PAPER

Analysing the distribution of questions in the gas law chapters of secondary and introductory college chemistry textbooks from the United States

Gabriel Gillette[a] and Michael J. Sanger*[b]

This study analysed the distribution of questions from the gas law chapters of four high school and four college chemistry textbooks based on six variables—Book Type (secondary *versus* introductory college), Cognitive Skill (lower-order *versus* higher-order), Question Format (calculation *versus* multiple-choice 10 *versus* short-answer), Question Placement (in-chapter *versus* end-of-chapter *versus* test-bank), Question Type (qualitative *versus* quantitative), and Representation (macroscopic *versus* particulate *versus* symbolic). The questions in these chapters were homogeneously distributed for the Cognitive Skill and the Representation variables, but showed differences in question distribution based on the Book Type, Question Format, Question Placement, and Question Type variables. The loglinear analysis method used 15 in this study provides one way to analyse the distribution of different types of questions appearing in chemistry textbooks, and these differences in question distribution can be helpful for textbook authors to evaluate the types of questions appearing in their textbooks and how they are presented, and can be helpful for chemistry instructors to determine how they need to adapt their instructional lessons to prepare students for course examinations or college/career placement examinations.

20 Introduction

For many secondary (high school) and tertiary (college) chemistry courses, the influence of the textbook cannot be overestimated. The textbook often dictates what chemical concepts the instructor teaches in the classroom and in what order 25 they are taught (Britton *et al.*, 1993; Tulip and Cook, 1993; Tobin *et al.*, 1994; Justi and Gilbert, 2002; Drechsler and Schmidt, 2005; Roth *et al.*, 2006; Dávila and Talanquer, 2010), and students and parents expect the textbook to be the main source of information (Hurd *et al.*, 1981; Chiang-Soong and Yager, 1993; 30 Weiss, 1993). Published textbooks are used in 90% of secondary science courses in the United States and 75% or more of the content is covered in class (Weiss, 1987); textbooks have become encyclopaedias of scientific facts that encourage teachers and students to focus on memorization (Gannaway and Stucke, 1996; 35 Stinner, 2001; Stern and Ahlgren, 2002; Resnick, 2007). In addition, several researchers have shown that inaccurate or imprecise language used by textbook authors can lead to student misconceptions in chemistry (Suidan *et al.*, 1995; Sanger and Greenbowe, 1999; Pedrosa and Dias, 2000; Drechsler and 40 Schmidt, 2005; Bergqvist *et al.*, 2013).

Since the end-of-chapter questions in textbooks can be used to assign homework grades and test banks (books provided by textbook publishers with a large number of questions related to the material covered in their textbooks) can be used to write

45 examinations that are used in deciding students' final course grades (Pappa and Tsaparlis, 2011), the types of questions asked in these textbooks can have a dramatic impact on the students' future (Boud, 1995; Ramsden, 2003; Danili and Reid, 2005; Schroeder *et al.*, 2012). With the advent of No Child Left Behind 50 (NCLB) and high-stakes testing in the United States, assessment of students can also have a dramatic impact on the future of the educational institutions since NCLB has provisions to hold schools, districts, and states accountable for their students' performance, and textbooks can play a role in these issues of 55 assessment (Colantonio, 2005; Witzel and Riccomini, 2007; Wright and Li, 2008). Therefore, it is important to analyse the types of questions being asked to students in these textbooks to determine what types of information or skills—content knowledge, thinking skills, calculation skills, etc.—are being 60 assessed (Davis and Hunkins, 1996; Justi and Gilbert, 2002; Dávila and Talanquer, 2010; Pappa and Tsaparlis, 2011).

Variables used to categorise textbook questions

In this study, we used six variables to categorise the questions present in the gas law chapters of eight chemistry textbooks from 65 the United States. These variables, and the categories within each variable, are summarized in Table 1. Analysing all of the questions in an entire textbook can be a daunting task, so we decided to focus our analysis on the gas law chapters of these chemistry books. We chose the gas law chapters because the

This journal is © The Royal Society of Chemistry [year]

Chem. Educ. Res. Pract., [year], **[vol]**, 00–00 | **1**

concepts discussed in these chapters contain both mathematical calculations and non-mathematical concepts, and these concepts lend themselves to discussions at the macroscopic (behaviour of gas samples), particulate (kinetic-molecular theory), and symbolic (calculations, analysis of graphs, etc.) levels.

Table 1 The six variables and their categories used in this study

Variable	Categories
Book Type (BT)	HS: Textbooks written for use in high-school (secondary) chemistry courses
	CO: Textbooks written for use in introductory college (tertiary) chemistry courses
Cognitive Skill (CS)	LO: Questions involving lower-order (simple recall or algorithmic calculation) skills
	HO: Questions involving higher-order (application, analysis, and synthesis) skills
Question Format (QF)	CL: Questions involving mathematical calculations or algebraic manipulations, often using simple algorithms
	MC: Questions where students are asked to choose answers from a list of provided choices (multiple-choice, true-false, matching)
	SA: Questions where students are asked to provide short-answers independently
Question Placement (QP)	IC: In-chapter questions appearing within the written text (worked-out examples, practice problems)
	EC: Questions appearing at the end of the chapter (review problems, exercises)
	TB: Questions appearing in the stand-alone test-bank written exclusively for teachers
Question Type (QT)	QL: Qualitative conceptual questions that do not require mathematical calculations/algorithms to be answered
	QN: Quantitative questions that require mathematical calculations/algorithms to be performed to answer the question
Representation (RP)	MS: Questions requiring answers focused at the macroscopic representation (observations based on the five senses)
	PT: Questions requiring answers focused at the particulate representation (behaviours and interactions of gaseous particles)
	SB: Questions requiring answers focused at the symbolic representation (use of symbols to stand for abstract ideas; *e.g.*, balanced chemical equations, mathematical formulas, graphs, etc.)

Book Type   Although several chemical education researchers have analysed secondary textbooks (Pedrosa and Dias, 2000; Drechsler and Schmidt, 2005; Gkitzia *et al.*, 2011; Bergqvist *et al.*, 2013) and introductory college-level textbooks (Suidan *et al.*, 1995; Sanger and Greenbowe, 1999; Pedrosa and Dias, 2000; Dávila and Talanquer, 2010; Österlund *et al.*, 2010; Gkitzia *et al.*, 2011; Pappa and Tsaparlis, 2011), few have compared the content and structure of these two types of books to each other. This comparison provides important information for both secondary and college instructors. Since many secondary school instructors are concerned about preparing their students to be successful in college chemistry courses (Tai *et al.*, 2006; Sadler and Tai, 2007), knowing how the textbooks (and, presumably, the instructional

format) for these two classes are different will help these instructors to adjust the instructional methods used in their secondary chemistry classes.   Similarly, knowing how the textbooks for secondary and college chemistry courses differ can help college instructors recognize content material that may not have been previously seen by the students enrolled in their classes. Pedrosa and Dias (2000) analysed the written text of five secondary and introductory college chemistry textbooks as possible sources of student alternative conceptions in equilibrium, and Gkitzia *et al.* (2011) conducted a detailed analysis of how five secondary and introductory college chemistry textbooks depict chemical representations.   However, neither study compared how the secondary and introductory college textbooks differed from each other.   Staver and Jumpe (1993) compared how seven high school (secondary), seven college preparatory, and fifteen introductory college chemistry textbooks introduce and discuss the concept of the mole.   They found that most high school textbooks define the mole as 6.02×10^{23} particles, while a majority of the college textbooks define the mole in terms of carbon-12.   They also found that most textbooks (both high school and college) described Avogadro's constant as an experimentally-determined value and most introduced the mole concept as a way to count particles that are too small to be directly weighed.

Cognitive Skill   The prevailing goal of most chemistry instruction is to improve students' critical-thinking, problem-solving, and decision-making skills (Zoller *et al.*, 1995 and 2002). These skills are collectively referred to as their higher-order cognitive skills. Zoller *et al.* (1995 and 2002) showed that college chemistry students' success at answering questions was highest for algorithmic (lower-order calculation) questions, intermediate for lower-order conceptual questions, and lowest for higher-order conceptual questions.   Zoller and Tsaparlis (1997) found that secondary chemistry students had more success in answering lower-order questions (recall and simple algorithmic calculations) than in answering higher-order qualitative and quantitative questions.   Stamovlasis *et al.* (2005) noted that secondary chemistry students' success in answering questions decreased from lower-order algorithmic calculations (highest) to lower-order conceptual questions to higher-order calculations (lowest). Finally, Papaphotis and Tsaparlis (2008) showed that secondary chemistry students performed better on algorithmic questions compared to conceptual questions based on the topic of quantum chemistry. Taken together, these results indicate that students perform better on lower-order *versus* higher-order questions.

