Analysis of classroom response system questions via four lenses in a General Chemistry course

Aaron D. Bruck and Marcy H. Towns

Received 29th December 2008, Accepted 31st July 2009
DOI: 10.1039/b920834h

General Chemistry lecture questions used in an electronic classroom response system (CRS) were analyzed using three theoretical frameworks and the pedagogical context in which they were presented. The analytical lenses included whether students were allowed to collaborate, Bloom's Taxonomy, a framework developed by Robinson and Nurrenbern, and an expanded framework discussed by Bretz, Smith and Nakhleh. Analysis via these frameworks allowed faculty to reflect upon question types used in the course, and to modify instruction by decreasing the number of lower order cognitive skill questions, and emphasizing higher order cognitive skill questions in subsequent semesters.

Keywords: undergraduate education, classroom response systems, reflective practice, pedagogy, clickers

Introduction

The General Chemistry 115 at Purdue University underwent major revisions in the fall of 2006. This is a large-lecture gateway course for engineering, science, and pharmacy majors with an enrollment of 2200-2400 students. An electronic classroom response system (CRS) was implemented in every section to facilitate interactive learning and formative assessment into the lectures. The goals for their use in Chemistry 115 were to encourage attendence, completion of reading before class, and active engagement in the course content.

The goal of this study was to analyze the CRS questions via frameworks and classification systems found in the literature to drive reflective practice. The results of the analysis were expected to promote reflection upon classroom practices and provide a route to data-driven modifications in subsequent semesters.

Classroom Response Systems in science teaching

One of the main themes in CRS research is the promotion of active learning in the classroom, regardless of class size. Ebert-May *et al.* (1997) found that using a classroom response system to involve students led to greater student confidence in doing science and analyzing data. Freeman *et al.* (2007) noted that the use of lecture systems based on active learning resulted in lower classroom attrition rates. Also, Burnstein and Lederman (2001) observed that the use of a classroom polling system *"greatly increas[ed] the participation of students in the lecture class."* (p. 10).

A CRS can also provide an important service to students through its ability to place students in a situation where they are asked a question, then given almost immediate feedback as to whether or not they are correct. This formative assessment has been cited as being highly valuable for student understanding and transfer, especially when the feedback is nearly immediate, as opposed to days later with a test or a quiz (Kulik and Kulik, 1988). Bangert-Drowns *et al.* (1991) added further support for its use as formative assessment by stating that feedback is more effective for students when it leads them to the correct answer, as opposed to simply telling a student whether they are correct or incorrect. Crossgrove and Curran (2008) added that students in their 'clicker' courses retained and transferred information better than students who were placed in a 'non-clicker' course. While students did not necessarily perform better in the CRS courses on exams during the semester, students in the CRS courses scored significantly higher on a test taken four months after the end of the semester. These findings agree with the previous discussion of active learning, with many authors citing decided advantages for students in CRS courses, despite large lecture classrooms where interaction is difficult to achieve (Ebert-May *et al.*, 1997; Klionsky, 2001; Judson and Sawada, 2002; Caldwell, 2007; Freeman, *et al.*, 2007).

The versatility of CRS has been well documented in a variety of science classrooms across all the major disciplines. Research on the effectiveness of these systems has been carried out in physics (Meltzer and Manivannan, 1996; James, 2006), chemistry (Bunce *et al.*, 2006; King and Joshi, 2007), and biology, where a large amount of research has been conducted and reported (Knight and Wood, 2005; Tanner and Allen, 2005; Crossgrove and Curran, 2008). In fact, a full review of clicker research in the field of chemistry was recently published by MacArthur and Jones (2008). Across this body of research a large number of positive effects have been reported.

The current investigation adds to this body of literature by explicating methods grounded in the literature that can be adapted and used by faculty to analyze CRS questions, thus promoting data-driven reflective practices.

Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, IN, USA. E-mail: mtowns@purdue.edu .

Frameworks

Four frameworks were used for analysis: whether or not students were allowed to collaborate on a question (solo versus buddy), Bloom's taxonomy (Bloom, 1956), Robinson and Nurrenbern's framework (Robinson and Nurrenbern, 2006), and Bretz, Smith, and Nakhleh's (Bretz *et al.*, 2004) framework, which expanded on Robinson and Nurrenbern prior work.

Our decision to investigate 'solo' questions (students were directed to work alone) versus 'buddy' questions (students were directed to collaborate) evolved for two reasons. First, the professors integrated individual 'solo' efforts and collaborative 'buddy' efforts into the lectures in each section of the course. Second, the data set also contained several pairs of questions in which students answered a question individually, then immediately answered the same question after discussing the question with a partner. We were interested in how that pedagogical structure might affect the student responses. Crouch and Mazur (2001) also used this pedagogical approach in their investigations on peer instruction. In their work, they used a similar system of having students work together to come to a consensus on a question, and used a polling system to track their answers. The results of this study were positive in terms of the effect of peer learning on student understanding. Fagen *et al.* (2002) confirmed this outcome, using a larger variety of classroom situations and offering some of their challenges in the analysis. Thus, our analysis is related to previous work in physics education research.

