpretation of coefficients and subscripts appeared lessened *via* student involvement with the illustration of submicroscopic diagrams (pictorial depictions of chemical reactions) to visualize proportional relationships between chemical quantities (Davidowitz and Chittleborough, 2009). These studies describe student struggles with representational competence likely impeding their ability to quantify chemical proportions (*e.g.*, deriving solutions to stoichiometry problems). A study published by Schank and Kozma in 2002 examined the impact of learning environments emphasizing representational competence *via* an interactive computer program that provides visualizations of chemical systems. In this study, chemistry teachers communicated the importance of representational competence and provided ideas for the use of these skill sets toward students' understanding of the law of mass conservation and how chemicals rearrange in a chemical reaction retaining stoichiometric proportions.

Similar to representational competence approaches for enhancing instruction to support student difficulty with systematic approaches, or the step-by-step solution processes used to arrange conversion factors, seemed to promote increasingly visual instruction (Gabel and Sherwood, 1983; Phillips, 2001). Phillips' (2001) dissertation describes the efficacy of graphic organizers toward supporting young students' transitions from Piagetian concrete-operational reasoning, or logic reasoning applied to physical entities, to that of formal-operational or abstract reasoning where deductive reasoning concerns hypothetical situations without the need to connect to a physical entity. An example of these forms of reasoning could involve how students solve the following prompt: "If Rob is taller than Jon and Jon is taller than Brandon, who is tallest?" If the student can deduce who is taller without creating a concrete representation, the student is likely relying on formal-operational reasoning. If however, the student illustrates a picture to support their reasoning, the student may be in the concrete-operational stage but can use the image to reason nonetheless. Connections between concrete- and formal-operational reasoning to students of low math aptitude scores have been made in the literature (Bender and Milakofsky, 1982; Lewis and Lewis, 2007). Phillips posits graphic organizers can support students of concrete-operational reasoning to progress toward more formal-operational reasoning. A chemistry example of a graphic organizer could be a solution flowchart where bidirectional arrows between mass and moles are labeled with the operations involving molar mass required to interconvert these values. Just as more visual instructional practices appeared to promote students' progression toward formal-operational reasoning by proving a framework from which students can connect abstract concepts, so too have increasingly visual instructional practices positively impacted the academic success of high school chemistry students presenting with relatively low proficiency in proportional reasoning when solving stoichiometry problems (Gabel and Sherwood, 1983). "Supplemental, less mathematical, and more visual approaches" were recommended for students who struggled with proportional reasoning (Gabel and Sherwood, 1983, p. 175). Interestingly, where visuals have been successful, practice with proportional reasoning in mathematical contexts was not observed to transfer to the context of chemistry (Scott, 2012; Ramful and Narod, 2014).

In these studies mathematical skill sets were not inherently transferred to analogous chemistry skill sets particularly where proportional reasoning was involved (as is observed with stoichiometry).

Thus, a final theme concerning the challenges observed amongst chemistry students in the literature was conceptual understanding regarding measures particular to the mole concept and its application to stoichiometry (BouJaoude and Barakat, 2003; Dahsah and Coll, 2007). In the works cited previously, students were unable to distinguish between relevant units of measure concerning stoichiometric conversions (*e.g.*, mass, moles, and molar mass). Analyses of textbooks and instructional materials suggest that some of these challenges could result from students' introduction to these quantities (Staver and Lumpe, 1995; Larson, 1997). A literature review conducted in 2002 also describes challenges with the conceptual understanding of stoichiometry as related to how these entities are introduced during instruction (Furió *et al.*, 2002). The review identifies a consensus in the literature concerning students' conceptual misunderstanding of units used in stoichiometry as attributable to students' misconstruction of the mole as a mass or number particular to the property of a substance such as atomic mass. From the perspective of students' solution processes, a tendency toward avoidance of the mole was observed amongst students and was thought to stem from this lack of conceptual understanding of terms in the context of students' solution processes (Schmidt, 1994). Instead, students were often observed to rely on proportional reasoning techniques (*e.g.*, cross-multiplication) to replace methods modeled step-by-step by their instructors suggesting students lean on their mathematical skill sets to compensate for a lack of understanding of calculations in the context of chemistry.

Ultimately, the themes described in the research literature concerning student challenges with the mole concept and stoichiometry cited above can be summarized as related to (1) representational competence, (2) systematic approaches, and (3) conceptual understanding. Each theme is summarized succinctly below.

(1) Representational competence: interpreting elemental symbols and chemical formulae or reactions to identify proportionality between chemical substances and inherent physical properties relevant to the problem-solving process.

(2) Systematic approaches: an algorithmic or step-by-step approach resulting in a solution that presents coherent units of measurement.

(3) Conceptual understanding: the underlying terminology and theories of science used to reason a viable solution deductively.

This review of the literature describing how low-performing chemistry students, and chemistry students as a whole, struggle with the mole concept and stoichiometry was used to develop a conceptual framework as to how these challenges may manifest and evaluate whether or not these challenges are a description of those observed amongst at-risk chemistry students.

### Purpose

Chemistry courses serve as a primary conduit in the STEM pipeline, and remediation of the differential outcomes observed for these students could serve to diversify the population of students who emerge successfully (Lee *et al.*, 2018). There is currently a lack of evidence on the topic matter, assessment strategies, and experiences contributory to the differential performance observed amongst chemistry students with low math aptitude scores, despite a sizable literature base relating math aptitude to chemistry performance suggesting this differential performance remains widespread (Coley, 1973; Pickering, 1975; Andrews and Andrews, 1979; Ozsogomonyan and Loftus, 1979; Rixse and Pickering, 1985; Spencer, 1996; Wagner *et al.*, 2002; Lewis and Lewis, 2007; Hall *et al.*, 2014; Ye *et al.*, 2016). Further, given the nearly ubiquitous adoption of compulsory prerequisite courses in mathematics before enrollment in college-level chemistry (Bialek and Botstein, 2004) it is clear that instructors and administrators are rightfully concerned about differential performance and would benefit from an expansion of the relevant evidence base.

This study sought to analyze the solution processes of 348 chemistry students (including 101 students with relatively low math aptitude scores) to mole concept and stoichiometry assessment items where differential performance has been previously identified at the research setting. The results can inform instructional and assessment practices while providing a greater insight as to the skill most strongly predicted by chemistry students' precollege math aptitude scores. Understanding how these students experience disproportionate difficulty in chemistry has the potential to inform practices that serve to ameliorate gaps in performance and support more equitable, evidence-based practices for intervention, instruction, and assessment. Without this knowledge, current practices may unintentionally contribute to the propagation of these inequitable outcomes in first-semester general chemistry.

## Research objectives

Two primary research objectives guided the development of this study:

(1) Describe the challenges observed amongst at-risk chemistry students in solution processes concerning the mole concept and stoichiometry assessment items.

(2) Compare and contrast the challenges that present amongst not-at-risk and at-risk chemistry students to describe potential causes for the differential performance observed.

## Methods

### Research setting

This study was conducted at a large, public research institution in the southeastern United States. Data collection occurred during a semester where class sizes ranged from 153 to 217 students for four classes of off-sequence (spring term), first-semester general chemistry. Classes were coordinated across four instructors with a shared textbook, learning objectives, syllabus, grading scheme, and online learning management platform. Students attended regular lecture sessions twice per week and a peer-led problem-solving session once per week

 

(Gosser *et al.*, 2006; Lewis, 2011). The textbook used for the classes was "Chemistry: A Molecular Approach" (Tro, 2014).

Students' grades were comprised of three interim exams (45% of total grade, 15% each exam), a final exam (25% of final grade), and participation driven grading systems (*e.g.*, clickers, participating in peer-led sessions, and completing online homework) for the remaining 30%. Exams were common across all classes and written by a committee of the students' instructors. Interim exams consisted of 20 multiple-choice assessment items with four distractors (five answer choices in total) and a series of six true-or-false items following the Measure of Linked Concepts format to emphasize the links across topics in the course (Ye *et al.*, 2015). Each multiple-choice item on the tests was worth seven points, and each true-or-false item was worth three points for a correct response or one point for selecting unsure (in an attempt to reduce chance guessing). The final, cumulative exam followed a similar format with 45 multiple-choice and ten true-or-false assessment items.

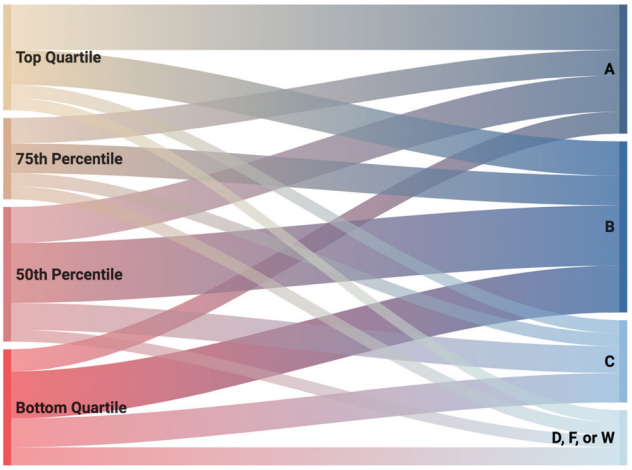


Fig. 1 Sankey diagram where the number of students within each quartile of math aptitude (left) are represented by the width of the bands connecting students to the final grades achieved in the course (right).

### Participants

Seven hundred fifty-five students were enrolled in the course during the study. Of these students, 634 were enrolled with a registered math aptitude score ranging from 440 to 800 with a mean score of 561. In 2018, the CollegeBoard reported a mean score of 531 on the math composite of the SAT (CollegeBoard, 2018). Five hundred forty-three students with registered math aptitude scores completed the cumulative final exam with a mean score of 70.0% and a standard deviation of 16.6%. Students who completed the final exam, were enrolled in the course with a registered math aptitude score and responded to each of the surveys following their interim exams ($n$ = 348) were observed as having a mean score of 66.1% on the final exam with a standard deviation of 16.1%. Names following student responses to these surveys as presented in this work are pseudonyms generated to protect the identity of the students and neither relate to nor assume each student's gender identity, race, or ethnicity. The university's Institutional Review Board approved the collection of data following institutional and local guidelines.

### Academic performance of students scoring at or below the 25th percentile on math aptitude scores

The relationship between math aptitude (by quartile) and final grade in the course was considered to determine the academic outcomes achieved by chemistry students of variable math aptitude. In the research setting, students could achieve an A, B, C, D, or F in grade or a W for students who chose to withdraw from the course. Adverse academic outcomes are considered as D, F, or W, as each of these grades prevents the student from enrolling in subsequent coursework until the course is retaken. The Sankey diagram in Fig. 1 depicts a visual used to evaluate the academic outcomes of students by quartile of math aptitude where the width of the connections between math aptitude quartile and the final grade achieved is equal to the frequency of students that match the presented intersectionality.

If math aptitude and academic chemistry performance were unrelated, each quartile of math aptitude would achieve one-quarter of each letter grade. Of the 634 students with math aptitude scores, 168 students (26.5%) were in the bottom quartile. The bottom-quartile students received 32.5% of the D, F, or W grades (including 31% of those who withdrew from the course) and 35.3% of the C grades but only 17.1% of the A grades. This data suggests that the previous findings in the research literature describing math aptitude as related to academic chemistry performance also holds in the current setting with the bottom quartile of math aptitude at a higher risk of not succeeding. To be clear, we are not proposing that chemistry performance and mathematics proficiency are unrelated, as multiple chemistry topics require quantification and mathematic manipulation. Instead, we hypothesize from the reviewed literature that students enrolled in general chemistry at the research setting have the requisite algebraic manipulation skills needed for general chemistry. Part of the goal of this work is to explore the feasibility of this hypothesis. If the hypothesis holds true, the math component of the SAT measures math skills in addition to what is needed in general chemistry. If so the relationship between math SAT and general chemistry performance demonstrated in Fig. 1 can be better understood and mitigated, which would reduce student attrition in the course and create equality of outcomes for students regardless of math aptitude (Lynch, 2000, p. 13). Alternatively, if the hypothesis fails and algebraic manipulation skills manifest as the reason for student struggles this would call for additional preparation aimed at developing this skill set.

### Development and design of post-assessment survey prompts

Following the finding that at-risk chemistry students at the research setting struggled most differentially on mole concept and stoichiometry assessment items (Ralph and Lewis, 2018), test items related to these topics were selected from those drafted by a committee of the course instructors. Items were

selected *a priori* (before the information was available as to students' performance on the item). Each assessment item was expected to require a mole-to-mole conversion in different contexts: mole concept, theoretical yield, gas laws, and changes in energy.

Survey prompts were administered *via* students' online learning management system. Following each of the three interim exams, students were asked to respond to open-ended prompts designed to elicit their solution processes within a week of their in-course interim chemistry tests. Surveys were administered following interim exams to receive students' solution processes after students prepared for, completed, and had access to the answer key for the assessment items presented in the surveys.

The first two prompts followed the first interim test, and the last two followed the second. No topics were identified as involving the mole concept or stoichiometry on the third interim test. Students were able to freely edit their answers on the survey until the established deadline. The survey prompts followed a consistent format as presented below.

### The mole concept item

For the test question (below), please explain how to arrive at the correct answer (''E'').

*How many moles of oxygen are in 15 g of $K_2SO_4$?*

*(A) 0.022 mol*

*(B) 0.043 mol*

*(C) 0.086 mol*

*(D) 0.17 mol*

***(E) 0.34 mol***

### The stoichiometry item

For the same test question (below), please explain how to arrive at the correct answer (''C'').

*Calculate the amount of $NH_3$ gas produced by reacting 45.6 g of $H_2$ gas with an excess of $N_2$ gas according to the following chemical equation.*

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

*(A) 85.6 g (B) 171 g (C)* ***257 g*** *(D) 385 g (E) 579 g*

### The gas laws item

For the same test question (below), please explain how to arrive at the correct answer (''A'').