Several of these studies also showed that students' abilities to solve lower-order questions involving simple recall or algorithmic calculations do not appear to be correlated to their ability to solve higher-order conceptual or calculation questions (Zoller *et al.*, 1995; Zoller and Tsaparlis, 1997; Zoller *et al,.* 2002; Stamovlasis *et al.*, 2005). These results do not suggest that the abilities to answer lower- and higher-order questions are incompatible in a learner or mutually exclusive, but simply that success in answering one type of questions does not appear to affect success in answering the other type of questions.

Zoller *et al.* (1995) concluded that traditional lecture-format instruction is incompatible with most higher-order learning, that

student success with simple recall questions and algorithmic calculations (typical of most exam questions appearing in secondary and college chemistry tests) does not demonstrate the development of higher-order skills nor does it indicate student mastery of chemistry concepts, and that college instruction needs to stop focusing on algorithmic and lower-order skill development and start focusing on higher-order oriented teaching strategies and pedagogies. Zoller and Tsaparlis (1997) and Stamovlasis *et al.* (2005) expressed the need for examinations to include both lower-order and higher-order questions in order to identify students operating at the lower- *versus* higher-order, to encourage students to develop their higher-order cognitive skills, and to allow instructors to determine whether their instructional methods are leading to improvements in their students' abilities to answer higher-order questions.

Presumably, textbooks whose questions focus on lower-order skills would be detrimental to the development of students' higher-order cognitive skills. Shepardson and Pizzini (1991) evaluated the cognitive skills needed to answer the questions in several junior high science textbooks, and found that 79% of these questions were at the input (lower-order) level, while 15% were at the lower-order processing level and 7% were at the output level (higher-order skills). They found no difference in the proportion of these questions based on the different textbooks analysed or on the science discipline (earth, general, life, or physical science) of the textbook. Pappa and Tsaparlis (2011), in their analysis of the questions in the intramolecular and intermolecular bonding chapters of ten chemistry textbooks, noted that 56% of the questions in these chapters involved declarative knowledge at the basic (recall) level, 30% of the questions involved procedural knowledge at the basic level, and 14% involved higher-level declarative knowledge. Dávila and Talanquer (2010) categorized the end-of-chapter questions in three introductory college chemistry textbooks from the United States according to Bloom's taxonomy (Bloom, 1956). These textbooks had 64% of their questions at the knowledge, comprehension, and application levels (collectively viewed as the lower-order levels) and 36% of their questions at the analysis, synthesis, and evaluation levels (the higher-order levels).

Question Format Johnstone and Ambusaidi (2000) discussed the advantages and disadvantages of fixed response testing (which can include matching or true-false, but most commonly involves multiple-choice questions). Although these tests appear to be objective, only the scoring truly objective, and tests involving these types of questions have disadvantages that must be considered. These issues include the possibility that students can guess the correct answer, the nature of the distractors chosen, negative discrimination, and the fact that changing the order of distractors in a question or the order of questions within a test can change student scores on these tests. They also report that surface-level learners tend to favour fixed response tests while independent learners favour free response tests, and they cite a study by Tamir (1990) showing that 30% of students choosing the right answer in a multiple-choice question chose it for an incorrect reason. In a subsequent paper, Johnstone and Ambusaidi (2001) described other forms of fixed-response questions that attempt to address the disadvantages of traditional multiple-choice questions mentioned in their previous work.

Recently, the Examinations Institute of the American Chemical Society has reported an effort to provide students with partial credit on their examinations instead of simply scoring the question right or wrong (Grunert *et al.*, 2013). They found that students on either extreme did not see much difference in their scores with or without partial credit; however, the middle-performing students showed the most potential for change but these changes were evenly split (half showing increased scores while the other half showing decreased scores).

Danili and Reid (2005) assessed secondary chemistry students using three different assessment formats: Multiple-choice questions, short-answer questions, and structural communication grid questions (Johnstone and Ambusaidi, 2001). These comparisons showed correlation coefficients ranging from 0.30 to 0.71. Students routinely performed better on the multiple-choice questions compared to the short-answer or the grid questions. The researchers noted that each form of the test appeared to be measuring different skills (recognizing answers for the multiple-choice questions, deriving answers for the short-answer questions, etc.), which might explain the lower than expected correlations. In any case, the researchers concluded that the best student found by one method is not necessarily the best student found by another method, and that instructors must take into account the validity of these tests and should consider what information is being tested by each question format. Zoller *et al.* (1995, 2002) reported that students' scores were highest for algorithmic questions (calculations), in the middle for lower-order conceptual questions (multiple-choice/true-false questions), and lowest for higher-order conceptual questions (short-answer explanations). It is possible that the format of these questions could be partially responsible for this ranking in student performance. In their analysis of the intramolecular and intermolecular bonding questions in ten chemistry textbooks, Pappa and Tsaparlis (2011) noted that almost all of the questions were categorised as closed-type (with more short-answer than multiple-choice questions), and most involved declarative knowledge at the recall level.

Question Placement Dávila and Talanquer (2010) categorised the end-of-chapter questions in the three top-selling U.S. introductory college chemistry textbooks. Each of these textbooks had a majority of their questions in only five of the 14 subcategories (recalling, explaining, inferring-predicting, executing-qualitative, and executing-quantitative); however, the three textbooks had different distributions of questions within these five categories. Few of these questions asked students to relate or convert between the particulate, macroscopic, and particulate views (Johnstone, 1993; Johnstone, 2010). Although this study analysed the end-of-chapter questions, it did not compare them to the questions placed within the text of the chapter or to the questions in the test banks. In fact, our search of the chemical education literature did not uncover any research studies that have compared the in-chapter, end-of-chapter, or test bank questions within a textbook. In their introduction, Pappa and Tsaparlis (2011) mentioned three types of assessment questions: (i) *Diagnostic* questions, used to diagnose students' learning difficulties and their causes, (ii) *formative* questions, used during a course to inform students and/or the teacher of learning gains, and (iii) *summative* questions, used to provide the

overall evaluation of the student by the teacher at the end of the course. While these three types of assessment aren't a perfect match for the three types of questions discussed in our study, they do mirror the goals of in-chapter questions (to help students learn the material presented within the textbook and help students diagnose any problems they might have), end-of-chapter questions (often used as homework assignments by teachers to assess student learning gains), and test bank questions (often used for chapter tests or final examinations by teachers to assess student learning and assign grades).

Question Type Several chemical education researchers have compared students' performance on quantitative *versus* qualitative questions. Most of these studies (Nurrenbern and Pickering, 1987; Pickering, 1990; Sawrey, 1990; Zoller *et al.*, 1995 and 2002) have shown that students' performance on quantitative calculations that can be solved using simple algorithms is much better than their performance on qualitative questions that require a conceptual understanding of chemistry topics. Unfortunately, the qualitative/quantitative nature of these questions is not the only difference in these two types of questions. The algorithmic questions are quantitative in nature but require lower-order cognitive skills to solve, while the conceptual questions are qualitative in nature but require higher-order cognitive skills to solve. In addition, many of these studies (Nurrenbern and Pickering, 1987; Pickering, 1990; Sawrey, 1990; Zoller *et al.*, 1995) present the quantitative questions as symbolic-level calculations using words (verbal) while they present the qualitative questions as particulate-level question involving pictures (visual). Therefore, it can be difficult to determine whether the quantitative/qualitative nature of the question is responsible for these differences or whether it is an issue of algorithmic/conceptual, symbolic/particulate, or verbal/visual differences in these two types of questions. Holme and Murphy (2011) compared students' performance on algorithmic and conceptual questions on the 2005 Paired-Questions First-Semester General Chemistry Exam and the 2007 Paired-Questions Second-Semester General Chemistry Exam. They found that the students' overall performance on the conceptual and algorithmic paired questions on these tests were roughly equivalent. Haláková and Proksa (2007) compared students' performance on two different types of qualitative conceptual questions—verbal and visual (pictorial). They found that their students had similar performances on the verbal (39%) and the visual questions (35%); these results led the researchers to conclude that the use of particulate pictures in the visual questions did not substantially affect their performance on the visual questions compared to the verbal questions that did not use these particulate pictures.