The second framework or lens was derived from Bloom's Taxonomy, which was first proposed by Benjamin Bloom in his work *Taxonomy of educational objectives, handbook I: The cognitive domain* (Bloom, 1956). The taxonomy has been cited frequently owing to its wide applicability to classroom situations in multiple disciplines.

A third analytical lens for question type was based on a document posted in the ChemEd DL's Question Bank. In this document Robinson and Nurrenbern (2006) outlined three different types of questions: 'recall' (which we renamed as *definition* to match the expanded framework used later), 'algorithmic', and 'higher-order'. Recall/definition questions *"ask students to recall facts, equations, or explanations (recall questions section, para. 1)."* Algorithmic questions require students to apply a process, usually a calculation or a familiar formula in order to produce an answer. Higher-order questions require a blend of information as well as a transfer of previously-learned information to new situations.

A fourth analytical lens was an extension of Robinson and Nurrenbern question types developed by Bretz *et al.* (2004). They expanded upon the Robinson and Nurrenbern framework by developing a finer grained description of algorithmic and conceptual questions. This description of question types developed by Bretz, Smith, and Nakhleh is displayed in Table 1.

In the Bretz, Smith, and Nakhleh (2004) framework the description of a definition question is identical to the Robinson and Nurrenbern framework. However, algorithmic questions are categorized by the type of domain conversion, analysis, or problem-solving approach as described in Table 1. Conceptual questions in the Bretz, Smith, and Nakhleh framework are classified by how conceptual knowledge is used, either to explain ideas or data, analyze representations, or predict of outcomes.

Table 1 Bretz, Smith and Nakhleh (2004) framework for determining question types

Question type and code	Description
Definition - D	*Recall/Definition* – recall facts, equations, or explanations
Algorithmic	
A-MaMi	*Algorithmic macroscopic-microscopic conversions* – conversions of macroscopic quantities (volumes or masses) to microscopic quantities.
A-MaD	*Algorithmic macroscopic-dimensional analysis* – unit conversions of macroscopic quantities
A-MiS	*Algorithmic microscopic-symbolic conversions* – application of stoichiometric relationships or other mathematical relationships to convert numbers of particles or moles of substances.
A-Mu	*Algorithmic multi-step* – multi-step problems with an application or algebraic manipulation of formulas.
Conceptual	
C-E	*Explanation of underlying ideas* – connecting observed phenomena with an explanation using underlying concepts
C-P	*Analysis of pictorial representations* – analyze representations of molecules or atoms to answer a question.
C-I	*Analysis or interpretation of data* – data given as a graph or table that requires analysis.
C-O	*Prediction of outcomes* – given a scenario that involves chemical and/or physical changes, predict the outcome.

Methods

In the fall of 2006, the Chemistry 115 course enrolled ~2400 students in seven sections taught by five professors. A common set of lectures and CRS questions were used in the course, which generated CRS data sets for each professor. These data sets contained a record of each question and each student's performance for the entire semester. Four data sets spanning three professors and 1100 students were used for analysis (the remaining three data sets were irretrievably corrupted).

Analysis process

The questions were classified as 'solo' or 'buddy' questions based upon how the professors used them (it was noted in the common PowerPoint lecture notes). Next, each question was analyzed using the three theoretical lenses: Bloom's Taxonomy, the Robinson and Nurrenbern framework, and the Bretz/Smith/Nakhleh framework.

Inter-rater reliability

After all of the questions had been coded in the four frameworks by the first author, a subset of thirty (22%) randomly chosen questions were coded by the second author.

Table 2 Number and percentage of questions classified using Bloom's Taxonomy and solo/buddy classification

Cognitive domain	Bloom's Taxonomy Number of questions	Bloom's Taxonomy Percentage of questions	Solo Number of questions	Solo Percentage of questions	Buddy Number of questions	Buddy Percentage of questions
Knowledge	49	35%	45	32%	4	3%
Comprehension	56	40%	44	32%	12	9%
Application	31	22%	23	17%	8	6%
Analysis	3	2%	1	1%	2	1%
Synthesis	0	0%	0	0%	0	0%
Evaluation	0	0%	0	0%	0	0%
Total	139	99%*	113	82%*	26	19%*

*Note: rounding produced totals different from 100%.