*If 126 L of $H_2(g)$ at 15 °C and a constant pressure of 0.920 atm reacts with excess $N_2(g)$ as shown below, how many moles of $NH_3(g)$ are produced?*

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

***(A) 3.27 mol $NH_3$***

*(B) 4.90 mol $NH_3$*

*(C) 7.36 mol $NH_3$*

*(D) 55.7 mol $NH_3$*

*(E) 84.0 mol $NH_3$*

### The changes in energy item

For the same test question (below), please explain how to arrive at the correct answer (''B'').

*Consider the following balanced chemical equation:*

*$4NH_3(g) + 5O_2(g) \rightarrow 4NO(g) + 6H_2O(l) \quad \Delta H_{rxn} = +1168$ kJ*

*How much heat is absorbed/released when 2.05 mol of $NH_3(g)$ reacts with excess $O_2(g)$ to produce $NO(g)$ and $H_2O(l)$?*

*(A) 5.99 × $10^2$ kJ of heat is released.*

***(B) 5.99 × $10^2$ kJ of heat is absorbed.***

*(C) 2.40 × $10^3$ kJ of heat is released.*

*(D) 2.40 × $10^3$ kJ of heat is absorbed.*

*(E) 1.02 × $10^4$ kJ of heat is absorbed.*

Students received 0.5% added to their final grades for the completion of each of the three surveys. Students' were also informed ''There are no correct answers to the survey, but repetitive or non-responsive answers may not result in extra points'' regarding the expectations for their receipt of these extra credit points.

### Coding students' responses

Responses were first collected into a database using the export function of the learning management system, merged with roster and assessment data files using SPSS (IBM Corp., 2017). Then, responses were imported to the qualitative analysis data software MAXQDA (VERBI Software, 2017) as a means to catalog and organize codes, memos, and logbook (comprised of a dated and timed series of entries wherein the researchers reflect on the data and the process by which the data has been analyzed) while interacting with the data. The first author devised a deductive coding scheme using themes observed from the literature base cited in the introduction. These works served as a conceptual framework concerning why students (in general) tend to struggle with the mole concept and stoichiometry assessment items. From this framework, a coding scheme was developed to describe the errors students make in their solution processes for the assessment items selected for the survey (see Table 1).

When challenges emerged beyond those described by the conceptual framework, the incorrect processes were used to generate one-to-four-word descriptions summarizing the challenges demonstrated by the students who authored them. Similar codes were then combined *via* pattern coding (Miles *et al.*, 2014) to group these summaries, assess their commonality, and consider each in the context of the conceptual framework in concerning the difficulties observed in students struggling with the mole concept and stoichiometry. Student solutions were coded so that one student's response could have multiple codes applied. This iterative process started with both authors coding common subsets of the data, discussing differences in the manner by which codes were applied, and altering the codebook to improve parsimonious and consistent coding of students' responses. Subsets of the data used for interrater were selected from inaccurate student responses of which there were 340 responses from the 247 not-at-risk (NAR) students and 199 responses from the 101 at-risk (AR) students. Coding rounds occurred by item (mole concept, stoichiometry, gas laws, and changes in energy) where rounds were randomized and selected so that the number of inaccurate responses from the NAR and AR cohorts was equal. For example, the coding round wherein codes (see Table 2) were applied to students' response processes for the mole concept item included 62 (31 from each group) of the 138 (or 45%) inaccurate responses. The researchers coded these items

 

Table 1 *A priori* coding scheme comprised of the challenges by which students have been observed to struggle with the presented mole concept and stoichiometry assessment items in past works

| Code | Code description |
|---|---|
| **Representational competence** | |
| Atomic *vs.* molecular | Atoms comprising species are misrepresented as covalent compounds using subscripts (*e.g.*, 4 O in potassium sulfate *vs.* $O_4$) |
| Misrepresented proportions | Proportions identified by the subscripts and/or coefficients of a chemical formula are misrepresented |
| Elemental symbols | Atomic masses are incorrectly assigned to elemental symbols as listed on the periodic table |
| Sign conventions | The meaning of + and − signs in the context of scientific notation or interpreting energy transfer is reversed |
| Nomenclature | Chemical symbols are misinterpreted when determining the predominate state of matter or name of a chemical species |
| **Systematic approaches** | |
| Inaccurate conversion factors | Selected conversion factors between units of measure (*e.g.*, mass to moles; moles to mass) are incapable of resulting in a viable measure for the converted value (*e.g.*, converting mass to moles using Avogadro's number) |
| Conversion arrangement | Numbers and units are arranged improperly result in units incapable of achieving a coherent number and unit of measure (*e.g.*, 12.1 moles O multiplied by 1 mol O/16.0 g O to achieve mass of O) |
| Sequence | Steps in the predicted solution path(s) are present but are rearranged in a manner that does not reflect a solution path that is chemically plausible |
| **Conceptual understanding** | |
| Interchemical unit identity | Moles of a compound are not presented as distinct from moles of a constituent element |
| Unit conflation | Differing units of measure are used interchangeably (*e.g.*, molar mass as mass) |
| Intrachemical unit identity | Terminologies inherent to particular units are misapplied (*e.g.*, using the term ''the number of atoms'' of a chemical to describe the moles of the same chemical) |

independently using a spreadsheet to record their applications of the codes, and later met to discuss the discrepancies and make amendments to the codes to promote greater clarity in their meaning and applicability. This cycle continued for each assessment item until rounds of coding emerged with few disagreements, each discussed by the authors until an agreement was achieved. In total, 539 inaccurate responses were collected and 196 (36.4%) were coded in rounds by both researchers until the discrepancies were rare enough to suggest the coding scheme was reliable for application to the remaining 343 inaccurate responses.

This methodology was regularly evaluated by the researchers for trustworthiness (Lincoln and Guba, 1986) using frameworks presented in the literature base for evaluating the quality of qualitative education research (Peshkin, 1993; Spencer *et al.*, 2003). Primary considerations included the alignment of the methodology with the research questions and previous literature, the execution of data collection and analysis, and the contributions these data and the presentation of relevant findings may provide to the literature. As a last effort to ensure the trustworthiness of the findings presented, the findings were shared with two undergraduate researchers (who were not participants in the study). These students demonstrated strong academic performance in their experiences with general chemistry, served as peer-leaders and tutors at the research institution and were able to provide a student's perspective for some of the challenges observed. Undergraduate researchers were given an overview of the methodology and deidentified samples of the dataset to summarize (these samples consisted of pastiches presenting three or more student responses grouped by code for each assessment item). Then, the students were given the results and discussion sections presented below and were asked to interject, argue, and debate their perspectives of the data where alignment was not achieved. The students were influential in this process and identified nuanced additions, contrapositives, and students' perspectives of the data and implications for assessment design.

# Results

The results presented below consist of students' assessment response data and solution response processes related to the topics of mole concept, stoichiometry, and applied stoichiometry (*e.g.*, stoichiometry used in the context of the gas laws). By-item statistics include the number of students who responded to the item ($n$), the percent by which students overall ($P$) selected the correct answer, and the standard deviations (SD) observed within each group demarcated by at-risk (AR) and not-at-risk (NAR). Below item statistics, the rationales for each of the distractors and the frequency of student selection ($n$ and %) are provided for each group. The mean differential between NAR and AR was calculated as the mean percentage of students in the not-at-risk student cohort subtracted by that of the at-risk cohort (MD = NAR − AR). These statistics are then contrasted with qualitative descriptions of students' challenges supported with participant-voiced pastiches comprised of exemplary student responses for each code color-coded *via* figure legend. Following the description of prevalent themes in the challenges observed amongst at-risk chemistry students, a qualitative comparison of those observed amongst the not-at-risk cohort is described.

## Challenges of at-risk chemistry students

**Interchangeable chemical identity.** At-risk chemistry students were commonly observed to interchange the chemical identity of a numerical value between different chemical

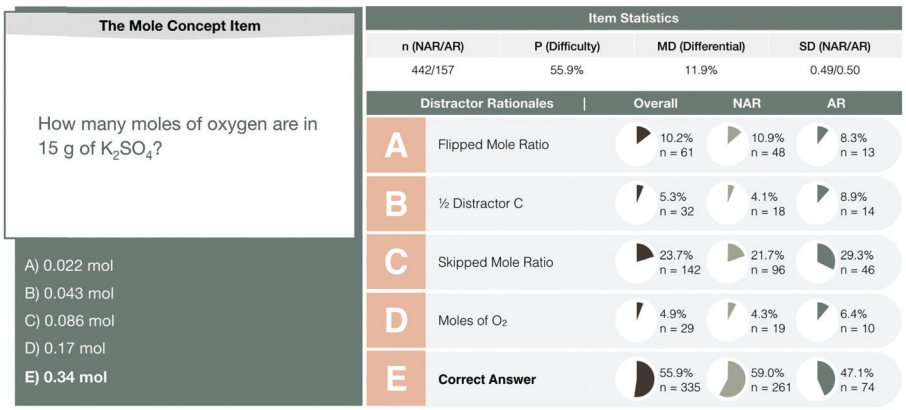


Fig. 2 The mole concept interim assessment item profile comparing the performance of not-at-risk (NAR) and at-risk (AR) chemistry students.



species sometimes *via* common units such as mass or moles. An error of this quality was coded as "interchangeable chemical identity." Take, for example, responses to the mole concept assessment item in which students were prompted to calculate the moles of oxygen in 15 grams of $K_2SO_4$ (see Fig. 2).

Two accurate solution processes to the mole concept item were described by students as depicted below.

The mole ratio process

$$(15\text{ g }K_2SO_4)\left(\frac{1\text{ mol }K_2SO_4}{174.27\text{ g }K_2SO_4}\right)\left(\frac{4\text{ mol O}}{1\text{ mol }K_2SO_4}\right) = 0.34\text{ mol O}$$

The mass percent process

$$(15\text{ g }K_2SO_4)\left(\frac{64.0\text{ g O}}{174.27\text{ g }K_2SO_4}\right)\left(\frac{1\text{ mol O}}{16.00\text{ g O}}\right) = 0.34\text{ mol O}$$

A presentation of interchangeable chemical identities was observed by students who divided the provided mass of $K_2SO_4$—15 grams—by the molar mass of oxygen.

> "Well, you'd set 15 grams of $K_2SO_4$ to proper convertion [*sic*], **with the molar mass of oxygen on the bottom**, set over the continuing equation. . ." (Erin)

Erin describes a conversion between 15 grams of potassium sulfate and "the molar mass of oxygen". One plausible interpretation of this solution process is displayed below.

$$(15\text{ g }K_2SO_4)\left(\frac{1\text{ mol O}}{16\text{ g O}}\right)$$

Whether Erin chose the atomic mass of oxygen or the molecular mass of oxygen (an error of consequence to be discussed), there was interchangeability in the chemical identity of the given mass and the atomic or molar mass included in the solution process. A student could perceive this process as resulting in the cancellation of grams to arrive at moles of oxygen *via* dimensional analysis. A variety of responses were observed to share this interchangeability (see Fig. 3) involving the mass or molar mass of oxygen applied to the mass or moles of potassium sulfate in solution processes. Much of the differential performance observed in Fig. 2, was attributed to the "skipped mole ratio" option (choice C) which matches the answer choice predicted by students who exemplified interchangeable chemical identity by using the molar mass of oxygen with the mass of potassium sulfate.

At-risk chemistry students were also observed to describe calculating the mass percent of oxygen in potassium sulfate and misattribute this value with the moles of oxygen in potassium sulfate as the two were near in numerical value for the assessment item (0.37 or 37% oxygen by mass and 0.34 mol of oxygen in potassium sulfate). Students' responses shown in Fig. 4 describe this inaccurate application of the mass percent algorithm and resulted in students either selecting the correct answer on the exam or presenting this process in their survey responses as what should have been enacted.

Interchangeable chemical identity was characteristic of solution processes for all four assessment items and indicates a characteristic challenge in a student's conceptual understanding of the meaning behind the numbers and units of

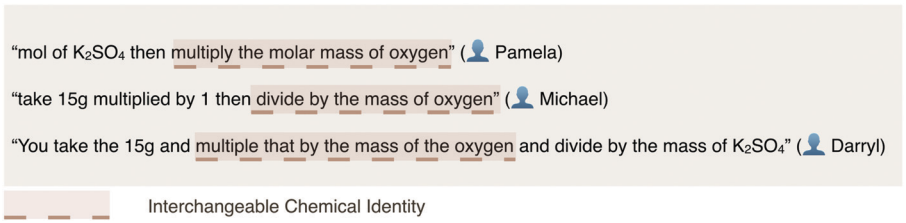


Fig. 3 Varieties by which students communicated interchangeability between numbers sharing the unit of mass but differing in chemical identity.

 

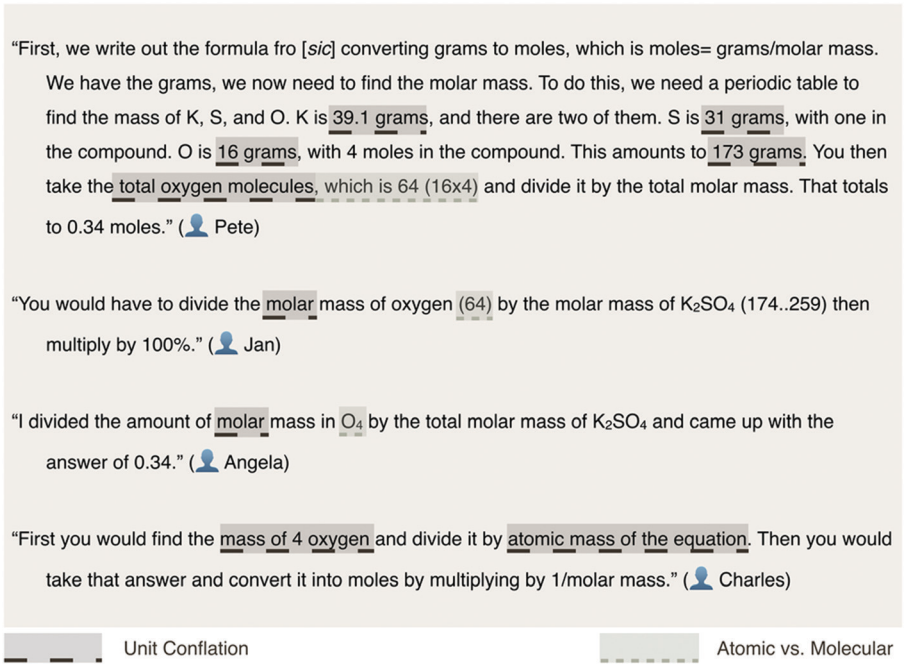


Fig. 4 Misapplying mass percentages to the moles or molar masses of oxygen.