Representation Chemists often use three representations—the macroscopic, particulate, and symbolic representations—to describe and explain chemical processes and phenomena (Johnstone, 1993; Gilbert and Treagust, 2009; Johnstone, 2010; Talanquer, 2011). The macroscopic representation is based on observations made using the five senses; the particulate representation describes the behaviour of atoms, molecules, and ions; and the symbolic representations uses symbols to stand for more abstract concepts involving chemical and mathematical relationships. As mentioned in the previous discussion on question type, several chemical education researchers have noted that students are better at answering symbolic algorithmic questions than they are at answering particulate conceptual questions (Nurrenbern and Pickering, 1987; Pickering, 1990; Sawrey, 1990; Zoller *et al.*, 1995). Unfortunately, we cannot definitively attribute these differences to the symbolic or particulate representations used in these questions.

Gkitzia *et al.* (2011) created a list of criteria for analysing and evaluating how chemistry textbooks depict chemical representations in their visual images. These criteria included the type of representation depicted in the visual, the interpretation of the surface features in the visual, how related the written text was to the visual representations, the presence and properties of captions accompanying the visuals, and the correlation between multiple representations presented in a single visual. The authors also presented data from the chemistry textbook used in 10th-grade classrooms in Greece based on these criteria. Overall, the images in the Greek 10th-grade book used the macroscopic representation 35% of the time, the particulate representation 28% of the time, and the symbolic representation 37% of the time. About 34% of the images used multiple, hybrid, or mixed representations involving more than one representation. Bergqvist *et al.* (2013) analysed the representations of chemical bonding models used by five secondary chemistry textbooks in Sweden to evaluate how the chemistry content was presented and to look for possible sources of students' misconceptions. The authors found that all of the textbooks still include descriptions that previous chemical education research has indicated could lead to student conceptual difficulties or misconceptions. Kumi *et al.* (2013) created a rubric to evaluate whether Newman Projections and Fischer projections (symbolic images) are accurately depicted in seven organic chemistry textbooks from the United States and how these textbooks help students interpret these symbolic images at the particulate level. The authors concluded that although many of the textbook images focus on the relationship between the symbolic images of the Newman and Fischer projections and particulate drawings of the molecules, students would be better served if the textbooks placed more emphasis on the introduction and conventions of these images, and the dynamic behaviour of the molecules being depicted. They also encouraged instructors to help students make these connections by actually using chemical models themselves instead of watching the instructor use them. Haláková and Proksa (2007) found that their students' performance on verbal (non-particulate) and visual (particulate) conceptual questions were similar, suggesting that the presence or absence of particulate pictures in the conceptual questions did not affect their ability to answer these conceptual questions.

Research questions

The goal of this study was to evaluate the types of questions presented in the gas law chapters of secondary and introductory college chemistry textbooks from the United States. Since this was our first study comparing the distribution of questions in chemistry textbooks, it was intended to be a descriptive study that describes how these distributions differ with respect to the different variables being compared. We found no prior studies making similar comparisons, and as such, we had no *a priori* research hypotheses based on a particular theoretical framework

or existing studies.

All of the quantitative (mathematical) questions in these chapters were categorised as using the symbolic representation, but the qualitative (non-mathematical) questions were categorised as using all three representations (macroscopic, particulate, and symbolic). So, we decided to evaluate the quantitative and qualitative questions separately (eliminating the Representation variable for the quantitative study since they were all symbolic questions). However, we were also interested in how the distribution of quantitative and qualitative questions are different, so we compared these two sets of questions to each other (eliminating the Representation variable for statistical reasons).

(1) For the qualitative questions the in gas law chapters, are there any significant associations among the five variables of Book Type, Cognitive Skill, Question Format, Question Placement, and Representation?

(2) For the quantitative questions the in gas law chapters, are there any significant associations among the four variables of Book Type, Cognitive Skill, Question Format, and Question Placement?

(3) For the combination of qualitative and quantitative questions the in gas law chapters, are there any significant associations among the five variables of Book Type, Cognitive Skill, Question Format, Question Placement, and Question Type?

Methods

Data collection and analysis

The questions from the gas law chapters of four secondary chemistry textbooks and their test banks (Davis *et al.*, 2002a and 2002b; Dingrando *et al.*, 2002a and 2002b; Phillips *et al.*, 2002a and 2002b; Wilbraham *et al.*, 2002a and 2002b) and introductory college chemistry textbooks and their test banks (Masterson and Hurley, 2000; Treichel, 2000; Zumdahl and Zumdahl, 2003; Zumdahl *et al.*, 2003; Zumdahl, 2004; Zumdahl and DeCoste, 2004; Brown *et al.*, 2006; Laurino *et al.*, 2006) were reviewed and analysed. For questions with multiple parts, each part was separately analysed. These books were chosen because they were approved for use in high school or Advanced Placement (which requires the use of college textbooks) chemistry courses by the Arkansas State Board of Education.

Each question ($N = 2313$) was categorized according to the six variables by the first author. To determine the reliability of the categorisation scheme, the second author categorised the end-of-chapter questions for one of the introductory college textbooks (Brown *et al.*, 2006) and the results from the two authors were compared. The inter-rater reliability was 0.86. The authors discussed all discrepancies and amended the categorisation scheme. Using this new scheme, both authors categorised the in-chapter questions for one of the secondary textbooks (Davis *et al.*, 2002a) and the inter-rater reliability was 0.96.

Data analysis

Each question was categorised according to the six variables, and the frequency (number) of questions appearing in the same category of all six independent variables was tallied. These nonparametric frequency data were analyzed using loglinear analysis (Tabachnick and Fidell, 1996), which is an extension of

the statistical method of multiway frequency analysis (Engleman, 2002). We used the LOGLIN function of the statistical program SYSTAT 10.2 (Engleman, 2002) to perform the loglinear analyses in this study (Sanger, 2008). Significant associations occur when the computer determines that the frequency data are not homogeneously distributed among the individual cells. To answer the three research questions, we performed three loglinear analysis calculations: (1) Using only the qualitative (QL) questions from the Question Type variable ($N = 740$), (2) Using only the quantitative (QN) questions from the Question Type variable ($N = 1573$), and (3) Using both the QL and QN questions ($N = 2313$). Since the qualitative and quantitative loglinear analyses were performed by collapsing the Question Type variable, this variable was not used in these analyses. All of the quantitative questions involving calculations were placed in the symbolic category of the Representation variable (none were categorised as macroscopic or particulate), so the Representation variable was not used in the quantitative or the combined loglinear analyses. Table 2 contains the results of the qualitative loglinear analysis, Table 3 contains the results of the quantitative loglinear analysis, and Table 4 contains the results of the combined loglinear analysis.

According to Tabachnick and Fidell (1996), multiway frequency analyses make no assumptions about population distributions and are remarkably free of limitations. However, they caution about having too many variables since it can be very difficult to explain (or understand) four-, five-, or higher-order

Table 2 Loglinear analysis output for the associations between the independent variables for the qualitative questions

Source	*df*	χ^2 value	*p* value
Book Type (BT)	1	9.69	0.00[a]
Cognitive Skill (CS)	1	1.52	0.22
Question Format (QF)	1	23.79	0.00[a]
Question Placement (QP)	2	20.85	0.00[a]
Representation (RP)	2	2.78	0.25
BT × CS	1	0.03	0.86
BT × QF	1	3.96	0.05[a]
BT × QP	2	11.69	0.00[a]
BT × RP	2	1.53	0.47
CS × QF	1	0.01	0.93
CS × QP	2	3.58	0.17
CS × RP	2	3.03	0.22
QF × QP	2	177.59	0.00[a]
QF × RP	2	1.44	0.49
QP × RP	4	0.40	0.98
BT × CS × QF	1	0.23	0.63
BT × CS × QP	2	0.97	0.61
BT × CS × RP	2	3.22	0.20
BT × QF × QP	2	11.78	0.00[a]
BT × QF × RP	2	0.01	0.99
BT × QP × RP	4	0.63	0.96
CS × QF × QP	2	1.80	0.41
CS × QF × RP	2	1.61	0.45
CS × QP × RP	4	1.85	0.76
QF × QP × RP	4	4.47	0.35
BT × CS × QF × QP	2	0.79	0.67
BT × CS × QF × RP	2	2.05	0.36
BT × CS × QP × RP	4	2.23	0.69
BT × QF × QP × RP	4	2.65	0.62
CS × QF × QP × RP	4	2.57	0.63
BT × CS × QF × QP × RP	4	1.64	0.80

[a]$p < 0.05$ corresponds to a significant association between these variables

This journal is © The Royal Society of Chemistry [year]

Chem. Educ. Res. Pract., [year], **[vol]**, 00–00 | **5**

Table 3 Loglinear analysis output for the associations between the independent variables for the quantitative questions