Table 3: Comparison of student performance on solo versus buddy questions by professor (α=0.05, significant difference if $p < 0.05$)

Instructor	Solo versus Buddy t	 *p*
Professor A	-3.319	0.002*
Professor B	0.049	0.961
Professor C	-0.579	0.564

The two raters compared their results and agreed on twenty-six of the the thirty codes for Bloom's taxonomy (87%), twenty-five of the thirty codes for the Robinson/Nurrenbern framework (83%), and twenty-three of the thirty codes for the Bretz/Smith/Nakhleh framework (77%). Overall, agreement was reached on seventy-four of the ninety total codes (82%).

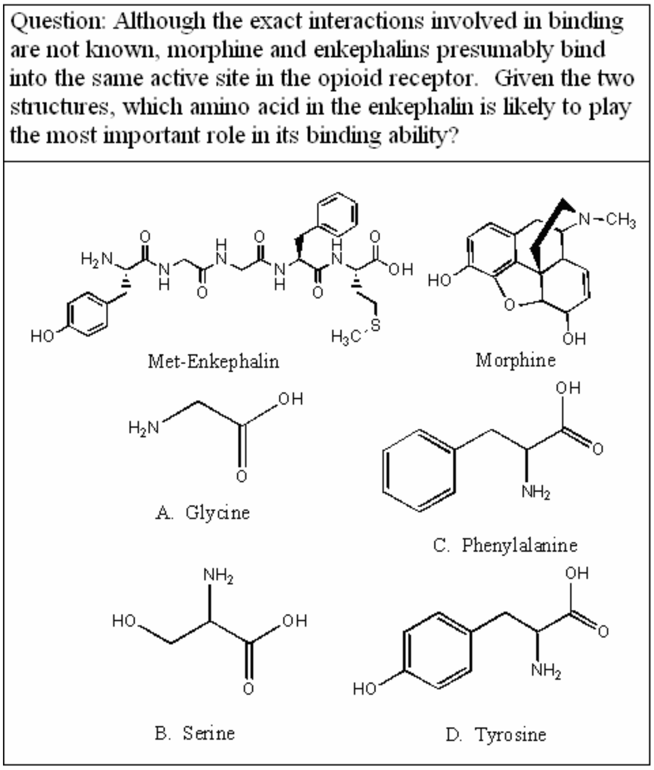
Results

Overall, a total of 153 questions were posed to students across the three lecture sections. However, due to retrieval errors in the CPS system, data were obtained for only 139 of the questions. Of these 139 questions, 68 questions (48.9%) were asked in the lectures of all three professors. Forty-one questions (29.5%) were asked in two of the three professors' lectures, and thirty questions (21.6%) were asked by only one of the professors. These 139 questions served as the basis of the analysis, both in the classification of questions and in the statistical comparisons of the classified groups.

As shown by Table 2, 105 of the 139 questions (75.5%) fell into the knowledge or comprehension category, indicating that the majority of questions were focused on lower order cognitive skills. Among higher order cognitive skills, only three analysis questions were posed to the students.

Over four times as many 'solo' questions as 'buddy' questions were asked across the semester. The comparison of student performance on 'solo' versus 'buddy' questions is given in Table 3. As demonstrated from the data in this table, there was no significant difference in performance on 'solo' compared to 'buddy' questions for two of the professors. However, students in Professor A's lecture performed significantly better ($p < 0.05$) on 'buddy' questions, where students were allowed to collaborate.

Paired solo-buddy questions such as the example given in Fig. 1 were also analyzed. In every case but one, students performed significantly better on questions where they were allowed to collaborate.



Question: Although the exact interactions involved in binding are not known, morphine and enkephalins presumably bind into the same active site in the opioid receptor. Given the two structures, which amino acid in the enkephalin is likely to play the most important role in its binding ability?

Fig. 1: An example of a paired solo-buddy question where students first responded to the question working alone, then responded again after discussion with a peer.

Analysis across question type from Bloom's Taxonomy revealed no significant differences for each professor, as shown in Table 4. Thus, student performance on knowledge, comprehension, application, and analysis questions was not shown to be significantly different, even though there was a preponderance of lower level questions.

The questions were analyzed using Robinson/Nurrenbern framework, as shown in Table 5. Slightly under half of the questions asked, 63 out of 139, were definition questions, which required students to recall *"facts, equations, or explanations"* (Robinson and Nurrenbern, 2006).

No statistically significant differences were found amongst student performance on definition, algorithmic, or conceptual questions, as shown in Table 6.

The questions were also analyzed using the Bretz/Smith/Nakhleh framework and the results are shown in Table 7 alongside the analysis based on Bloom's Taxonomy.