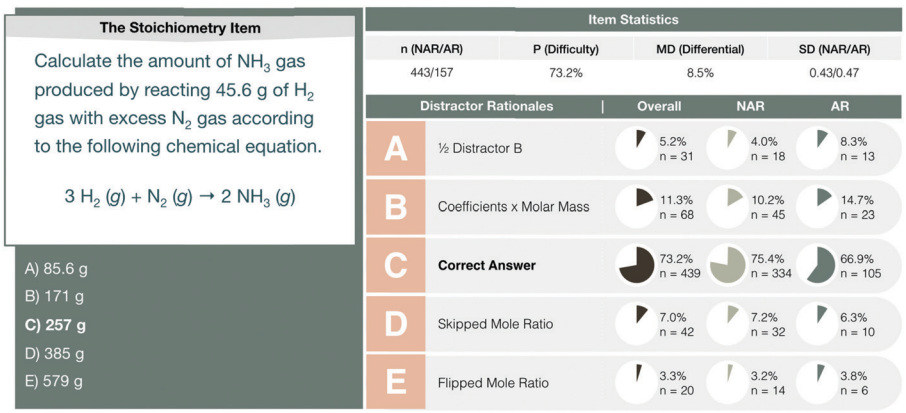


Fig. 5 The stoichiometry interim assessment item profile comparing the performance of not-at-risk (NAR) and at-risk (AR) chemistry students.

a solution process. In the stoichiometry assessment item, students were tasked with calculating the mass of $NH_3$ produced from 45.6 g of $H_2$ and excess $N_2$ (see Fig. 5).

A predicted solution path for the item could be described as follows.

$$(45.6\,\mathrm{g\,H_2})\left(\frac{1\,\mathrm{mol\,H_2}}{2.02\,\mathrm{g\,H_2}}\right)\left(\frac{2\,\mathrm{mol\,NH_3}}{3\,\mathrm{mol\,H_2}}\right)\left(\frac{17.03\,\mathrm{g\,NH_3}}{1\,\mathrm{mol\,NH_3}}\right)=257\,\mathrm{g\,NH_3}$$

Students' processes involved the interchangeability of chemical identities sharing a common unit of measure (often mass, molar mass, or moles) as observed in their responses to other assessment items.

"Convert 45.6 g [*of* $H_2$] givens [*sic*] to moles by dividing it by the total mass of $NH_2$ [*sic*]. Then multiply the moles found by the 2 that are in hydrogen, and lastly convert your new moles back to 'g'" (Kevin)

$$(45.6\ \mathrm{g\ H_2})\left(\frac{1\ \mathrm{mol\ NH_2}}{\mathrm{g\ NH_2}}\right)(2\ \mathrm{mol\ H_2})$$

Here the interchangeability between $H_2$ and "$NH_2$" seems reinforced by both the cancellation of grams and the quality of $NH_2$ as containing molecular hydrogen. Kevin's response was coded as both interchangeable chemical identity and an error in the interpretation of chemical symbolism.

**Atomic *vs.* molecular.** The code labeled "atomic *vs.* molecular" was most pronounced where the comprising elements of a compound were capable of forming a diatomic molecule and were involved in a component of students' solution processes.

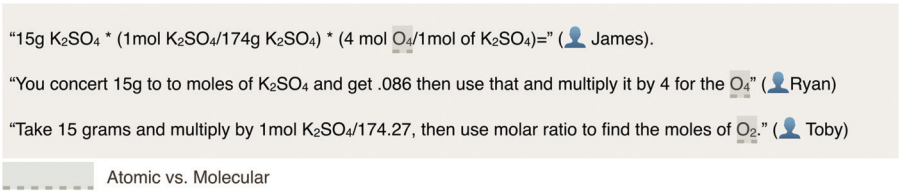


Fig. 6 Returning to the mole concept assessment item, the code "atomic *vs.* molecular" was prevalent.



In Fig. 6, responses to the mole concept assessment item reflect this challenge with representational competence.

Here, students tend to represent 4 mol of oxygen atoms in potassium sulfate as either $O_4$ or $O_2$. One plausible interpretation of these errors in the solution process could concern the stem of this item. Students could have misinterpreted "moles of oxygen" with "moles of molecular oxygen" and thereby selected "D" as a plausible response to the item. Only twenty-nine students (4.8%) selected answer choice "D" discounting this as a common interpretation of the question but the question stem would be improved by specifying "oxygen atoms". As shown in Fig. 6, students commonly misrepresented oxygen as $O_4$. James and Ryan still utilized a coefficient of 4 with $O_4$ and thus this misconception did not prevent their selecting the correct answer to the assessment item. Inaccurate solution processes that arrive at correct answer choices could reinforce students' misconceptions and students' future interpretations of chemical symbolism could be negatively impacted as a result.

**The role of molar mass.** Save for the mole concept assessment item, all other stoichiometry-related assessment items involve the application of stoichiometric coefficients interpreted from a balanced chemical reaction to solve an assessment item. For the stoichiometry assessment item in Fig. 5, students struggled in their attempts to discern the masses, moles, molar masses, and stoichiometric proportions between chemical species and the roles these values serve within systematic approaches. One manner by which this challenge manifested was with confusion related to the use of coefficients from a balanced chemical reaction in the determination of molar masses.

As observed in all three student elaborations presented in Fig. 7, the role of a coefficient was unclear. Reflected by Stanley and Meredith's responses, students were observed to apply stoichiometric ratios to their calculations of the molar mass of a compound. At-risk students manifested this misconception as to the utility of coefficients and were observed to achieve the correct answer *via* an inaccurate conception of the algorithm. There were other solution processes, however, wherein at-risk students applied stoichiometric coefficients to molar masses in a manner that did not result in the attainment of a correct answer (see Holly's response). The stoichiometry assessment item presented with a low mean differential (MD = 8.5%) and was observed to attract 11.3% of the students to answer choice B in Fig. 5. The rationale of this distractor reflected students' use of coefficients in the balanced chemical reaction to calculate the molar mass of a substance (*e.g.*, $H_2$ would have, not 2.016 g $mol^{-1}$, but 6.048 g $mol^{-1}$ following the multiplication of the molar mass by a coefficient of 3 from the balanced chemical reaction). This distractor was responsible for 4.5% of the overall 8.5% mean differential observed.

Challenges with the interpretation and use of molar mass were also observed throughout the assessment items on the second

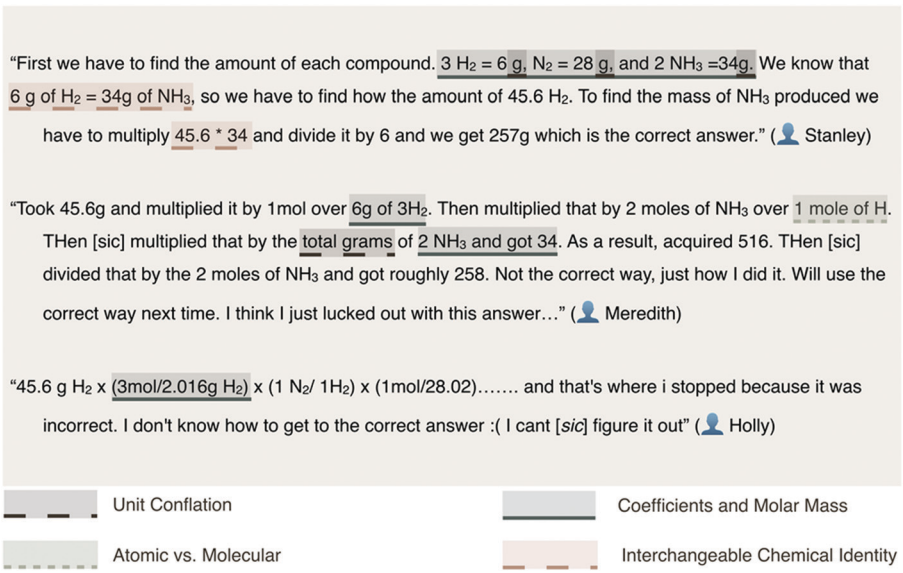


Fig. 7 Students were applying coefficients representing stoichiometric proportions from balanced chemical reactions to the molar masses of chemicals.

 

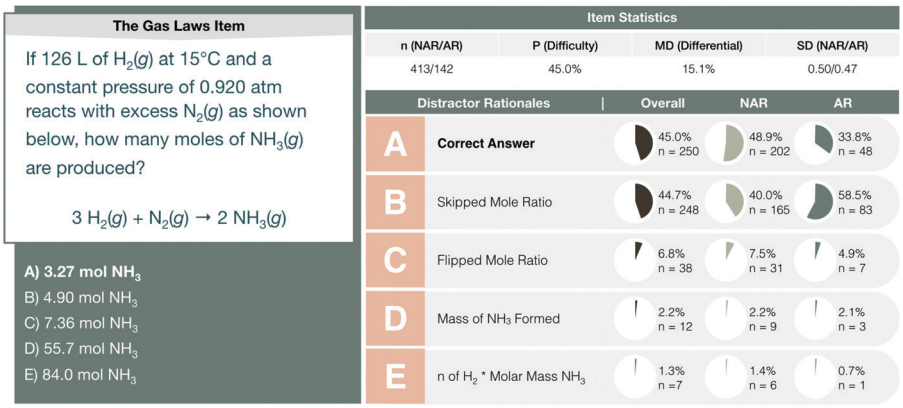


Fig. 8 The gas laws interim assessment item profile comparing the performance of not-at-risk (NAR) and at-risk (AR) chemistry students.



interim exam. The gas law item provides the values necessary to compute moles of molecular hydrogen gas using the ideal gas law and asks students to determine the moles of ammonia gas produced given excess molecular nitrogen gas (see Fig. 8).

The provided chemical reaction was the same as that of the stoichiometry item from the first interim test where a predicted solution path for the item could be depicted as follows.

$$n_{H_2} = \frac{(0.920\ \text{atm})(126\ \text{L})}{(0.08206\ \text{L atm mol}^{-1}\ \text{K}^{-1})(15 + 273.15\ \text{K})}$$

$$n_{NH_3} = (4.90\ \text{mol H}_2)\left(\frac{2\ \text{mol NH}_3}{3\ \text{mol H}_2}\right) = 3.27\ \text{mol NH}_3$$

Students were observed to calculate the moles of $H_2$ with little difficulty. Their ability to algebraically manipulate the ideal gas law equation for moles appeared unencumbered by having scored in the bottom quartile of math aptitude scores. Challenges arose when converting the moles of $H_2$ gas calculated to moles of $NH_3$. Often, this difficulty involved determining whether molar mass plays a role in discerning a final answer as demonstrated in Fig. 9.

In parallel, the changes in energy assessment item is predicted to have a solution path that does not inherently require the use of the molar mass for ammonia, yet students demonstrate the use of this value as a prominent point-of-confusion. For this item, students are provided the following chemical reaction and heat of reaction (see Fig. 10).

Students were asked to calculate the amount of heat absorbed or released when 2.05 mol of $NH_3(g)$ reacts with excess $O_2(g)$ to produce $NO(g)$ and $H_2O(l)$. A theoretical solution path to the item is depicted below.

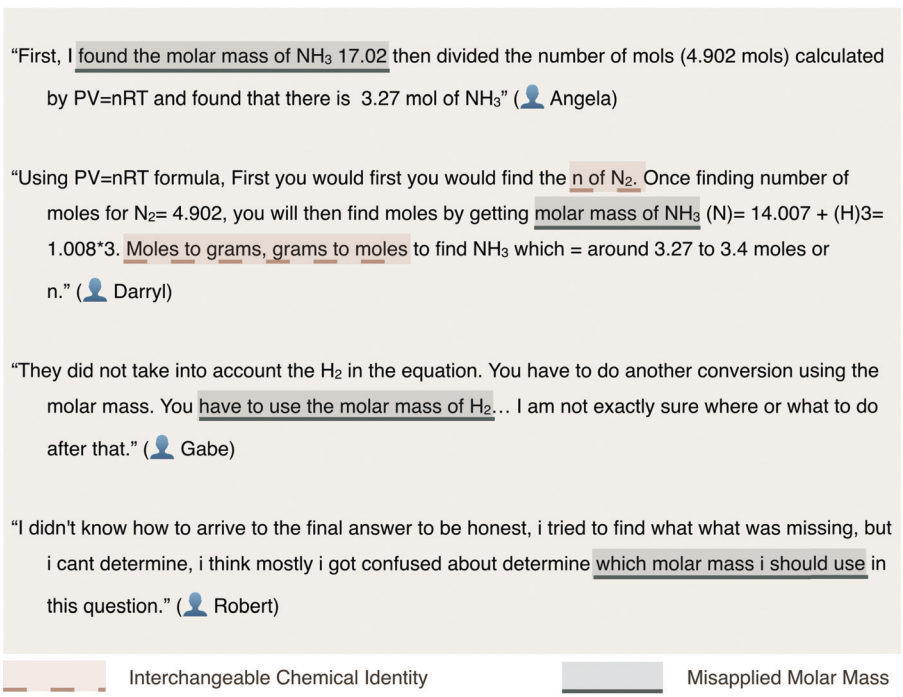


Fig. 9 Misapplications of molar mass in response to the gas law assessment item.