Source	df	χ^2 value	p value
Book Type (BT)	1	0.10	0.75
Cognitive Skill (CS)	1	0.03	0.86
Question Format (QF)	1	127.16	0.00[a]
Question Placement (QP)	2	35.80	0.00[a]
BT × CS	1	0.55	0.46
BT × QF	1	1.52	0.22
BT × QP	2	1.75	0.42
CS × QF	1	0.02	0.89
CS × QP	2	0.18	0.91
QF × QP	2	812.73	0.00[a]
BT × CS × QF	1	0.00	0.99
BT × CS × QP	2	1.45	0.48
BT × QF × QP	2	7.12	0.03[a]
CS × QF × QP	2	0.96	0.62
BT × CS × QF × QP	2	0.06	0.97

[a]$p < 0.05$ corresponds to a significant association between these variables

Table 4 Loglinear analysis output for the associations between the independent variables for the qualitative and quantitative (combined) questions

Source	df	χ^2 value	p value
Book Type (BT)	1	5.76	0.02[a]
Cognitive Skill (CS)	1	0.13	0.72
Question Format (QF)	1	48.25	0.00[a]
Question Placement (QP)	2	21.41	0.00[a]
Question Type (QT)	1	4.31	0.04[a]
BT × CS	1	0.06	0.80
BT × QF	1	3.59	0.06
BT × QP	2	7.93	0.02[a]
BT × QT	1	4.01	0.05[a]
CS × QF	1	0.04	0.84
CS × QP	2	1.33	0.51
CS × QT	1	0.02	0.90
QF × QP	2	1306.69	0.00[a]
QF × QT	1	6.51	0.01[a]
QP × QT	2	0.80	0.67
BT × CS × QF	1	0.19	0.66
BT × CS × QP	2	2.08	0.35
BT × CS × QT	1	0.61	0.44
BT × QF × QP	2	15.37	0.00[a]
BT × QF × QT	1	0.04	0.84
BT × QP × QT	2	1.78	0.41
CS × QF × QP	2	2.02	0.36
CS × QF × QT	1	0.15	0.70
CS × QP × QT	2	2.12	0.35
QF × QP × QT	2	0.47	0.79
BT × CS × QF × QP	2	0.59	0.75
BT × CS × QF × QT	1	0.21	0.65
BT × CS × QP × QT	1	0.09	0.96
BT × QF × QP × QT	2	0.54	0.76
CS × QF × QP × QT	2	0.18	0.91
BT × CS × QF × QP × QT	2	0.32	0.85

[a]$p < 0.05$ corresponds to a significant association between these variables

difficult to explain (or understand) four-, five-, or higher-order associations. Tabachnick and Fidell also described four practical issues affecting the power of loglinear results that should be addressed. (1) All of the categories must be mutually exclusive, so that each question appears in only one cell; this requirement has been met in the three analyses. (2) There should be at least five times as many questions as cells in the study; in the three loglinear analyses in this study, there are an average of 10.3, 65.5, and 48.2 questions per cell, respectively. (3) When cases are rare (less than five), the marginal frequencies may not be evenly distributed. The expected cell frequencies for all two-way associations should be examined to assure that all are greater than zero, and no more than 20% are less than five. The two-way associations contained frequencies of zero for 1 of the 57 cells (2%) in the qualitative loglinear analysis, 2 of the 30 cells (7%) in the quantitative loglinear analyis, and 1 of the 48 cells (2%) in the combined loglinear analysis; no other cells had frequencies less than five. For each analysis, these zero frequencies occurred in the QF × QP association—none of the eight textbooks contained any multiple-choice questions (qualitative or quantitative) within the chapter text (MC/IC) and none contained multiple-choice quantitative questions at the end of the chapter (MC/EC). The only way to address this loss of power would be to analyse more textbooks (Tabachnick and Fidell, 1996). However, an analysis of the in-chapter and end-of-chapter gas law questions in two additional secondary and two additional introductory college chemistry textbooks showed no multiple-choice questions within the chapters or at the end of the chapters. Therefore, we have decided to accept the loss of power associated with having zero entries in these cells. (4) When performing unsaturated tests, there may be substantial differences between observed and expected frequencies that make it impossible for the proposed model to adequately fit the data. This is not an issue in this study since we performed saturated tests.

Results

The output for the results of the qualitative, the quantitative, and the combined loglinear analyses appear in Tables 2-4, respectively. For each association, SYSTAT calculates standard normal deviate scores for each cell within the test. These values represent z-scores that can be used to determine the influence of each cell on the overall effect (Tabachnick and Fidell, 1996). A standard normal deviate score above +1.96 means that the observed frequency of a cell is significantly higher than would be expected if the questions were homogeneously distributed, and a standard normal deviate score below –1.96 means that the observed frequency of a cell is significantly lower than would be expected if the questions were homogeneously distributed, at the $p = 0.05$ level.

Significant associations for the qualitative questions

Although this loglinear analysis included five variables, only three variables (Book Type, Question Format, and Question Placement) were involved in the significant associations; the distribution of qualitative questions in these textbooks were not significantly different based on the cognitive skill (lower-order *versus* higher-order questions) or representation (macroscopic *versus* particulate *versus* symbolic) used in these questions.

• The BT association indicated that there were more qualitative questions in the secondary textbooks (an average of 109 questions per book) than in the introductory college textbooks (76 questions per book, on average), $\chi^2(1) = 9.69$, $p < 0.01$.

• The QF association indicated that there were more short-answer questions (59% of the total) than multiple-choice questions (41% of the total) in these textbooks, $\chi^2(1) = 23.79$, $p < 0.01$.

• The QP association indicated that there were more qualitative

questions in the test banks (49% of the total) and fewer in-chapter questions (13% of the total), $\chi^2(2) = 20.85$, $p < 0.01$.

• The BT × QF association indicated that the secondary books have fewer multiple-choice questions and more short-answer questions than the introductory college textbooks, $\chi^2(1) = 3.96$, $p = 0.05$. However, this is contradicted by the actual data—the qualitative questions in the secondary textbooks were 46% multiple-choice and 54% short-answer while those in the introductory college textbooks were 34% multiple-choice and 66% short-answer.

This contradiction is due to the way loglinear calculations are performed (Tabachnick and Fidell, 1996). With the saturated model, the highest-order association is tested first then the next highest-order associations are tested with the effect of the highest-order association removed (partialed out). The process continues from highest- to lowest-order associations with the effects of all higher-order associations removed before a test is performed. Because of this partialing effect, the results of a significant one- or two-way association test may contradict the actual distribution of questions. Since this significant association contradicts the actual distribution of questions in the study, we will not discuss this association further.

• The BT × QP association indicated that the secondary books had fewer (22%) qualitative questions at the end of the chapter and more (59%) in the test banks, while the introductory college books had more (60%) at the end of the chapter and fewer (34%) in the test banks, $\chi^2(2) = 11.69$, $p < 0.01$.

• The QF × QP association indicated that the textbooks had fewer qualitative questions within the chapters (0%) and at the end of the chapters (2%) in the multiple-choice format, and more (83%) qualitative questions in the test banks using the multiple-choice format, $\chi^2(2) = 177.59$, $p < 0.01$.

• The BT × QF × QP association indicated that the secondary books had more end-of-chapter questions in the multiple-choice format (5%) compared to the introductory college textbooks (0%). The secondary books also had fewer test bank questions in the multiple-choice format (76%) compared to the introductory college textbooks (100%), $\chi^2(2) = 11.78$, $p < 0.01$.

Significant associations for the quantitative questions

Although this loglinear analysis included four variables, only three variables (Book Type, Question Format, and Question Placement) were involved in the significant associations, and these are the same three variables involved in the significant associations for the qualitative questions. The distribution of qualitative questions in these textbooks were not significantly different based on the cognitive skill used in the question (lower-order or higher-order); the Representation variable was not used in this loglinear analysis because all of the quantitative questions involved in this analysis were categorised as symbolic questions.

• The QF association indicated that there were more quantitative questions posed as calculations (77% of the total) than as multiple choice (23% of the total), $\chi^2(1) = 127.16$, $p < 0.01$.

• The QP association indicated that there were more quantitative questions in the test banks (28% of the total) and fewer in-chapter questions (20% of the total), $\chi^2(2) = 35.80$, $p < 0.01$.

• The QF × QP association indicated that fewer of the quantitative questions within the chapters (0%) and at the end of the chapters (0%) were in the multiple-choice format, but more

(83%) of quantitative questions in the test banks were in the multiple-choice format, $\chi^2(2) = 812.73$, $p < 0.01$.

• The BT × QF × QP association indicated that the introductory college books had more of their test bank questions in the multiple-choice format (97%) compared to the secondary textbooks (63%), $\chi^2(2) = 7.12$, $p = 0.03$.

Significant associations for the combination of qualitative and quantitative questions

As with the other two loglinear analyses, the distribution of all questions in these textbooks were not significantly different based on the Cognitive Skill variable (lower-order *versus* higher-order). The other four variables (Book Type, Question Format, Question Placement, and Question Type) were involved in the significant associations.

• The residuals in the BT association indicated that there were more questions in the secondary textbooks (an average of 259 questions per book) compared to the introductory college textbooks (an average of 320 questions per book), $\chi^2(1) = 5.76$, $p = 0.02$. However, the residuals indicate a difference that is counterintuitive (i.e., 259 is greater than 320). Since this significant association contradicts the actual distribution of questions in the study, we will not discuss this association further.

• The QF association indicated that there were more short-answer qualitative and calculation-format quantitative questions (71% of the total) than multiple-choice questions (29% of the total), $\chi^2(1) = 48.25$, $p < 0.01$.

• The QP association indicated that there were more questions in the test banks (35% of the total) and fewer in-chapter questions (18% of the total) in these textbooks, $\chi^2(2) = 21.41$, $p < 0.01$.

• The QT association indicated that there were more quantitative questions (197 quantitative questions per book, on average) and fewer qualitative questions (an average of 93 qualitative questions per book) in these textbooks, $\chi^2(1) = 4.31$, $p = 0.04$.

• The BT × QP association indicated that the secondary books had fewer (31%) questions at the end of the chapter and more (42%) in the test banks, while the introductory college books had more (61%) questions at the end of the chapter and fewer (29%) in the test banks, $\chi^2(2) = 7.93$, $p = 0.02$.

• The QF × QP association indicated that fewer of the questions within the chapters (0%) and at the end of the chapters (<1%) were in the multiple-choice format, and more (83%) of questions in the test banks were in the multiple-choice format, $\chi^2(2) = 1306.69$, $p < 0.01$.

• The BT × QT association indicated that the secondary books had more qualitative and fewer quantitative questions (an average of 109 and 150 per book, respectively) than the introductory college books (which had an average of 76 qualitative and 244 quantitative questions per book), $\chi^2(1) = 4.01$, $p = 0.05$.

• The QF × QT association indicated that more of the qualitative questions used the multiple-choice format (42% of the total) compared to the quantitative questions (23% of the total), $\chi^2(1) = 6.51$, $p = 0.01$.

• The BT × QF × QP association indicated that the secondary books had more end-of-chapter questions in the multiple-choice format (2%) compared to the introductory college textbooks (0%). The secondary books also had fewer test bank questions in the multiple-choice format (71%) compared to the introductory college textbooks (98%), $\chi^2(2) = 15.37$, $p < 0.01$.

Discussion

The loglinear analyses for the qualitative questions, the quantitative questions, and the combination of qualitative and quantitative questions in the gas law chapters of these eight books showed significant associations for the variables of Book Type, Question Format, Question Placement, and Question Type; however, no significant associations were found for the variables of Cognitive Skill or Representation.

Secondary and introductory college textbooks

The significant associations in the loglinear analyses performed in this study demonstrated three major differences in the distribution of questions in the secondary *versus* the introductory college textbooks. The BT association (qualitative) and the BT × QT association (combined) indicated that the secondary textbooks had more qualitative questions and fewer quantitative calculation questions than the introductory college textbooks. The BT × QP associations (qualitative, combined) showed that the secondary textbooks had more questions in the test banks and fewer questions at the end of the chapters than the introductory college textbooks. The BT × QF × QP associations for all three analyses revealed that the secondary textbooks had a higher percentage of multiple-choice questions at the end of the chapters and a lower percentage of multiple-choice questions in the text banks compared to the introductory college textbooks.

Unfortunately, it is difficult to attribute these differences between the distribution of questions in the secondary and introductory college chemistry textbooks to a difference in the types of learning that secondary and introductory college teachers value from their students. In fact, it is common for secondary chemistry books in the United States to be written by college professors and not by high school instructors (the intended users of these books). Since this is the first study to identify key differences in the distribution of questions in secondary and introductory college chemistry textbooks, it's not obvious that the secondary and introductory college textbook authors are even aware of these differences, or whether these differences are in fact deliberate.

Lower-order and higher-order cognitive skills

The three loglinear analyses showed no significant associations involving the Cognitive Skill variable. This indicates that the lower- and higher-order questions were homogeneously distributed with respect to the other variables. In this study, 1278 of the 2313 questions (55%) were categorised as lower-order and 1035 questions (45%) were categorised as higher-order. Zoller *et al.* (1995) has argued that chemistry instruction needs to stop focusing on algorithmic and lower-order skills and start focusing on higher-order oriented teaching strategies and pedagogies. Shepardson and Pizzini (1991) found that 93% of the questions in junior high science textbooks were lower-order and 7% were at higher-order, Pappa and Tsaparlis (2011) found that 86% of the questions in the intramolecular and intermolecular bonding chapters of ten chemistry textbooks were lower-order and 14% were higher-order, and Dávila and Talanquer (2010) found that 64% of the end-of-chapter questions in three introductory college chemistry textbooks were lower-order and 36% were higher-order. Collectively, these results corroborate the assertion from

Zoller *et al.* (1995) that instruction (and the textbooks used as part of instruction) are still focusing on lower-order instead of higher-order skills. Our analysis also shows more lower-order questions in these eight textbooks compared to higher-order questions; however, this difference is much smaller than that seen in the other studies. These studies looked at different book types (junior high *versus* college books *versus* a mix of high school and college books) and different chemistry topics (junior high science *versus* intramolecular and intermolecular bonding *versus* college chemistry *versus* gas laws). Unfortunately, it is difficult to attribute these differences to any one variable, and these differences could also be the result of different categorisation schemes used to identify lower-order and higher-order questions.

Calculation, multiple-choice, and short-answer formats

All three loglinear analyses showed the same three significant associations regarding the Question Format variable. The QF associations showed that the textbooks had fewer questions in the multiple-choice format and more questions in the short answer or calculation format. The QF × QP associations revealed that there were fewer multiple-choice questions within the chapters and at the end of the chapters and more multiple-choice questions in the text banks compared to short-answer or calculation questions. The BT × QF × QP associations indicated that differences found for the QF × QP associations described in the last sentence were more extreme for the introductory college textbooks than for the secondary textbooks. In addition, the analysis for the combined questions showed a QF × QT association: A higher proportion of the qualitative questions used the multiple-choice format compared to the short-answer format, while a lower proportion of the quantitative questions used the multiple-choice format compared to the calculation format.

Of the 668 multiple-choice questions appearing the eight textbooks, none of them appeared within the chapters and only five appeared in the end-of-chapter questions; over 99% of the multiple-choice questions appeared in the test banks, and they represented 83% of all questions asked in the test banks (8% were short-answer and 9% were calculations). Danili and Reid (2005) reported that using different question formats on a test leads to measures of different students skills, and that the best student found using one question format is not necessarily the best student found using another format. It can be implied from that study that learning to solve questions using one format will not necessarily assist students in answering questions using another format. If this is true, the distribution of questions found from our study can particularly troublesome for students using these textbooks. The in-chapter and end-of-chapter questions in these textbooks have essentially no multiple-choice questions, so as part of their instruction in chemistry these students only learn to answer short-answer qualitative questions and calculation-based quantitative questions. If their instructor chooses to ask multiple-choice questions on the chapter tests or final examination (on their own, or using the textbook's test bank, which contains predominantly multiple-choice questions) or if these students are asked to take a standardized multiple-choice test like those from the American Chemical Society Examinations Institute (Holme and Murphy, 2011; Schroeder *et al.*, 2012; Grunert *et al.*, 2013), these students may be at a disadvantage because they were never

In-chapter, end-of-chapter, and test-bank questions

given the chance to answer multiple-choice questions as part of their instruction. If students are expected to answer multiple-choice questions as part of their summative assessments, then the textbooks should provide examples of multiple-choice questions at the end of the chapters for students to practice.