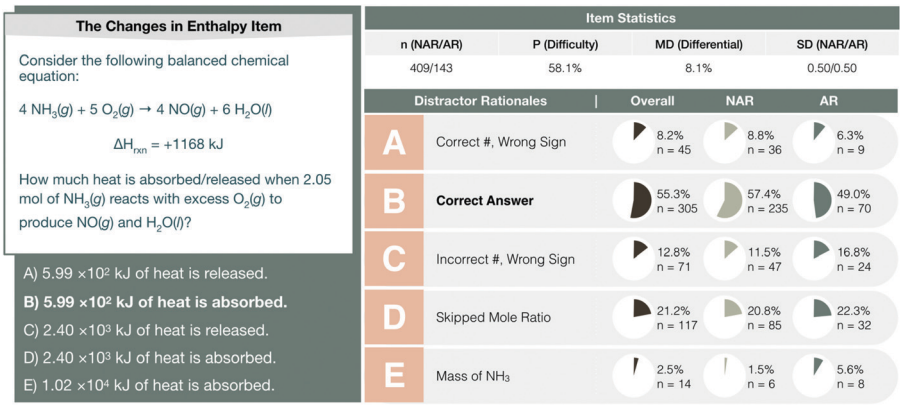


Fig. 10 The changes in energy interim assessment item profile comparing the performance of not-at-risk (NAR) and at-risk (AR) chemistry students.



$$(2.05\ \text{mol NH}_3)\left(\frac{+1168\ \text{kJ}}{4\ \text{mol NH}_3}\right) = +559\ \text{kJ}$$

Here, confusions over the application of molar mass arose once again.

Solution processes similar to that of Pete and Angela (see Fig. 11) would result in an accurate response *via* an inaccurate conception of the process. Our interpretation of Angela's solution process is illustrated below and may demonstrate the degree to which students rely on algorithmic solution processes without conceptual understanding.

$$2.05\ \text{mol NH}_3\left(\frac{17.03\ \text{g NH}_3}{1\ \text{mol NH}_3}\right) = 34.9\ \text{g NH}_3$$

$$\frac{34.9\ \text{g NH}_3}{\left(4\ \text{mol NH}_3 \times \frac{17.03\ \text{g NH}_3}{1\ \text{mol NH}_3}\right)} = 0.512\ \text{mol NH}_3$$

$$0.512\ \text{mol NH}_3\left(\frac{1168\ \text{kJ}}{1\ \text{mol NH}_3}\right) = 5.99 \times 10^2\ \text{kJ}$$

Angela's application of the molar mass of ammonia twice was, presumably, to execute an algorithm by which the mass of

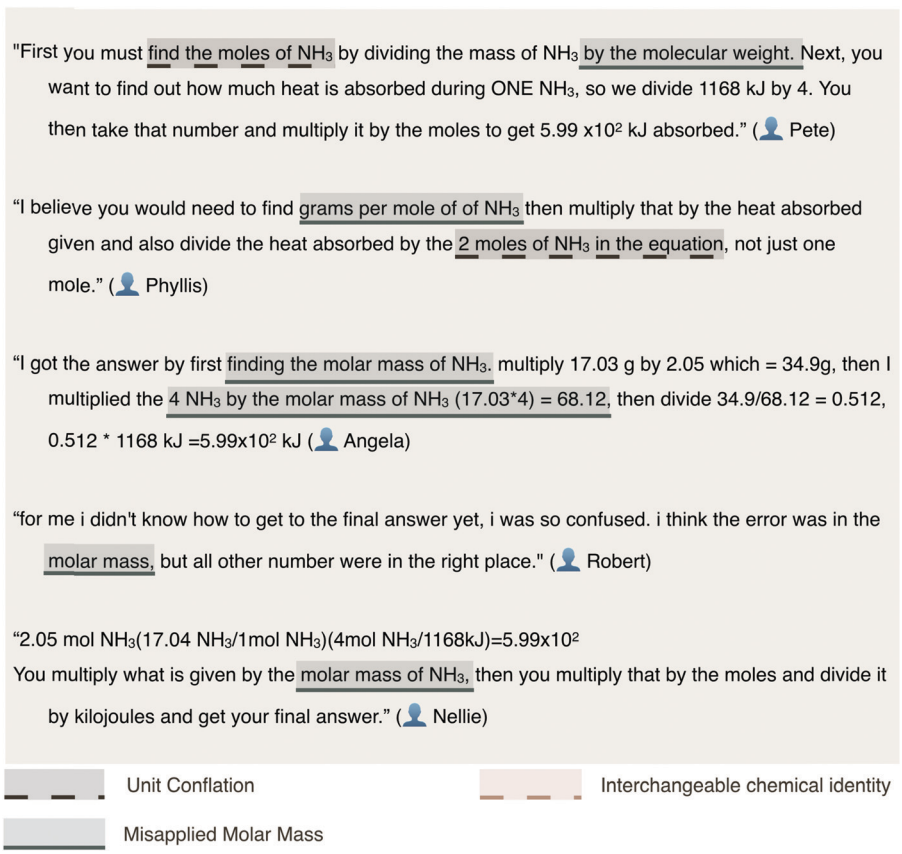


Fig. 11 Misapplications of molar mass in response to the changes in energy assessment item.

 

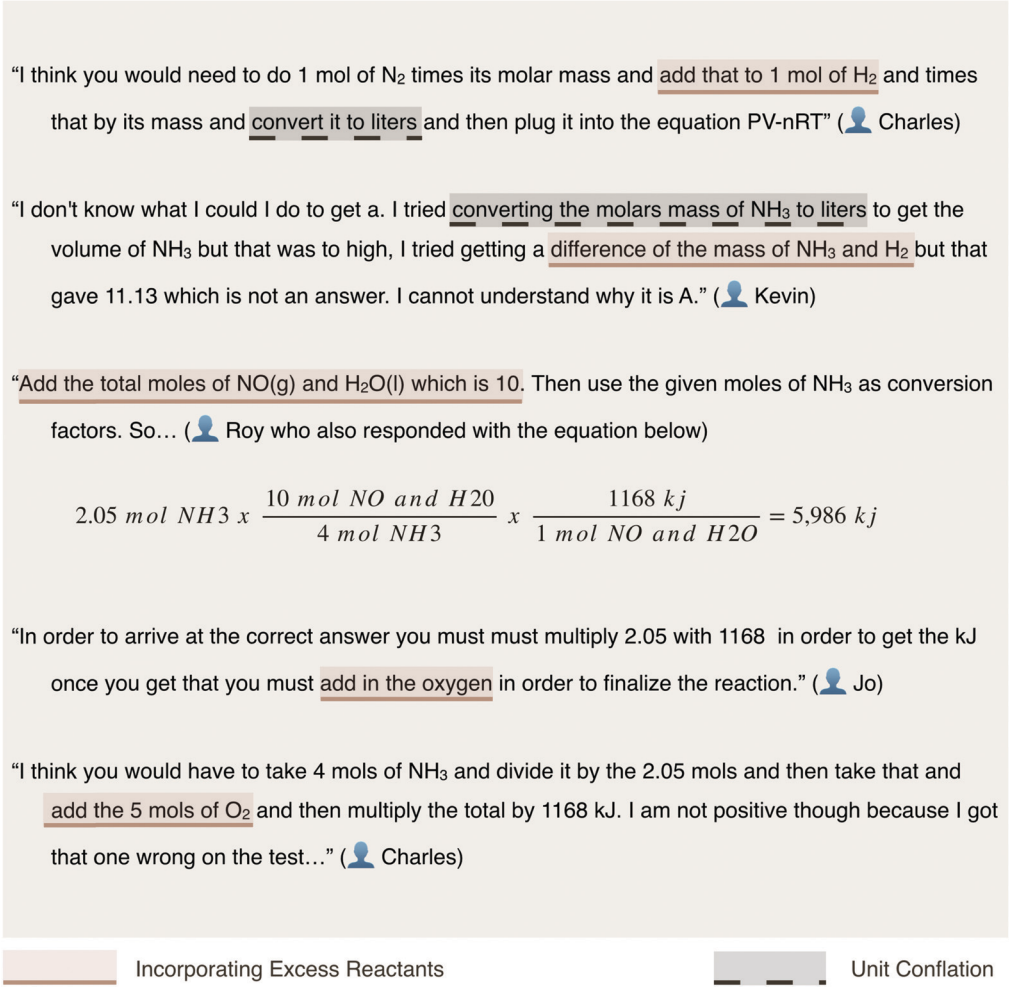


Fig. 12 Interchangeable chemical identities manifested *via* uncharacteristic solution processes involving excess reactant to the gas laws (first two responses) and changes in energy (last three responses) assessment items.

$NH_3$ was required. This difficulty in determining whether or not to use the molar mass may be indicative of the difficulty associated with distinguishing moles of one chemical species from moles of another. Alternatively, the challenge could be related to the use of the moles of $NH_3$ to reconcile the connection between energy and mass or moles. The persistence of this challenge was commonly observed (see responses authored by Angela and Robert in Fig. 9 and 11). These difficulties were most notable in the changes in energy assessment item as students struggled to differentiate between the stoichiometric coefficient (*e.g.*, the 4 mol of $NH_3$ required for the balanced chemical reaction) and the amount in moles of ammonia consumed in the chemical reaction (see responses by Pete and Phyllis in Fig. 11).

**"In excess" as a source of confusion.** Inaccurate solution processes of at-risk chemistry students for the applied stoichiometry assessment items also involved the incorporation of chemicals prompted as "in excess" suggesting that this phrase did not inform students as to whether a theoretical yield or a discerning the limiting reactant algorithm should be executed. Exemplar responses demonstrate how students invoke algorithms that seem to emphasize the conservation of matter using excess reactants (Fig. 12) or include excess reactants in the determination of theoretical yield (Fig. 13).

Charles, an at-risk student whose solution processes are shown for both the gas laws and changes in energy items in Fig. 12, demonstrates persistent difficulty with discerning whether or not to implement excess reactants toward arriving at an answer. "In excess" is a phrase commonly observed amongst the assessment items related to stoichiometry. The phrase is intended to communicate an abundance of a reactant suggesting one need not consider the reactant present in excess as limiting to the theoretical yield of a chemical reaction. As none of the four assessment items require students to discern the limiting reactant in a calculation of a produced mass, challenges with limiting reactants were not expected to emerge within students' responses. Yet, these challenges were prevalent amongst the stoichiometry, the gas laws, and changes in energy assessment items. Often, students included the excess reactant in their calculations. Manifestations of these errors are presented in each pair of exemplary student responses in Fig. 13

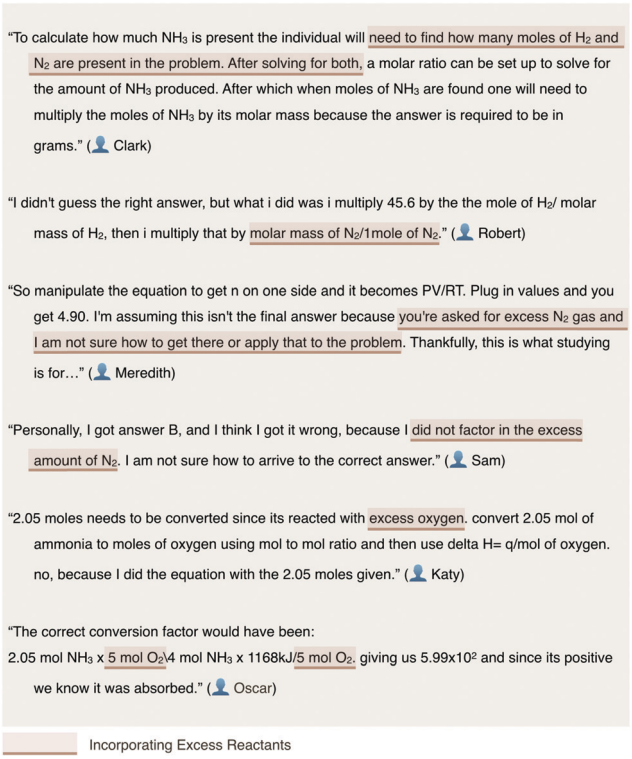


Fig. 13 Incorporating a limiting reactants algorithm for stoichiometry (Clark and Robert), the gas laws (Meredith and Sam), and changes in energy (Katy and Oscar) assessment items describing all other reactants as in excess.

representative of the stoichiometry, gas laws, and changes in energy assessment items, respectively.

Each response demonstrates confusion as to how to incorporate the ''excess'' reactant, which was commonly observed throughout the students' elaborations. If the instructors intend to measure proficiency with associating an abundant reactant as having no impact on the theoretical yield of a chemical reaction in addition to solving gas law stoichiometry or energy changes of a chemical reaction, then these assessment items are well placed. It may be more plausible that the intent was to include ''excess'' as a means of eliminating the additional challenge of discerning limiting reactants within the context of the stoichiometric calculations. This simplification does not seem to correspond with students' elaborations of the item in which they indicate confusion over the phrase and attempt incorporation within the algorithms by which problems are solved.

### A comparison of the challenges observed between not and at-risk chemistry students

In evaluating potential sources for differential performance, the prevalence of challenges with these items were considered with respect to each code (see Table 2).

Overall, the most common challenges observed involve interchangeable chemical identity, unit conflation, and incorporating excess reactants. Of the students struggling with these concepts, the challenges and the prevalence by which they affect students does not seem to explain the differential performance observed. In the process of qualitatively reviewing student responses, it was clear that some students were able to arrive at a correct answer for the prompt *via* a chemically implausible solution process. For example, the exemplar responses of at-risk students in Fig. 4 demonstrate an error wherein students equate the moles of oxygen with the mass ratio of oxygen in potassium sulfate to arrive at a correct answer for the mole concept assessment item *via* an incorrect process.

Responses were coded for ''incorrect process, correct answer'' if the student was observed to apply unnecessary, such as converting the moles of $NH_3$ provided in the changes in energy item into grams and back into moles, or chemically implausible, as described above with regard to the difference between mass ratio and moles of a constituent element, steps within their solution processes. Students who conflated units of measure (*e.g.*, describe the molar mass of a substance) or referred to chemicals by inaccurate nomenclatures (*e.g.*, describing $NH_3$ as NH) but otherwise were observed to arrive at a correct answer were not included in this measure as these students represented aspects of their processes inaccurately but did not execute inaccurate processes to arrive at a correct numerical answer. The quantitative results of these codes can be found in Table 3.

Each assessment item presented a unique variety of pathways by which students arrive at accurate numerical answers *via* inaccurate processes. To demonstrate the range of pathways, a qualitative description of not-at-risk students' responses processes follows.

**The mole concept item.** Fig. 14 presents exemplar responses of not-at-risk chemistry students for the mole concept item wherein students arrive at a correct numerical answer *via* an incorrect process.

Here, Ricky and Alex describe their solution processes following their descriptions with the calculations by which they arrived at their answer to the prompts. Ricky adopts the mass percent solution path incorporating the mass of 4 moles of oxygen atoms (64 g). The process can be summarized as follows:

(1) first convert the provided mass of $K_2SO_4$ to moles by dividing the molar mass of the compound,

(2) then, multiply the moles of $K_2SO_4$ by the mass of 4 moles of oxygen atoms,

(3) finally, divide by the molar mass of 1 mole of oxygen atoms.

The step in Ricky's solution process involving the multiplication of the moles and mass of two different chemical compounds demonstrates the conceptual challenge of interchangeable chemical identity. The instances in which students present all the steps of a predicted solution path but rearrange the order of these processes to one that is chemically implausible was coded in Fig. 14 as an error in sequence. Ricky, thereby, attains a solution to the assessment item that would be marked correct on the exam but was arrived at *via* a chemically implausible solution process that reflects a lack of conceptual understanding as to the meaning of the numerical values used in the algorithm.

 

Table 2 Codes (organized by theme) with prevalence measured as the percentage each code comprises of the total number of inaccuracies identified in students' solution processes ($n_{NAR}$ = 393, $n_{AR}$ = 234)

| Codes (by theme) | Code descriptions | NAR (%) | AR (%) |
|---|---|---|---|
| **Representational competence** | | | |
| Atomic *vs.* molecular | Atoms comprising species are misrepresented as covalent compounds using subscripts (*e.g.*, 4 O in potassium sulfate *vs.* $O_4$) | 4 | 3 |
| Misrepresented proportions | Proportions identified by the subscripts and/or coefficients of a chemical formula are misrepresented | 5 | 4 |
| Elemental symbols | Atomic masses are incorrectly assigned to elemental symbols as listed on the periodic table | 1 | 1 |
| Nomenclature | Chemical symbols are misinterpreted when determining the predominate state of matter or name of a chemical species | 2 | 3 |
| **Systematic approaches** | | | |
| Inaccurate conversion factors | Selected conversion factors between units of measure (*e.g.*, mass to moles; moles to mass) are incapable of resulting in a viable measure for the converted value (*e.g.*, converting mass to moles using Avogadro's number) | 5 | 5 |
| Misapplied molar mass | Molar masses are incorporated incorrectly and/or are unnecessary for the process described | 8 | 10 |
| Conversion arrangement | Numbers and units are arranged improperly result in units incapable of achieving a coherent number and unit of measure (*e.g.*, 12.1 moles O multiplied by 1 mol O/16.0 g O to achieve mass of O) | 5 | 7 |
| Sequence | Steps in the predicted solution path(s) are present but are rearranged in a manner that does not reflect a solution path that is chemically plausible | 6 | 3 |
| Coefficient and molar mass | Stoichiometric coefficients are incorrectly incorporated in the calculation of molar mass (*e.g.*, mass of $H_2$ with a coefficient of 2 is equal to $2 \times 2.0158$ g mol$^{-1}$ or 2.0158 g mol$^{-1}$ divided by 2). | 3 | 3 |
| ±Moles or mass | Often used for seemingly uncharacteristic solution pathways, the sum of some values of reactants to equal that of the products is used | 3 | 3 |
| Interchangeable chemical identity | Conversions between numerical values representative of different chemicals are conducted demonstrating a lack of identity reflective of the interrelation of one chemical to another (*e.g.*, 15 g of $K_2SO_4$ divided by atomic mass of oxygen to find moles of oxygen) | 19 | 15 |
| **Conceptual understanding** | | | |
| Interchemical unit identity | Moles of a compound are not presented as distinct from moles of a constituent element | 1 | 0 |
| Unit conflation | Differing units of measure are used interchangeably (*e.g.*, molar mass as mass) | 11 | 14 |
| Intrachemical unit identity | Terminologies inherent to particular units are misapplied (*e.g.*, using the term "the number of atoms" of a chemical to describe the moles of the same chemical) | 8 | 9 |
| Incorporating excess reactants | Despite the expected cue of "excess" in the prompt of the item task, solution processes are described that include determining the limiting reactant or the inclusion of the excess reactant in the solution process | 13 | 11 |
| Distinct measures | Conversions between numerical values representative of the same chemical, a lack of identity toward conversions that would be reflective of the interrelation of one unit to another is demonstrated (*e.g.*, equating the moles of oxygen in 15 g $K_2SO_4$ with the mass percent of oxygen in 15 g of $K_2SO_4$) | 6 | 9 |

Table 3 Prevalence for students who arrived at accurate numerical values for the assessment items implementing an inaccurate process

| Assessment item | NAR (%) | AR (%) | MD (NAR − AR) (%) |
|---|---|---|---|
| Mole concept item | 24 | 15 | 9 |
| Stoichiometry item | 15 | 13 | 2 |
| The gas laws item | 13 | 4 | 9 |
| Changes in energy item | 31 | 21 | 10 |

Alex also demonstrates a similar error *via* the predicted mole ratio path. Alex's solution process involves the multiplication of 15 grams of $K_2SO_4$ by the mole-to-mole ratio of 4 mol of oxygen atoms to 1 mol of potassium sulfate. The use of chemically implausible processes that result in the correct answers was common amongst not-at-risk chemistry students in solving the mole concept item (24% or 21 out of 86 inaccurate responses for this item). This compared to that of at-risk chemistry students who arrive at correct answers *via* inaccurate processes for the mole concept item (15% or 8 of 52 inaccurate responses) explains some of the differential performance observed between these two groups of students.

**The stoichiometry item.** One hundred fifty-four inaccurate solution processes were presented amongst the students (100 NAR, 54 AR) for the stoichiometry assessment item. While not-at-risk and at-risk cohorts manifested similar challenges with their solution processes, multiplying coefficients by the molar mass of chemicals and rearranging a chemically plausible sequence for the solution processes were two common routes by which students arose at a correct answer *via* an incorrect process (see Fig. 15).

Sequence (or the rearrangement of a solution process that would otherwise be accurate) was observed as a common method by which students arrived at an accurate value utilizing a solution pathway that does not attribute meaning or identity to the numerical values involved. Consider the depicted solution process designed to reflect an error made by Trudy and Donna.

$$(45.6\,\mathrm{g\,H_2})\left(\frac{1\,\mathrm{mol\,H_2}}{2.02\,\mathrm{g\,H_2}}\right)\left(\frac{17.03\,\mathrm{g\,NH_3}}{1\,\mathrm{mol\,NH_3}}\right)\left(\frac{2\,\mathrm{mol\,H_2}}{3\,\mathrm{mol\,NH_3}}\right)=257\,\mathrm{g\,NH_3}$$

While an accurate value is acquired, students apply the mole ratio following the multiplication of the attained moles of $H_2$ by the molar mass of $NH_3$ demonstrating a lack of chemical identity to these measures. Oliver describes another inaccurate solution path involving the incorporation of coefficients in the calculation of molar mass as depicted below.

$$(45.6\ \mathrm{g\ H_2})\left(\frac{2\cdot\dfrac{17.03\ \mathrm{g\ NH_3}}{1\ \mathrm{mol\ NH_3}}}{3\cdot\dfrac{2.02\ \mathrm{g\ H_2}}{1\ \mathrm{mol\ H_2}}}\right)=257\ \mathrm{g\ NH_3}$$

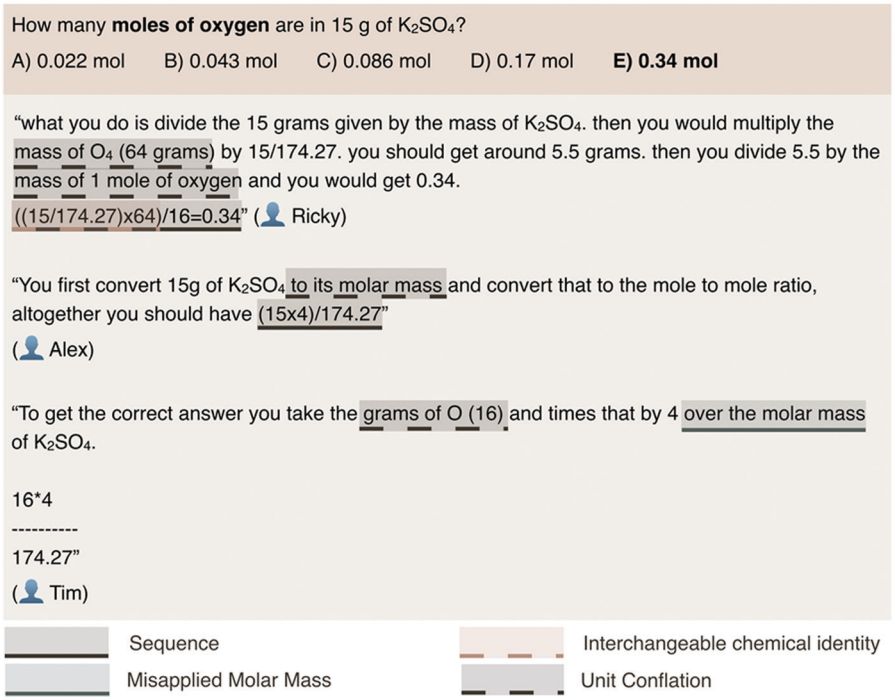


Fig. 14 Common and chemically implausible solution strategies along with representational competence challenges observed amongst the not-at-risk chemistry students.

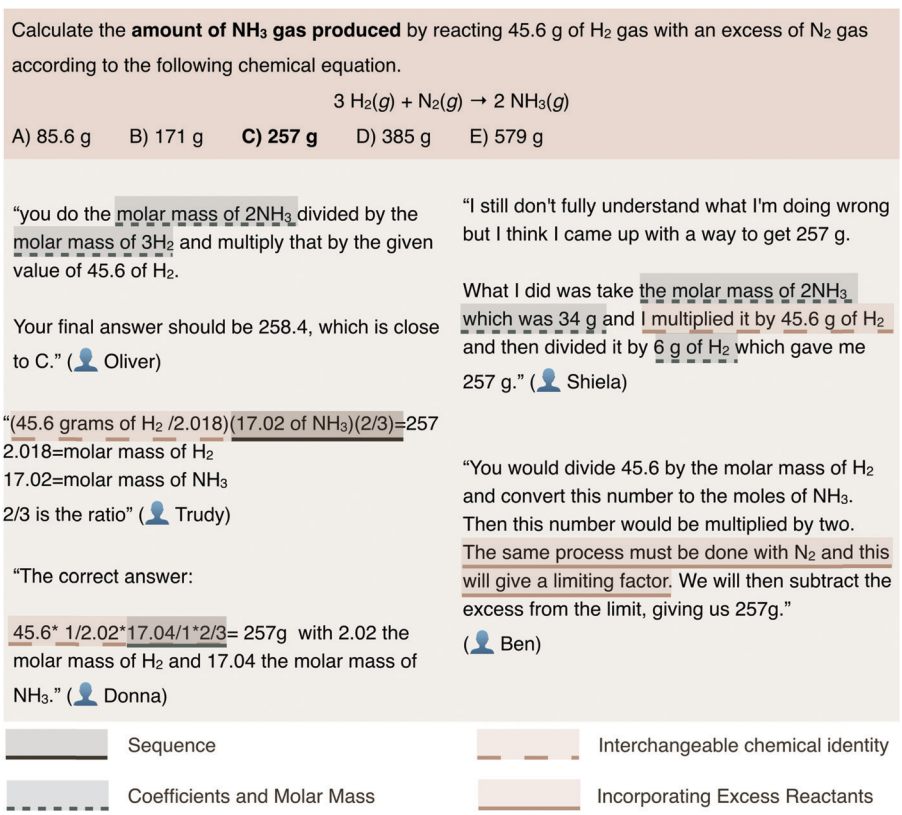


Fig. 15 Inaccurate conceptions of algorithms communicated by not-at-risk chemistry students that result in accurate responses to the stoichiometry assessment item.

 

This process functionally removes unit and chemistry identity for the numerical values used in the calculation. The lack of order to not-at-risk students' algorithms suggest that their conceptual understanding of these algorithms may not be more advanced than their at-risk peers, but rather not-at-risk students are more familiar with the execution of a predictable series of calculations with little meaning attributed to the numerical values that arise. A lack of order relegates the process of sequential unit cancelation as the primary means by which to solve a prompt, regardless of the chemical plausibility of a solution processes. These occurrences serve as another example by which items designed to assess a proficiency with attaining a final numerical have a potential to reward chemically implausible algorithm execution.

**The gas laws item.** Responses to the gas laws item demonstrate additional manifestations of not-at-risk students employing algorithms that arrive at the correct numerical value but fail to demonstrate a chemically plausible solution process (see Fig. 16).