In-chapter, end-of-chapter, and test-bank questions

The significant associations in this study demonstrated four major differences in the distribution of questions based on the Question Placement variable. The QP associations for all three analyses indicated that the textbooks had fewer in-chapter questions and more test-bank questions. The BT × QP associations (qualitative, combined) revealed that the secondary textbooks had fewer end-of-chapter questions and more test-bank questions compared to the introductory college textbooks. The QF × QP associations for all three analyses indicated that there were fewer multiple-choice questions within the chapters and at the end of the chapters and more multiple-choice questions in the text banks compared to short-answer or calculation questions. The BT × QF × QP associations showed that differences found for the QF × QP associations were more extreme for the introductory college textbooks compared to the secondary textbooks.

Since the textbook authors' goals for the in-chapter, end-of-chapter, and test-bank questions are different, it should not be surprising that the distribution of questions in these three spots of the textbook is different. In-chapter questions are used to help students learn new material and help them diagnose any problems they might have. There are fewer of these questions in the textbooks and these questions are more likely to use the short-answer or calculation format than the multiple-choice format. Presumably, students would need to see fewer examples of these questions to diagnose any problems they might have and this diagnosis would be easier if the answer is depicted as a short-answer or a calculation. End-of-chapter questions are used as homework assignments to assess students' learning gains. It is not clear why secondary textbooks would have more of these questions than introductory college textbooks or why these questions do not appear in multiple-choice format. Test-bank questions are used to assess student learning and to assign grades. There are more of these questions in the textbooks, and this could be done to provide instructors with a larger number of questions for their examinations. It is not clear why secondary textbooks would have fewer of these questions than introductory college textbooks or why these questions appear in predominantly multiple-choice format.

Qualitative and quantitative questions

The Question Type variable was only used in the third loglinear analysis involving both the qualitative and quantitative questions. The QT association revealed that the textbooks had fewer qualitative questions and more quantitative questions. The BT × QT association indicated that secondary textbooks had more qualitative questions and fewer quantitative questions compared to the introductory college textbooks. The QF × QT association showed that the qualitative questions had a higher proportion of multiple-choice questions compared to the quantitative questions, and that there were a higher proportion of quantitative questions in the calculation format and a lower proportion of qualitative questions in the short-answer format.

Several chemical education researchers (Nurrenbern and Pickering, 1987; Pickering, 1990; Sawrey, 1990; Zoller *et al.*, 1995) have shown that students' performance on quantitative calculations using simple algorithms is better than their performance on qualitative questions requiring conceptual understanding. Prior to these studies, many chemistry instructors in the United States asked their students to solve only quantitative questions and assumed that mastery of these questions implied a mastery of the chemistry concepts related to these topics. However, these studies forced chemistry instructors to re-evaluate this assumption. As a result, many instructors relinquished their reliance on quantitative questions to favour a mix of qualitative and quantitative questions. However, this study shows that these books still ask twice as many quantitative questions as qualitative questions, and that this difference is larger for the introductory college textbooks than the secondary textbooks.

Because the secondary and introductory college chemistry textbooks serve different student populations, it isn't unreasonable that they should have a different distribution of qualitative and quantitative questions. The fact that the introductory college textbooks have fewer qualitative questions than the secondary textbooks might imply that the textbook authors (or college instructors) have assumed that college students have developed this conceptual understanding in their secondary chemistry courses and that not as many of these questions are needed in the introductory college textbooks. Similarly, having more quantitative questions in the introductory college textbooks might suggest that the textbook authors (college instructors) believe that college students need more practice in solving mathematical problems or that these types of questions are more valued/valuable and are more important in helping introductory college students be successful in college chemistry courses or their future career goals.

Macroscopic, particulate, and symbolic representations

The Representation variable was only used in the loglinear analysis for the qualitative questions. However, this analysis showed no significant associations involving the Representation variable, indicating that the qualitative questions involving the macroscopic, particulate, or symbolic representation were homogeneously distributed with respect to the other variables. Of the 740 qualitative questions, 273 (37%) used the macroscopic representation, 248 (33%) used the particulate representation, and 219 (30%) used the symbolic representation. The gas law chapters were chosen for this study because these chapters discuss important concepts at all three representational levels. Gkitzia *et al.* (2011) analysed the images appearing in a 10th-grade book used in Greece and found that these images used the macroscopic representation 35% of the time, the particulate representation 28% of the time, and the symbolic representation 37% of the time; their study differed from ours in that it allowed questions to be categorised using a combination of two or more representations. Although this study and ours analysed different topics (questions in the gas law chapters *versus* visuals throughout the entire book) and categorised the questions differently (one representation *versus* multiple representations when appropriate), they both found that all three representations were being used in these textbooks between 28-37% of the time (close to the value of 33% expected for three representations).

Conclusions

This study compared the distribution of questions appearing in the gas law chapters of four secondary and four introductory college chemistry textbooks from the United States based on six variables—Book Type (secondary or introductory college), Cognitive Skill (lower-order or higher-order), Question Format (calculation, multiple-choice, or short-answer), Question Placement (in-chapter, end-of-chapter, or test bank), Question Type (qualitative or quantitative), and Representation (macroscopic, particulate, or symbolic). The statistical analysis indicated that the lower- and higher-order questions were homogenously distributed throughout these chapters, and that the macroscopic, particulate, and symbolic representation used for the qualitative (non-mathematical) questions were also homogeneously distributed throughout these chapters. This study also showed that the distribution of questions in the textbooks were different for the secondary *versus* the introductory college textbooks, for the calculation *versus* the multiple-choice *versus* the short-answer question formats, for the in-chapter *versus* the end-of-chapter *versus* the test bank questions, and for the qualitative *versus* the quantitative questions. Although the statistical method used in this analysis assumes that the questions in these textbooks should be homogeneously distributed for every variable, that may not be optimal or even preferred from an instructional perspective, especially if the different categories within the variable have different goals. For example, secondary and introductory college textbooks are aimed at different populations, so the distribution and types of question used in these textbooks may not need to be the same. The in-chapter, end-of-chapter, and test-bank questions also have different purposes and goals, so it is reasonable to think that the distribution and types of questions in each part of the textbook should not be the same either.

The significant associations identified in this study are unique to the gas law chapters within the eight textbooks used in this study. If different textbooks were analysed, or even if different chapters within these same eight textbooks were analysed, it is likely that different significant associations would be found. The loglinear analysis method used in this study provides a useful way to analyse the types of questions appearing in chemistry textbooks, and we would encourage similar comparisons using textbooks from other countries, and using chemistry topics other than the properties of gases and gas laws.

Implications for textbook authors and instructors

To our knowledge, this is the first study to identify differences in the distribution of questions within chemistry textbooks based on variables related to the types of questions asked, and we found several significant associations related to question distribution. We do not believe that textbook authors should change their textbooks to eliminate all of these associations. In fact, if the variables serve different student populations (secondary or introductory college) or different instructional goals (in-chapter, end-of-chapter, or test bank; qualitative or quantitative), then there is every reason to expect different distributions of these questions within the textbooks. However, we encourage textbook authors to evaluate these associations and to decide whether there are pedagogical reasons for these differences (associations) to

exist. Should the authors find differences that are unnecessary or even detrimental to the students using these textbooks, we would encourage the authors to eliminate these differences. We would also encourage chemistry instructors using these textbooks to be aware of the differences in the distribution of questions, and to decide whether these differences are detrimental to their students. Should instructors find differences in the types of questions asked by their textbook that they view as detrimental, these instructors should change their instructional lessons to compensate for this detrimental difference.

As an example of a possible detrimental difference, the QF × QP association found in all three loglinear analyses indicated that the questions within the chapters and at the end of the chapters were almost exclusively in the calculation or short-answer format, but the test-banks contained predominantly multiple-choice questions. The lack of multiple-choice questions at the end of the chapter is inconsistent with the end-of-chapter goal of preparing students for course examinations if their instructors give multiple-choice examinations. It may also hinder students' performance on standardized multiple-choice assessments (e.g., No Child Left Behind assessments; American Chemical Society examinations; college/career placement examinations like SAT, ACT, MCAT, or GRE; etc.). In addition, the small number of calculation or short-answer questions in the test banks also limits their usefulness for teachers who do not wish to use multiple-choice questions. Worse yet, the bias toward multiple-choice questions in the test banks could force instructors to write chapter tests or final examinations that are mostly multiple-choice if the instructor wants to use the test bank to write examinations. If the textbook authors decide that this difference is detrimental to students, we would encourage them to include multiple-choice questions at the end of the chapters and to incorporate calculation and short-answer questions in the test banks. If chemistry instructors find this difference to be detrimental to their students, then they would need to supplement their homework assignments to include multiple-choice questions so that students could practice answering multiple-choice questions (if the instructors are planning to give multiple-choice assessments to their students) or they would need to create chapter test and final examination questions that are in the calculation and short-answer format.