While both groups of chemistry students were described as commonly interchanging the chemical identities of numerical values (*e.g.*, moles of $H_2$ *vs.* moles of $NH_3$ in the context of the gas law assessment item), not-at-risk students who err in their solution processes tended to apply proportions used to interconvert chemical species in a variety of ways. For example, the not-at-risk student responses in Fig. 16 refer to the conversion of 126 L of $H_2$ to 84 L of $NH_3$ as the conversion necessary to relate molecular hydrogen and ammonia rather than applying a mole-to-mole ratio. In the case of students like Rachel, Malcolm, Karen, and Phillip (see Fig. 16), students demonstrate a conceptual understanding of chemical identity by applying a mole ratio in their solution processes but seem to assume that temperature is held constant as the reaction proceeds and thus the mole ratio can be applied to the volume.

A similar variety of this code can be observed in Karen's solution process. Here, the mole ratio is not derived from the coefficients of the balanced chemical reaction but rather the subscripts of the formula wherein moles of hydrogen atoms in ammonia are 3 and in molecular hydrogen are 2. However, this error in the context of the problem reflects the challenges all students face with interpreting chemical symbolism in addition to conceptualizing how the depicted quantities interrelate. Of the inaccurate responses observed from not-at-risk students 13% arrived at a correct answer *via* a chemically implausible solution processes wherein only 4% of at-risk chemistry students were observed to do the same. These responses suggest a shortcoming in conceptual understanding of the processes by which answers are obtained.

**The changes in energy item.** Students solving the changes in energy assessment item (see Fig. 17) appeared more comfortable converting the 2.05 mol of $NH_3$ provided into mass and

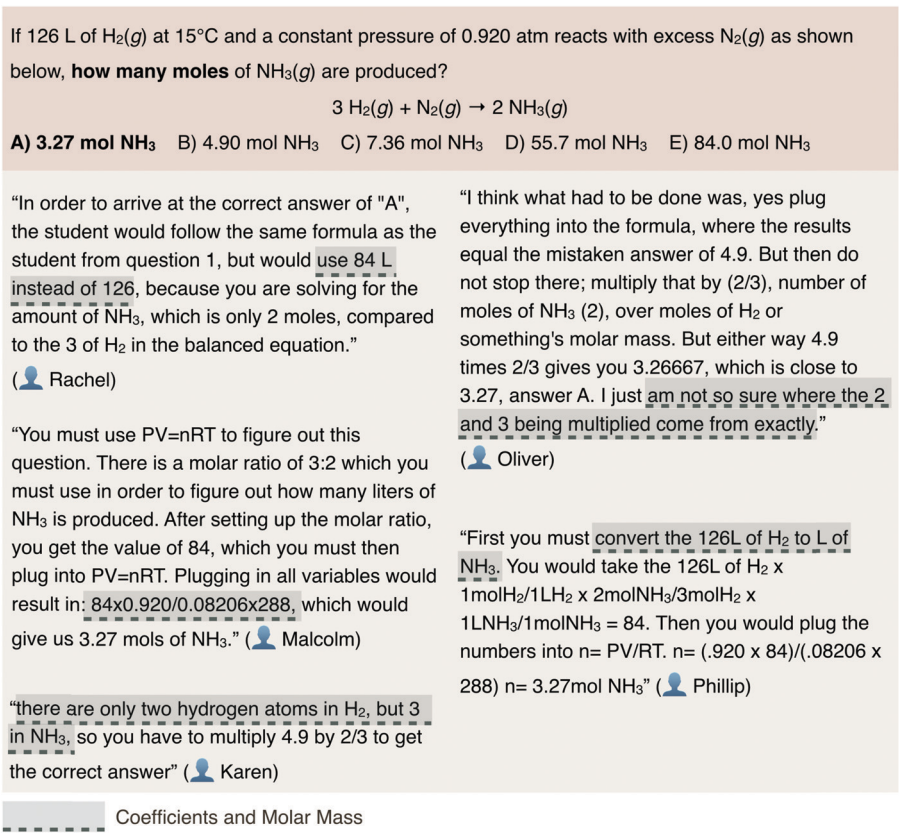


Fig. 16 Inaccurate conceptions of algorithms communicated by not-at-risk chemistry students that result in accurate responses to the gas laws assessment item.

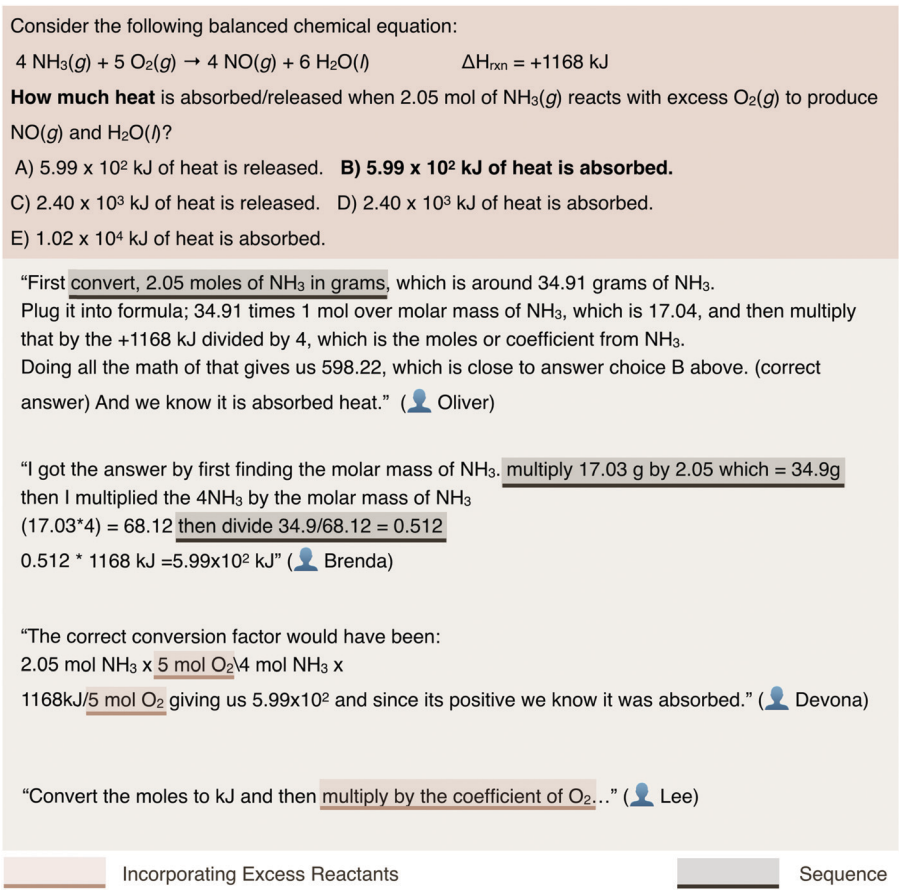


Fig. 17 Inaccurate conceptions of algorithms communicated by not-at-risk chemistry students that result in accurate responses to the changes in energy assessment item.

then solving the problem (even recalculating moles) perhaps as a means to rationalize their responses. This strict adherence to algorithms incorporating molar mass for an item where molar mass was not necessary supports the idea that students have attained proficiencies, not with the understanding of these processes, but with their execution. Exemplar responses are provided in Fig. 17.

Similarly to the other assessment items, 31% of the not-at-risk students arrived at a correct answer *via* a chemically implausible solution process (compared to 21% observed amongst at-risk students). Oliver's responses were chosen as exemplary for the last three figures as conceptual challenges with molar mass and its application was a common challenge observed throughout students' solution processes. This student demonstrates how some not-at-risk students successfully develop algorithms that, from the perspective of an instructor may demonstrate little conceptual proficiency yet result in accurate responses on multiple-choice exams designed to measure students' ability to attain a numerical response. Further, Oliver's repetition of this chemically implausible algorithm from test 1 to test 2 reflects the concern that assessment items designed to measure a student's ability to attain a numerical value *via* a step-by-step process could encourage and reward conceptually implausible processes.

# Discussion

## Findings of the study in the context of prior research

Challenges observed amongst at-risk chemistry students commonly related to conceptualizing the unit and chemical identities of numerical values in the context of their solution processes. Descriptions of students' challenges with the mole concept and stoichiometry in past works used as the conceptual framework for this study were summarized into themes of representational competence, systematic approaches, and conceptual understanding. Within these works, these themes often appear as three distinct challenges for the students; however, the authors posit a more pluralistic view of the challenges students face with the mole concept and stoichiometry. Each of the three themes manifested within students' struggles to conceptualize the processes used to solve the problems. For example, students were often observed to describe molar mass as mass and applying gram units to these values but were able to apply these values properly to achieve an accurate and chemically plausible answer to the problem (see Chris, David, and Gareth's responses to the mole concept prompt). In another research study, the impact of terminology and conflicting descriptions were thought to result from a lack of a consistent conception of the mole as presented in textbooks

 

and instruction (Furió *et al.*, 2002). It is possible that students' challenges to conceptualize the mole begin with the difficulty the scientific community as a whole has had in communicating its meaning and utility to chemists. Another plausible explanation for students' challenges with the mole relates to how this topic is assessed. At-risk chemistry students were observed to struggle with attributing chemical identities to numerical values, particularly when units were shared between chemicals. For example, consider the stoichiometry assessment item (Fig. 5) in which students were asked to convert 45.6 grams of molecular hydrogen to grams of ammonia. At-risk students were often observed to start the problem using the following process:

$$45.6\text{ g } H_2\left(\frac{1\text{ mol } NH_3}{17.03\text{ g } NH_3}\right)$$

Here, neither a terminological understanding of mass nor the algebraic arrangement of units for cancellation impacted a student's ability to conduct this conversion accurately. What appears to be lacking is a conceptual understanding of the process in which the molar mass of one chemical cannot be used to convert a mass of another chemical to moles. While chemical interchangeability was more common amongst at-risk chemistry students, difficulties with conceptualizing algorithms were regularly observed throughout. Similar observations were described in a study conducted by Staver and Lumpe (1995) wherein the extent of prior coursework in mathematics and chemistry was inconsequential concerning the lack of conceptualization amongst the students. Regardless of incoming preparation in mathematics or prior coursework in chemistry, students did not attain the conceptual understanding intended following their assessment of this topic.

> "This subject's responses show a high level of awareness of the mole as a vehicle for moving back and forth between the macro and atomic/molecular levels, but he is unable to explain the numerical identity issue and does not correctly work either of the problems. Surmounting this barrier requires much more than an awareness; it requires that students overcome... insufficient understanding of the concepts and use of memorized algorithms, rules, or other information." (Staver and Lumpe, 1995, p. 190)

Concerning representational competence, errors were observed in students' interpretations of the chemical symbolism provided; however, these errors rarely influenced the successes of their algorithmic solution processes. For example, students solving the mole concept item often described the moles of oxygen as O, $O_2$, or $O_4$ while retaining the coefficient of 4 in their calculations (see Fig. 4 and 14). The works of Gabel and Sherwood (1983), Potgieter and Davidowitz (2011), and Phillips (2001) suggest low-performing students and students who struggle with proportional reasoning could benefit *via* the provision of more visual instruction (*e.g.*, graphic organizers and diagrams of the particulate nature of matter). The challenges in the context of this study rarely reflected issues with interpreting chemical symbolism (see nomenclature and elemental symbols in Table 2) and were more commonly concerning how to use these values in an algorithmic process. Thus, these visuals seem less likely to support students' algorithmic success in solving these problems but may aptly address the more pertinent challenge in the students' conceptual understanding of stoichiometry. Representational competence and conceptual understanding most commonly converged in students' confusion concerning the distinct roles of coefficients and molar mass in calculating stoichiometric proportions (see Fig. 11, 15 and 16). Here students apply coefficients to the molar masses of chemicals and can typically arrive at correct answers. While reaching a correct answer, these solution processes demonstrate the conflation of mass and moles.

The knowledge demonstrated enacts an algorithmic property where students execute a step-by-step process learned and practiced throughout the semester. Assessments, as observed in the research setting, can reinforce these algorithmic strategies by requesting students replicate these algorithms in the selection of a single numerical product. Thus, the algorithmic nature of assessments designed to measure the mole concept and stoichiometry may result in fostering chemically implausible algorithms that inequitably favor students of higher math aptitude and fail to measure conceptual understanding. A study concerning students' approaches to solving stoichiometry problems concluded that while students with high proportional reasoning abilities used algorithmic reasoning strategies more frequently than their peers, students overall did not understand the chemical concepts on which the problems were based (Gabel *et al.*, 1984). The current study found similar outcomes as the algorithmic processes of not-at-risk students more commonly resulted in the achievement of desirable assessment scores, some of the differential performance observed could be interpreted as students with higher math aptitude scores achieving greater comfort in applying and devising algorithms regardless of their chemical plausibility. Overall the items involving mole-to-mole conversions, 70 (21%) of the 340 chemically implausible solution processes presented by not-at-risk chemistry students resulted in a correct answer for the items. Comparatively, 26 (13%) of the 199 inaccurate responses by the at-risk chemistry students accomplished the same result. As students of both groups were discovered to share similar misconceptions related to these topics (see Table 2), it is possible *how* the mole concept and stoichiometry are assessed creates and reinforces the differential performance observed.

This implication is of consequence to instructors, researchers, and institutions interested in more equitable assessment design and closing performance gaps observed amongst students of variable incoming preparation. The combination of systematic approaches and conceptual understanding toward students' conceptualization of the algorithms used in addition to the similarities observed amongst not-at-risk and at-risk chemistry students in a lack of conceptual understanding of stoichiometry were informative. However, the inequitable successes following the implementation of chemically implausible algorithms observed as a potential source for the differential performance observed between these group is a concerning finding of this study. Ultimately, the work presented furthers the literature

concerning at-risk chemistry students (or those with low math aptitude scores) and in the challenges these students experience with the mole concept and stoichiometry.

Practical adaptations to the assessment items designed to measure proficiency with these topics could reduce the inequity observed by reducing algorithmic reliance and emphasizing the conceptualization of the processes by which students engage in solving stoichiometry problems. Such adaptations to assessment items could (1) improve the learning experiences of students by providing more detailed feedback related to their performance as a result of the distractors they've selected on multiple-choice assessment items, (2) provide instructors with tractable data facilitating responsive changes to instruction following assessments, and (3) provide a means to potentially reduce the differential performance observed on these topics amongst cohorts of diverse student preparation. These objectives may be attained *via* the alignment of assessment prompts with the tasks they elicit. Engaging in this practice could improve the confidence in cognitive validity, the relationship between what an assessment aims to measure and what it elicits from the student, of the data collected from assessment items (Ayala, 2002; Geranpayeh and Taylor, 2013). Practical suggestions for adaptations to assessment items to promote conceptual understanding of the processes used to solve stoichiometry problems in response to the findings of this study are discussed below and present opportunities for future exploration.

### Implications for more equitable assessment design

The processes used to solve the items were far more conceptual than expected and involved processes that intertwine students' representational competence and conceptual aspects related to the unit and chemical identity of numerical values used. This skill set is highly process-oriented suggesting that one way to align the task to the item and improve the cognitive validity of data acquired from multiple-choice questions on the mole concept or stoichiometry is to assess the process and not the product. For example, consider the mole concept item to which students' solution processes were analyzed (Fig. 18, left tile) in this study redesigned to items that emphasize process-over-product (Fig. 18, center and right tiles) in addition to clarifying the interpretation of the stem by replacing the phrase ''moles of oxygen'' with ''moles of oxygen atoms''.

In the original assessment item (left), not-at-risk students were often observed to engage in solution processes that, although chemically implausible, could result in an accurate response (see responses authored by Alex and Ricky in Fig. 14). The data collected from such a product-oriented assessment item may be biased in favor of not-at-risk chemistry students who have shown an ability to arrive at accurate responses *via* inaccurate conceptions of an algorithm. The adaptations of the mole concept item (left tile of Fig. 18) to more process-oriented representations (center and right tiles of Fig. 18) could have a number of benefits. Process-oriented assessment items can elicit unique aspects of students' solution process. The center tile focuses on misidentifying the mass percent of oxygen atoms in potassium sulfate as the moles of oxygen atoms in the compound (choice D) as observed in Fig. 4 and interchangeable chemical identities (A and C) as observed in Fig. 3. The right tile focuses on the atomic *versus* molecular representations of oxygen observed in Fig. 6. The distractors of process-oriented assessment items can also present a variety of answer choices that achieve a shared numerical value. Note the right tile where answer choices B and E would both result in the correct numerical value yet choice B would identify the number incorrectly as moles of $O_4$. The items in the center and right tiles promote student attention to the solution process.

From the instructor perspective, consider the data received regarding student proficiency with this concept across the three items. Instructors would be able to explicitly identify errors in the processes by which students engage including whether students confuse molecular with atomic oxygen, struggle with interchangeable chemical identities, or are unsure of the sequence by which chemically plausible solution processes occur *via* the distractors selected by their students. Pyburn *et al.* (2014) describe the impact distractors have on learning experiences of students taking multiple-choice tests and the benefit that follows the purposeful alignment of distractors

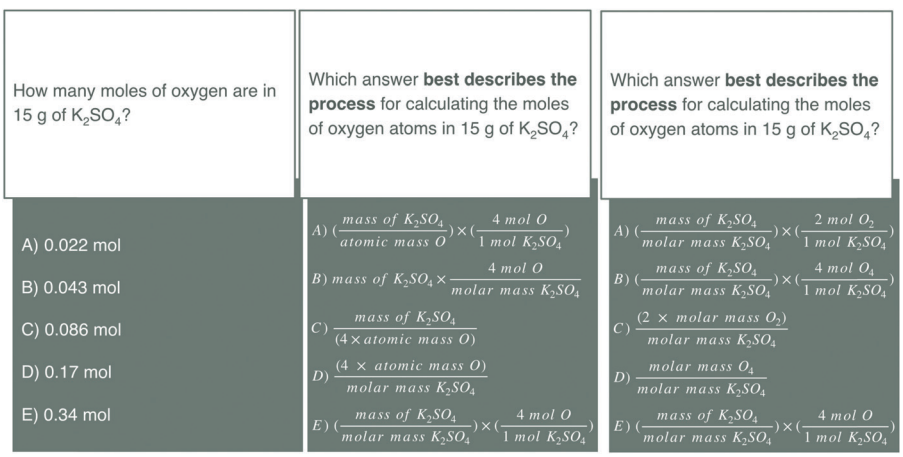


Fig. 18 Potential adaptation of distractors for a product-oriented assessment item to that of a process-oriented assessment item.

 

with common alternative conceptions (or in the case of this study, chemically implausible solution process) when provided to students as feedback. Instructors could insert process-related challenges they have observed amongst their students or those observed in the presented study and gauge from the data the degree to which these challenges impact students' conception of the topic. Further, by assessing process, the potential for reinforcing incorrect strategies that arrive in the correct numerical answer, as noted in students' response processes, is avoided along with the inequitably this introduces to cohorts of students with variable preparation in mathematics.

### Other implications for improved assessment design

One error prevalent amongst inaccurate responses of the students overall was confusion regarding the interpretation of "in excess" or "excess." The stoichiometry, gas laws, and changes in energy assessment items uses the phrase "excess" to indicate that one of the two reactants involved in the chemical reaction is not limiting and thereby the limiting reactants algorithm is not necessary for determining the theoretical yield of the reaction described. Students often did not attribute "excess" with "the other reactant is limiting" and instead expressed solution pathways that at best were inefficient but resulted in accurate responses and at worst confused students to the point where they were unable to respond.

Should the intent be to measure students' proficiency with determining theoretical yield and not connecting the phrase "excess" to the algorithm then eliminating the use of "excess" would be beneficial. Ultimately, researchers and instructors may consider the inclusion of "excess" as a confounding variable that poses a threat to the validity data collected on assessments of theoretical yield. Three possible options for adapting items to reduce this threat to the validity data collected are shown in Fig. 19.

The left and center tiles in Fig. 19 provide options for maintaining the reaction without the use of the word excess. Additionally researchers and educators may replace the chemical reactions related to calculations of theoretical yield with that of decomposition reactions (Fig. 19, right tile) such that no competing reactants are present that could influence students' solution processes.

One final consideration for the improvement of assessment design where the chemical identity and units associated with a numerical value are of importance is the inclusion of units with chemical identity in the distractors. The mole concept assessment item, for example, lists numerical values with corresponding answer choices without a description of chemical identity. It may change the manner in which students respond. For example, 23.7% of students selected distractor C (Fig. 2) of "0.086 mol" but may have changed their response if the answer choice read "0.086 mol of O". In particular, this is possible for students who reference that this value is indicative of the moles of potassium sulfate and not oxygen. Such a change would be reflective of their proficiency with the material and could serve as a learning opportunity for the student. Further, students may learn from the assessment and the data collected regarding their proficiency should common numerical values be paired with different units and chemicals identities. For example, should the distractors in Fig. 2 be replaced with: "0.022 mol of O", "0.086 mol of $O_4$", "0.086 mol of $O_2$" and "0.34 mol of $O_2$" with the correct answer "0.34 mol of O" it would require a student to consider both the numerical responses and the chemical representation.

### Pedagogical implications

Adaptations to assessment design offer potential advantages to improving the accuracy of data generated from the assessment and ameliorating differential performance. Response processes observed in the dataset also provide guidance for pedagogy and preparation of instructional materials. In light of the interchanging of chemical identity that was observed, instruction and associated materials may benefit by explicitly incorporating chemical identity along with units (*e.g.* grams per mol $NH_3$) during the presentation of worked problems. By explicitly noting chemical identity students may avoid the interchange of chemical identity and form a stronger conceptual link between the numerical values and their relevance toward the chemical phenomenon. While presenting worked problems, multi-step factor-label representations could be replaced with solution processes that pause after each conversion to describe the chemical meaning of the numerical value (see below).

Multi-step, factor-level representation

$$11.27\ \text{g}\ \text{N}_2 \times \frac{1\ \text{mol}\ \text{N}_2}{28.01\ \text{g}\ \text{N}_2} \times \frac{2\ \text{mol}\ \text{NH}_3}{1\ \text{mol}\ \text{N}_2} \times \frac{17.03\ \text{g}\ \text{NH}_3}{1\ \text{mol}\ \text{NH}_3} = 13.70\ \text{g}\ \text{NH}_3$$

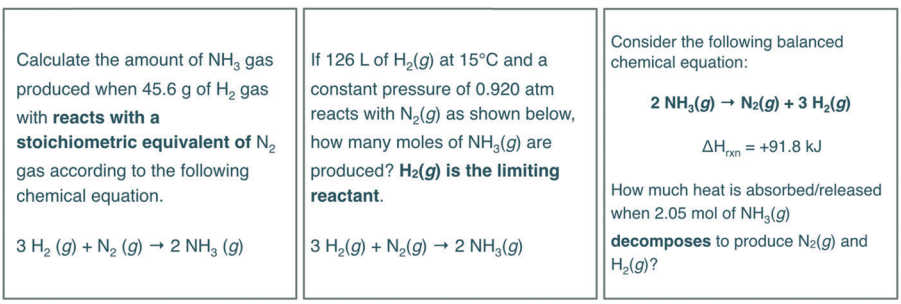


Fig. 19 Suggestions for adapting assessment items in Fig. 5, 8 and 10 toward items to avoid eliciting a limiting reactants algorithm from students as a result of reading the term "excess".

Each step presented with units and chemical

$$11.27\ \text{g N}_2 \times \frac{1\ \text{mol N}_2}{28.01\ \text{g N}_2} = 0.4024\ \text{mol N}_2$$

$$0.4024\ \text{mol N}_2 \times \frac{2\ \text{mol NH}_3}{1\ \text{mol N}_2} = 0.8047\ \text{mol NH}_3$$

$$0.8047\ \text{mol NH}_3 \times \frac{17.03\ \text{g NH}_3}{1\ \text{mol NH}_3} = 13.70\ \text{g NH}_3$$

A plausible explanation for the prevalence of the code "intra-chemical unit identity" observed commonly throughout students' solution processes is the conflation between moles as a coefficient in a chemical equation and as a starting value in the prompt of an item. Another way to further support students in delineating these quantities may be to present tabular representations of the values determined when presenting worked problems and encouraging students to do the same in their work (see Table 4 for an example). For example, if the stoichiometry assessment item was presented as a worked problem, instructors could model creating and completing the table while working the problem. Subsequent student assignments could then be scaffold to first complete instructor-provided tables and later for students' to generate and complete their own tables.

Didactic proposals for the use of tables in stoichiometry calculations can be found in the chemistry education (Watkins, 2003) and engineering education (Serafin, 2006) research literature bases. Further research concerning students' response processes to table stoichiometry may be a helpful start for evaluating whether or not this strategy could promote conceptual understanding of students' solution processes when solving stoichiometry problems.

Finally, instruction may benefit by addressing students' use of molar mass in the context of chemical reactions. Students' processes described multiplying molar mass by the stoichiometric coefficient. Conventionally, introduction of molar mass calculations takes place without the context of chemical reactions. Later when stoichiometry is introduced, student proficiency with molar mass is presumed without explicit mention for how chemical coefficients fail to alter the molar mass. Instead stoichiometry offers an opportunity to reinforce the definition of molar mass as a conversion factor that is not influenced by the number of moles *via* its operationalization as the mass of one mole.

Table 4 Example tabular representation of worked problem

| | $N_2(g)$ | $H_2(g)$ | → | $NH_3(g)$ |
|---|---|---|---|---|
| Molar mass | | | | |
| Initial mass | 11.27 g | 25.19 g | | 0 g |
| Initial moles | | | | |
| Change in moles | | | | |
| Final moles | | | | |
| Final mass | | | | |

## Limitations and future works

The data collected in this study reflect the responses of chemistry students within a single semester at a research setting, which consists of a single institution. While the intent of the study is not to provide generalizable characterizations of the challenges chemistry students face when solving mole concept and stoichiometry questions, the authors sought to generate hypotheses regarding the nature of at-risk student struggles and the causes of differential performance observed with these topics. The information elicited from students regarding mole concept and stoichiometry assessment items were (as argued in the study) highly dependent on the items selected for review. All of the questions selected may be considered "traditional" items and are product-oriented, multiple-choice items that rely on chemical formulas and balanced chemical reactions to convey information as to the proportions by which chemicals react. As representational competence played a role in the lack of conceptual understanding conveyed, future explorations of students' representational competence and other forms of chemical representation (*e.g.*, submicroscopic diagrams) may be analyzed to discern whether this skill set contributes to differential performance or the overall difficulty of an assessment item.

The consistency among assessment items in both representation and the use of the particular chemical reaction depicting the synthesis of ammonia was intentional to reduce conflating variables when reviewing students' performance on a subset of tasks. This decision could also limit the information observed regarding challenges students' experience with these items and sources of differential performance observed. Items were designed naturalistically by the course instructors and were chosen *a priori* to students' performance and differential performance. Thus, items intentionally selected as a result of high mean differential could provide a different viewpoint of the challenges that separate these two groups of students. Given the information collected, researchers could seek to design assessment items with the intent of identifying a particular challenge with a conceptual model in mind for how at-risk chemistry students struggle and explore differential performance based on assessment design. Other limitations of the study include the role of motivation and implicit assumptions made regarding students' responses. Students exhibited a great range of responses to the items, and one could argue that some of the errors observed were the result of students' apathy toward completing the survey with greater detail. While the consistency of the errors observed and the nature of detail given for the responses suggest students' misconceptions were not necessarily derived from motivation, these concerns informed decisions made by the authors in the treatment of the data. Only students who responded to all three surveys following interim exams were included in the analysis presented above. Additionally, student responses too vague to evaluate students' efforts such as "I don't know" or "just use stoichiometry" were not considered in the description of the challenges students faced with their solution processes.