Appendix 1. Examples of questions evaluated in this study

For each question shown below, the category for each of the six variables is listed in the order they appear in Table 1.

Book Type and Cognitive Skill: The first question comes from an introductory college textbook and is an example of a lower-order question (Brown *et al.*, 2006, p. 423); the second question comes from a secondary chemistry textbook (Phillips *et al.*, 2002a, p. 398) and is an example of a higher-order question.

Sample Exercise 10.13a: A sample of O_2 gas initially at STP is compressed to a smaller volume at constant temperature. What effect does this change have on (a) the average kinetic energy of the O_2 molecules? [CO, LO, SA, IC, QL, PT]

Section Review 5: **Aerosol Cans** Use the kinetic theory to explain why pressurized cans carry the message "Do not incinerate". [HS, HO, SA, IC, QL, PT]

Question Format and Question Placement: These questions from the same introductory college chemistry textbook (Brown *et al.*, 2006, pp. 403, 432-433) and its test bank (Laurino *et al.*, 2006, p. 357) differ by the QF and the QT variables. The in-chapter question uses the calculation format, the end-of-chapter question uses the short-answer format, and the test bank question uses the multiple-choice format.

Practice Exercise 10.1a: In countries that use the metric system, such as Canada, atmospheric pressure in weather reports is given in units of kPa. Convert a pressure of 745 torr to kPa. [CO, LO, CL, IC, QN, SB]

10.7b: Consider the drawing below [a Boltzmann distribution with a steeper curve at lower molecular speeds labelled 'A' and a flatter curve at higher molecular speeds labelled 'B']. If A and B refer to the same gas at two different temperatures, which represents the higher temperature? [CO, LO, SA, EC, QN, SB]

10. A sample of a gas (5.0 mol) at 1.0 atm is expanded at constant temperature from 10 L to 15 L. The final pressure is ___ atm. A) 1.5, B) 7.5, C) 0.67, D) 3.3, E 15. [CO, LO, MC, TB, QN, SB]

Question Type: These questions from the same introductory college chemistry textbook (Masterton and Hurley, 2000, p. 140) differ by the QT variable. The first question is a quantitative calculation (CL/QN) of a ratio of average translational energies and the second question is a qualitative comparison (SA/QL) of the average translational energies of two gases in a single sample.

57c. Given that 1.00 mol of neon and 1.00 mol of hydrogen chloride gas are in separate containers at the same temperature and pressure, calculate each of the following ratios. (c) average translational energy Ne/average translational energy HCl. [CO, LO, CL, EC, QN, SB]

58a. A mixture of 3.5 mol of Kr and 3.9 mol of He occupies a 10.00 L container at 300 K. What gas has the larger (a) average translational energy? [CO, LO, SA, EC, QL, SB]

Representation: These questions from the same secondary chemistry textbook (Wilbraham *et al.*, 2002a, p. 356) differ only by the RP variable.

45. Heating a contained gas that is held at a constant volume increases its pressure. Why? [HS, HO, SA, EC, QL, PT]

50a. Describe what happens to the volume of a balloon when it is taken outside on a cold winter day. [HS, HO, SA, EC, QL, MS]

69. What can you conclude about the nature of the relationship between two variables with a quotient that is a constant? [HS, HO, SA, EC, QL, SB]

Notes and references

[a] West Memphis High School, West Memphis, AR 72301 USA; E-mail: ggillette@wmsd.net
[b] Department of Chemistry, Middle Tennessee State University, P.O. Box 68, Murfreesboro, TN 37132 USA. Fax: (615) 494-7693; Tel: (615) 904-8558; E-mail: michael.sanger@mtsu.edu

Bergqvist A., Drechsler M., De Jong O. and Rundgren S. -N. C., (2013), Representations of chemical bonding models in school textbooks—help or hindrance for understanding?, *Chem. Educ. Res. Pract.*, **14**, 589-606.

Bloom B. S., (1956), *Taxonomy of educational objectives: The classification of educational goals* (pp. 201-207). New York: Longmans Green.

Boud D., (1995), Assessment and learning: Contradictory or complementary? In Knight P. (ed.), *Assessment for learning in higher education*, London: Kogan Page Ltd.

Britton B. K., Woodward A. and Binkley M. (ed.), (1993), *Learning from textbooks theory and practice*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown T. L., LeMay H. E., and Bursten. B. E., (2006), *Chemistry: The central science* (10th ed.), Upper Saddle River, NJ: Pearson Prentice Hall.

Chiang-Soong B. and Yager R. E., (1993), The inclusion of STS material in the most frequently used secondary science textbooks in the U.S., *J. Res. Sci. Teach.*, **30**, 339-349.

Colantonio J. N., (2005), Assessment for a learning society, *Principal Leadership*, **6(2)**, 22-26.

Danili E. and Reid N., (2005), Assessment formats: Do they make a difference?, *Chem. Educ. Res. Pract.*, **6**, 204-212.

Dávila K. and Talanquer V., (2010), Classifying end-of-chapter questions and problems for selected general chemistry textbooks used in the United States, *J. Chem. Educ.*, **87**, 97-101.

Davis O. L., Jr. and Hunkins F. P., (1996), Textbook questions: What thinking processes do they foster?, *Peabody J. Educ.*, **43**, 285-292.

Davis R. E., Metcalfe H. C., Williams J. E. and Castaka J. E., (2002a), *Modern chemistry*. Austin, TX: Holt, Rinehart, and Winston.

Davis R. E., Metcalfe H. C., Williams J. E. and Castaka J. E., (2002b), *Modern chemistry*: *Assessment item listing*. Austin, TX: Holt, Rinehart, and Winston.

Dingrando L., Gregg K., Hainen, N. and Winstrom C., (2002a), *Chemistry: Matter and change*. Blacklick, OH: Glencoe McGraw-Hill.

Dingrando L., Gregg K., Hainen, N. and Winstrom C., (2002b), *Chemistry: Matter and change: Chapter assessment*. Blacklist, OH: Glencoe McGraw-Hill.

Drechsler M. and Schmidt H. -J., (2005), Textbooks' and teachers' understanding of acid-base models used in chemistry teaching, *Chem. Educ. Res. Pract.*, **6**, 19-35.

Engleman L., (2002), Loglinear models. In *STAT® 10.2 Statistics I* (pp. 618-647), Richmond, CA: SYSTAT® Software, Inc.

Gannaway S. P. and Stucke A., (1996), New literature suggests that we don't have to teach everything in the textbook, *J. Chem. Educ.*, **73**, 773-775.

Gilbert J. K. and Treagust D. (ed.), (2009), *Multiple representations in chemical education*, Dordrecht: Springer-Verlag.

Gkitzia V., Salta K. and Tzougraki C., (2011), Development and application of suitable criteria for the evaluation of chemical representations in school textbooks, *Chem. Educ. Res. Pract.*, **12**, 5-14.

Grunert M. L., Raker J. R., Murphy K. L. and Holme T. A., (2013), Polytomous versus dichotomous scoring on multiple-choice examinations: Development of a rubric for rating partial credit, *J. Chem. Educ.*, **90**, 1310-1315.

Haláková Z. and Proksa M., (2007), Two kinds of conceptual problems in chemistry teaching, *J. Chem. Educ.*, **84**, 172-174.

Holme T. and Murphy K., (2011), Assessing conceptual and algorithmic knowledge in general chemistry with ACS exams, *J. Chem. Educ.*, **88**, 1217-1222.

Hurd P. D., Robinson J. T., McConnell M. C. and Ross N. M., Jr., (1981), *The status of middle school and junior high school science; Technical Report, Vol. 1*, Center for Educational Research and Evaluation, Louisville, CO.

Johnstone A. H., (1993), The development of chemistry teaching, *J. Chem. Educ.*, **70**, 701-704.

Johnstone A. H., (2010), You can't get there from here, *J. Chem. Educ.*, **87**, 22-29.

Johnstone A. H. and Ambusaidi A., (2000), Fixed response: What are we testing?, *Chem. Educ. Res. Pract.*, **1**, 323-328.

Johnstone A. H. and Ambusaidi A., (2001), Fixed response questions with a difference, *Chem. Educ. Res. Pract.*, **2**, 313-327.

Justi R. S. and Gilbert J. K., (2002), Models and modelling in chemical education, in Gilbert J., De Jong O., Justi R., Treagust D. and Van Driel J. (ed.), *Chemical education: Towards research-based practice* (pp. 213-234). Dordrecht: Kluwer.

Kumi B. C., Olimpo J. T., Bartlett, F. and Dixon B. L., (2011), Evaluating the effectiveness of organic chemistry textbooks in promoting representational fluency and understanding of 2D-3D diagrammatic relationships, *Chem. Educ. Res. Pract.*, **14**, 177-187.

Laurino J. P., Cannon D. J., Richter H. and Cooke E., (2006), *Chemistry: The central science: Test item file* (10th ed.), Upper Saddle River, NJ: Pearson Prentice Hall.

Masterton W. L. and Hurley C. N., (2000), *Chemistry: Principles and reactions* (4th ed.), Belmont, CA: Brooks/Cole Thomson Learning.

Nurrenbern S. C. and Pickering M., (1987), Concept learning *versus* problem solving: Is there a difference?, *J. Chem. Educ.*, **64**, 508-510.

Österlund L. -L., Berg A. and Ekborg M., (2010), Redox models in chemistry textbooks for the upper secondary school: Friend or foe? *Chem. Educ. Res. Pract.*, **11**, 182-192.

Papaphotis G. and Tsaparlis G., (2008), Conceptual *versus* algorithmic learning in high school chemistry: The case of basic quantum chemical concepts. Part 2. Students' common errors, misconceptions and difficulties in understanding, *Chem. Educ. Res. Pract.*, **9**, 332-340.

Pappa E. T. and Tsaparlis G., (2011), Evaluation of questions in general chemistry textbooks according to the form of the questions and the Question-Answer Relationship (QAR): the case of intra- and intermolecular chemical bonding, *Chem. Educ. Res. Pract.*, **12**, 262-270.

Pedrosa M. A. and Dias M. H., (2000), Chemistry textbook approaches to chemical equilibrium and student alternative conceptions, *Chem. Educ. Res. Pract.*, **1**, 227-236.

Phillips J. S., Strozak V. S. and Winstrom C., (2002a), *Chemistry, concepts and applications*. Blacklick, OH: Glencoe McGraw-Hill.

Phillips J. S., Strozak V. S. and Winstrom C., (2002a), *Chemistry, concepts and applications: Chapter assessment*. Blacklick, OH: Glencoe McGraw-Hill.

Pickering M., (1990), Further studies on concept learning *versus* problem solving, *J. Chem. Educ.*, **67**, 254-255.

Ramsden P., (2003), *Learning to teach in higher education* (2nd ed.), London: Routledge Falmer.

Resnick L. B., (2007), Science education that makes sense, In *Research points: Essential information for education policy*, vol. 5, issue 1, Washington, DC: American Educational Research Association.

Roth K. J., Druker S. L., Garnier H. E., Lemmens M., Chen C., Kawanaka T., Rasmussen D., Trubacova S., Warvi, D., Okamoto, Y., Stigler, J. and Gallimore, R., (2006), Teaching science in five countries: Results from the TIMSS 1999 video study. Washington, DC: National Centre for Education Statistics.

Sadler P. M. and Tai R. H., (2007), High school chemistry instructional practices and their association with college chemistry grades, *J. Chem. Educ.*, **84**, 1040-1046.

Sanger M. J., (2008), Using inferential statistics to answer quantitative chemical education research questions. In Bunce D. M. and Cole R. S. (ed.), *Nuts and bolts of chemical education research* (pp. 101-133), Washington, DC: American Chemical Society.

Sanger M. J. and Greenbowe T. J., (1999), An analysis of college chemistry textbooks as sources of misconceptions and errors in electrochemistry, *J. Chem. Educ.*, **76**, 853-860.

Sawry B. A., (1990), Concept learning *versus* problem solving: Revisited, *J. Chem. Educ.*, **67**, 253-254.

Schroeder J., Murphy K. L. and Holme T. A., (2012), Investigating factors that influence item performance on ACS exams, *J. Chem. Educ.*, **89**, 346-350.

Shepardson D. P. and Pizzini E. L., (1991), Questioning levels of junior high school science textbooks and their implication for learning textual information, *Sci. Educ.*, **75**, 673-682.

Stamovlasis D., Tsaparlis G., Kamilatos C., Papaoikonomou, D. and Zarotiadou E., (2005), Conceptual understanding *versus* algorithmic problem solving: Further evidence from a national chemistry examination, *Chem. Educ. Res. Pract.*, **6**, 104-118.

Staver J. R. and Lumpe A. T., (1993), A content analysis of the presentation of the mole concept in chemistry textbooks, *J. Res. Sci. Teach.*, **30**, 321-337.

Stern L. and Ahlgren A., (2002), Analysis of students' assessments in middle school curriculum materials: Aiming precisely at benchmarks and standards, *J. Res. Sci. Teach.*, **39**, 889-910.

Stinner A. J., (2001), Linking 'The book of nature' and 'The book of science': Using circular motion as an exemplar beyond the textbook, *Sci. and Educ.*, **10**, 323-344.

Suidan L., Badenhoop J. K., Glendening E. D. and Weinhold F., (1995), Common textbook and teaching misrepresentations of Lewis structures, *J. Chem. Educ.*, **72**, 583-586.

Tabachnick B. G. and Fidell L. S., (1996), *Using multivariate analysis* (3rd ed., pp. 239-319), New York: Harper Collins.

Tai R. H., Ward R. B. and Sadler P. M., (2006), High school chemistry content background of introductory college chemistry students and its association with college chemistry grades, *J. Chem. Educ.*, **83**, 1703-1711.

Talanquer V., (2011), Macro, submicro, and symbolic: The many faces of the chemistry "triplet", *Sci. Educ.*, **33**, 179-195.

Tamir P., (1990), Justifying the selection of answers in multiple choice items, *Int. J. Sci. Educ.*, **12**, 563-573.

Tobin K., Tippins D. J. and Gallard A. J., (1994), Research on instructional strategies for teaching science, In Gabel D. L. (ed.), *Handbook of research on science teaching and learning* (pp. 45-70). New York: Macmillan.

Treichel D., (2000), *Chemistry: Principles and reactions test bank* (4th ed.), Belmont, CA: Harcourt College Publishers.

Tulip D. and Cook A., (1993), Teacher and student usage of science textbooks, *Res. Sci. Educ.*, **23**, 302–307.

Weiss I. R., (1987), *Report of the 1985-1986 national survey of science and mathematics education*, Research Triangle Park, NC: Center for Education Research and Evaluation.

Weiss I. R., (1993), Science teachers rely on the textbook. In Yager R. E. (ed.), *What research says to the science teacher, vol. 7: The science, technology, society movement* (pp. 35-41), Washington, DC: National Science Teachers Association.

Wilbraham A. C., Staley D. D., Matta M. S. and Waterman E. L., (2002a), *Chemistry*. Upper Saddle River, NJ: Addison-Wesley Prentice Hall.

Wilbraham A. C., Staley D. D., Matta M. S. and Waterman E. L., (2002a), *Chemistry: Review Module—Chapters 9-12*. Upper Saddle River, NJ: Addison-Wesley Prentice Hall.

Witzel B. S. and Riccomini P. J., (2007), Optimizing math curriculum to meet the learning needs of students, *Prev. Sch. Failure*, **52**, 13-18.

Wright W. E. and Li X., (2008), High-stakes math tests: How *No Child Left Behind* leaves newcomer English language learner behind, *Lang. Policy*, **7**, 237-266.

Zoller U., Dori Y. J. and Lubezky A., (2002), Algorithmic, LOCS, and HOCS (chemistry) exam questions: Performance and attitudes of college students, *Int. J. Sci. Educ.*, **24**, 185-203.

Zoller U., Lubezky A., Nakhleh M. B., Tessier B. and Dori Y. J., (1995), Success on algorithmic and LOCS *vs.* conceptual chemistry exam questions, *J. Chem. Educ.*, **72**, 987-989.

Zoller U. and Tsaparlis G., (1997), Higher and lower-order cognitive skills: The case for chemistry, *Res. Sci. Educ.*, **27**, 117-130.

Zumdahl S. S., (2004), *Introductory chemistry: A foundation* (5th ed.), Boston, MA: Houghton Mifflin.

Zumdahl S. S. and DeCoste D. J., (2004), *Introductory chemistry: A foundation: Test bank* (5th ed.), Boston, MA: Houghton Mifflin.

Zumdahl S. S. and Zumdahl S. A., (2003), *Chemistry* (6th ed.), Houghton Mifflin: Boston, MA.

Zumdahl S. S., Zumdahl S. A., and DeCoste D. J., (2003), *Chemistry: Test item file* (6th ed.), Houghton Mifflin: Boston, MA.