A final limitation concerning the data elicited by the survey prompts was the frequency by which researchers had to make

 

implicit assumptions about the unit, chemical, or numerical identity of a value provided by the students. For example, a student could start their solution process to the mole concept item as "take 15 grams and divide by the molar mass" which was interpreted by the researchers to mean "divide 15 grams of $K_2SO_4$ by the molar mass of $K_2SO_4$". While the data collected more efficiently provided the solution processes for hundreds of students in the course, in-depth interviews with a subset of students are a target for future exploration as therein exists the capability for researchers to ask follow-up questions regarding the identities of the values they present in their solution processes. Suggestions for adaptations of assessment items including an emphasis in process observed in Fig. 18, adaptations to prompts presented in Fig. 19, and table stoichiometry as exemplified by Table 4 have not been explicitly tested by the researchers. Response processes for these changes in addition to reports on student outcomes are intended in future works but do not encompass the scope of the presented works intended to identify how students struggle with the stoichiometry assessment items on which at-risk chemistry students were identified to perform differently. Ultimately, the authors hope to issue a call for future research as to the impact of assessment design on differential performance. Whether by in-depth interviews or a compiled database of items presenting with variable differentials, greater nuance as to moderators of difficulty and differential performance could support assessment design and inform interventions or prerequisite practices issued for students of variable preparation entering this critical point in the STEM coursework.

## Conclusion

Despite comparable proficiencies with the representational competency, systematic approaches, and conceptual understanding required for the mole concept and stoichiometry assessment items (see Table 2), students' success at implementing chemically implausible algorithms appears the most apparent source of differential performance between chemistry students of lower and higher math aptitude scores. Not-at-risk chemistry students were observed to make similar errors to those of their at-risk peers often regarding the conceptual understanding of the algorithms used to solve mole concept and stoichiometry assessment items. However, a difference was observed in the commonality by which not-at-risk students used chemically implausible solution paths to arrive at the correct numerical value elicited by the prompt of these assessment items when compared to their at-risk peers. As the differential performance observed does not seem to reflect differences amongst the students' knowledge of chemistry and instead favors students entering with more preparation in the rote execution of algorithms, current practices involving the assessment of stoichiometry is inequitable. The primary challenges observed in students' solution processes with the mole concept and stoichiometry assessment items were: (1) interchanging chemical identities of the numerical values used in stoichiometry, (2) distinguishing representational differences between atomic and molecular representations of elements capable of forming diatomic molecules, (3) applying mass percentages to moles or molar mass, (4) applying stoichiometric coefficients in the determination of molar mass, (5) inserting molar mass in algorithms requiring only mole-to-mole ratios, and (6) applying solution processes related to identifying limiting reactants where all other reactants are in excess. To reduce the inequity concerning how this topic is assessed, instructors and researchers should actively seek out processes that invoke a conceptual understanding of stoichiometry. Potential changes to instruction and assessment of stoichiometry including the use of table stoichiometry as a scaffold for students, replacement of product-oriented assessment items (*e.g.*, calculate $x$) with process-oriented assessment items (*e.g.*, which process describes how to calculate $x$) and other proposed multiple-choice formats are presented with an intent to support future research aimed at addressing the observed performance gaps in this gateway STEM course.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

The authors thank the undergraduate researchers (Gustavo A. Velasquez and Yvonne Nguyen) who provided a student-perspective to the challenges observed in students' response processes as described in methodology presented for this manuscript. Appreciation is also extended to the instructors and students at the research setting and the anonymous reviewers of this manuscript. Partial support for this work was provided by the National Science Foundation's Improving Undergraduate STEM Education (IUSE) program under DUE-1712164 and Florida-Georgia Louis Stokes Alliance for Minority Participation Bridge to the Doctorate award HRD-1612347. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

Andrews M. H. and Andrews L. J., (1979), First-year chemistry grades and SAT-M scores, *J. Chem. Educ.*, **56**(4), 231–232, DOI: 10.1021/ed056p231.

Ayala C. C., (2002), *On the cognitive validity of science performance assessments*, Stanford University, [WWW Document]. URL https://search.proquest.com/docview/305523863?accountid=14745%0A (accessed 1.12.19).

Bender D. S. and Milakofsky L., (1982), College chemistry and Piaget: the relationship of aptitude and achievement measures, *J. Res. Sci. Teach.*, **19**(3), 205–216, DOI: 10.1002/tea.3660190303.

Bialek W. and Botstein D., (2004), Introductory science and mathematics education for 21st-century biologists, *Science*, **303**(5659), 788–790, DOI: 10.1126/science.1095480.

BouJaoude S. and Barakat H., (2003), Students' problem solving strategies in stoichiometry and their relationships to conceptual understanding and learning approaches, *Electron. J. Sci. Educ.*, **7**(3), 42, [WWW Document], URL http://ejse.southwestern.edu/article/view/7709/5476 (accessed 11.13.18).

Coley N. R., (1973), Prediction of success in general chemistry in a community college, *J. Chem. Educ.*, **50**(9), 613, DOI: 10.1021/ed050p613.

College Board, (2018), *SAT Suite of Assessments Annual Report*, [WWW Document] URL https://reports.collegeboard.org/pdf/2018-total-group-sat-suite-assessments-annual-report.pdf (accessed 22.03.2019).

Dahsah C. and Coll R. K., (2007), Thai Grade 10 and 11 students' conceptual understanding and ability to solve stoichiometry problems, *Res. Sci. Technol. Educ.*, **25**(2), 227–241, DOI: 10.1080/02635140701250808.

Davidowitz B. and Chittleborough G., (2009), Linking the macroscopic and sub-microscopic levels: diagrams, in Gilbert J. K. and Treagust D. (ed.), *Multiple representations in chemical education*, vol. 4, pp. 169–194, DOI: 10.1007/978-1-4020-8872-8_9.

Davidowitz B., Chittleborough G. and Murray E., (2010), Student-generated submicro diagrams: a useful tool for teaching and learning chemical equations and stoichiometry, *Chem. Educ. Res. Pract.*, **11**(3), 154–164, DOI: 10.1039/C005464J.

Donovan W. J. and Wheland E. R., (2009), Comparisons of success and retention in a general chemistry course before and after the adoption of a mathematics prerequisite, *Sch. Sci. Math.*, **109**(7), 371–382, DOI: 10.1111/j.1949-8594.2009.tb17868.x.

Furió C., Azcona R. and Guisasola J., (2002), The learning and teaching of the concepts ''amount of substance'' and ''mole'': a review of the literature, *Chem. Educ. Res. Pract.*, **3**(3), 277–292, DOI: 10.1039/B2RP90023H.

Gabel D. L. and Sherwood R. D., (1983), Facilitating problem solving in high school chemistry, *J. Res. Sci. Teach.*, **20**(2), 163–177, DOI: 10.1002/tea.3660200207.

Gabel D. L., Sherwood R. D. and Enochs L., (1984), Problem-solving skills of high school chemistry students, *J. Res. Sci. Teach.*, **21**(2), 221–233, DOI: 10.1002/tea.3660210212.

Geranpayeh A. and Taylor L., (2013), *Examining listening: research and practice in assessing second language listening*, Cambridge University Press.

Gosser D. K., Strozak V. and Cracolice M. S., (2006), *Peer-led team learning: general chemistry*, Upper Saddle River, NJ: Prentice Hall.

Hailikari T. K. and Nevgi A., (2010), How to diagnose at-risk students in chemistry: the case of prior knowledge assessment, *Int. J. Sci. Educ.*, **32**(15), 2079–2095, DOI: 10.1080/09500690903369654.

Hall D. M., Curtin-Soydan A. J. and Canelas D. A., (2014), The science advancement through group engagement program (SAGE): leveling the playing field and increasing retention in science, *J. Chem. Educ.*, **91**(1), 37–47, DOI: 10.1021/ed400075n.

IBM Corp., (2017), Released IBM SPSS Statistics for Macintosh, Version 25.0, Armonk, NY: IBM Corp.

Kozma R. and Russell J., (2005), Students becoming chemists: developing representational competence, in *Visualization in science education*, Springer: Dordrecht, pp. 121–145.

Larson J. O., (1997), Constructing understandings of the mole concept: interactions of chemistry text, teacher and learners, *Annu. Meet. Natl. Assoc. Res. Sci. Teach., 70th*, Oak Brook, IL.

Lee S., Crane B. R., Ruttledge T., Guelce D., Yee E. F. and Lenetsky M., (2018), Patching a leak in an R1 university gateway STEM course, *PLoS One*, **13**(9), 1–16, DOI: 10.1371/journal.pone.0202041.

Lewis S. E., (2011), Retention and reform: an evaluation of peer-led team learning, *J. Chem. Educ.*, **88**(6), 703–707, DOI: 10.1021/ed100689m.

Lewis S. E. and Lewis J. E., (2007), Predicting at-risk students in general chemistry: comparing formal thought to a general achievement measure, *Chem. Educ. Res. Pract.*, **8**(1), 32–51, DOI: 10.1039/B6RP90018F.

Lincoln Y. S. and Guba E. G., (1986), But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation, in Williams D. D. (ed.), *Naturalistic evaluation*, San Francisco, CA: Jossey-Bass, pp. 73–84.

Lynch S. J., (2000), *Equity and science education reform*, Lawrence Erlbaum Associates: Mahwah, NJ.

Miles M. B., Huberman A. M. and Saldaña J., (2014), *Qualitative data analysis: a methods sourcebook*, Thousand Oaks, California: SAGE Publications, Inc.

Ozsogomonyan A. and Loftus D., (1979), Predictors of general chemistry grades, *J. Chem. Educ.*, **56**(3), 173–175, DOI: 10.1021/ed056p173.

Peshkin A., (1993), The Goodness of Qualitative Research, *Educ. Res.*, **22**(2), 23–29, DOI: 10.3102/0013189X022002023.

Phillips C. K., (2001), The effects of an integrated computer program on math and reading improvement in grades three through five, The University of Tennessee [WWW Document]. URL https://doi.org/10.16953/deusbed.74839 (accessed 12.08.18).

Pickering M., (1975), Helping the high risk freshman chemist, *J. Chem. Educ.*, **52**(8), 512, DOI: 10.1021/ed052p512.

Potgieter M. and Davidowitz B., (2011), Preparedness for tertiary chemistry: multiple applications of the Chemistry Competence Test for diagnostic and prediction purposes, *Chem. Educ. Res. Pract.*, **12**(2), 193–204, DOI: 10.1039/c1rp90024b.

Pyburn D. T., Pazicni S., Benassi V. A. and Tappin E. M., (2014), The testing effect: an intervention on behalf of low-skilled comprehenders in general chemistry, *J. Chem. Educ.*, **91**(12), 2045–2057, DOI: 10.1021/ed4009045.

Ralph V. R. and Lewis S. E., (2018), Chemistry topics posing incommensurate difficulty to students with low math aptitude scores, *Chem. Educ. Res. Pract.*, **19**(3), 867–884, DOI: 10.1039/C8RP00115D.

Ramful A. and Narod F. B., (2014), Proportional reasoning in the learning of chemistry: levels of complexity, *Math. Educ. Res. J.*, **26**(1), 25–46, DOI: 10.1007/s13394-013-0110-7.

 

Rixse J. S. and Pickering M., (1985), Freshman chemistry as a predictor of future academic success, *J. Chem. Educ.*, **62**(4), 313, DOI: 10.1021/ed062p313.

Schank P. and Kozma R., (2002), Learning chemistry though the use of a representation-based knowledge building environment, *J. Comput. Math. Sci. Teach.*, **21**(3), 253–279, Norfolk: VA: Association for the Advancement of Computing in Education (AACE). [WWW Document]. URL https://www.learntechlib.org/primary/p/9262/ (accessed 19.04.2019).

Schmidt H., (1994), Stoichiometric problem solving in high school chemistry, *Int. J. Sci. Educ.*, **16**(2), 191–200.

Scott F. J., (2012), Is mathematics to blame? An investigation into high school students' difficulty in performing calculations in chemistry, *Chem. Educ. Res. Pract.*, **13**(3), 330–336, DOI: 10.1039/c2rp00001f.

Serafin M., (2006), *Engineering's approach to teach stoichiometry*, Paper presented at 9th International Conference on Engineering Education: San Juan, PR [WWW Document]. URL http://www.icee.usm.edu/ICEE/conferences/ICEE2006/papers/3359.pdf (accessed 18.04.2019).

Spencer H. E., (1996), Mathematical SAT test scores and college chemistry grades, *J. Chem. Educ.*, **73**(12), 1150, DOI: 10.1021/ed073p1150.

Spencer L., Ritchie J., Lewis J. and Dillion L., (2003), Quality in Qualitative Evaluation: A framework for assessing research evidence, *Government Chief Social Researcher's Office*, **5**(1–4), 182–185, DOI: 10.1016/S0009-9120(72)80029-8.

Staver J. R. and Lumpe A. T., (1995), Two investigations of students' understanding of the mole concept and its use in problem solving, *J. Res. Sci. Teach.*, **32**(2), 177–193, DOI: 10.1002/tea.3660320207.

Tro N. J., (2014), *Chemistry: A Molecular Approach*, Boston, Columbus: Pearson Education.

VERBI Software, (2017), *MAXQDA 2018 [computer software]*, Berlin, Germany: VERBI Software.

Wagner E. P., Sasser H. and DiBiase W. J., (2002), Predicting students at risk in general chemistry using pre-semester assessments and demographic information, *J. Chem. Educ.*, **79**(6), 749, DOI: 10.1021/ed079p749.

Watkins S. F., (2003), Applying the reaction table method for chemical reaction problems (stoichiometry and equilibrium), *J. Chem. Educ.*, **80**(6), 658–661, DOI: 10.1021/ed080p658.

Ye L., Oueini R. and Lewis S. E., (2015), Developing and implementing an assessment technique to measure linked concepts, *J. Chem. Educ.*, **92**(11), 1807–1812, DOI: 10.1021/acs.jchemed.5b00161.

Ye L., Shuniak C., Oueini R., Robert J. and Lewis S., (2016), Can they succeed? Exploring at-risk students' study habits in college general chemistry, *Chem. Educ. Res. Pract.*, **17**(4), 878–892, DOI: 10.1039/C6RP00101G.