Table 4 Comparison of student performance using Bloom's Taxonomy for each professor (α=0.05, significant difference if $p < 0.05$)

Instructor	Bloom's Taxonomy	
	F	*p*
Professor A	0.442	0.723
Professor B	1.315	0.274
Professor C	2.227	0.089

Table 5 Side-by-side comparison of the questions analysed according to Bloom's Taxonomy and the Robinson/Nurrenbern framework.

Bloom's Taxonomy		Robinson/Nurrenbern Framework Number and percentage of each question					
Cognitive Domain	Number	Definition		Algorithmic		Conceptual (higher order)	
Knowledge	49	47	34%	2	1%	0	0%
Comprehension	56	15	11%	13	9%	28	20%
Application	31	1	1%	28	20%	2	1%
Analysis	3	0	0%	1	1%	2	1%
Synthesis	0	0	0%	0	0%	0	0%
Evaluation	0	0	0%	0	0%	0	0%
Total*	139	63	46%	44	31%	32	23%

*Note: The rounding of individual figures created a total percentage greater than 100%

Table 6 Comparison of student performance by percentage correct on definition, algorithmic, and conceptual questions, as categorized using the Robinson/Nurrenbern Framework
(α=0.05, significant difference if $p < 0.05$)

Professor	Definition	Algorithmic	Conceptual	F	*p*
Professor A	80.0%	84.6%	86.7%	1.127	0.330
Professor B	67.3%	66.3%	72.4%	0.583	0.560
Professor C	68.0%	64.6%	68.4%	0.343	0.710

Eight of the questions that were originally categorized as algorithmic in the initial analysis were reassigned in the expanded framework because of their inability to fit into the algorithmic subcategories. Two were reassigned as definition questions; the other six were placed in conceptual subcategories, with three being assigned as analysis of pictoral representations (C-P) and three others assigned as analysis or interpretation of data (C-I).

The largest category of the algorithmic questions, 16 out of 36, required students to convert a microscopic into a symbolic representation of their knowledge. In the conceptual category, 22 out of 38 questions required students to analyze pictorial representations.

Student performance on each category of question for the Bretz/Smith/Nakhleh framework was analyzed for each professor in the course. As shown in Table 8, across all professors there were no statistically significant differences in how students performed on questions classified using the Bretz/Smith/Nakhleh framework.

Discussion

Three frameworks for analyzing questions—Bloom's Taxonomy, Robinson/Nurrenbern and Bretz/Smith/Nakhleh—demonstrated a preponderance of lower order, definition and/or recall-based questions. Part of what drove the question writing was a desire to have the students read the material prior to class. Thus, many of the questions associated with the readings were knowledge and comprehension questions. In subsequent semesters an effort was made to decrease the emphasis on lower order cognitive skills, and to include more algorithmic and conceptual questions that required higher order cognitive skills.

The Robinson/Nurrenbern and Bretz/Smith/Nakhleh frameworks aided deeper reflection on classroom practices, because they revealed what *types* of algorithmic and conceptual questions were asked. Among the algorithmic questions the most frequently asked were 'microscopic to symbolic' problem tasks. An example of this type of question beyond stoichiometric conversions would be finding the radius of a barium atom from unit cell dimensions. The most frequently asked conceptual questions focused on having students apply concepts to pictorial representations of molecules, such as the question shown in Fig. 1, to assess their understanding. Students answered definition, algorithmic, and higher-order conceptual questions equally well, as demonstrated by the data in Tables 4, 5 and 8.

Paired 'solo-buddy' questions were used in the course to encourage interactivity. Analysis of the paired 'solo' and 'buddy' questions revealed that the percentage of students able to choose the correct response significantly increased in all but one case when they were encouraged to discuss their response with a peer. This movement of students from an incorrect to a correct reponse indicates that students are able to evaluate their conceptual understanding and problem solving approaches, and use the discussion to move in the direction of a correct response. This result is congruent with the findings of Crouch and Mazur (2001) in their study of peer learning outcomes.

Conclusions

Theoretical frameworks such as Bloom's taxonomy, Robinson/Nurrenbern, or Bretz/Smith/Nakhleh can be used to facilitate reflection upon classroom practices. Based upon the results of this study, some of the professors changed classroom practices in subsequent semesters, de-emphasizing definition oriented questions. Although this may have removed the encouragement to read the text in advance, it allowed professors to focus more attention on algorithmic and conceptual questions. This action provided better alignment between the coverage of conceptual questions during lectures and in exams.

Table 7 Bloom's Taxonomy and Bretz/Smith/Nakhleh Framework side-by-side comparison of questions

Bloom's Taxonomy		Bretz/Smith/Nakhleh Framework Number and percentage of each question																	
Cognitive domain	Number	Definition		A-MaMi		A-MaD		A-MiS		A-Mu		C-E		C-P		C-I		C-O	
Knowledge	49	49	35%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
Comprehension	56	15	11%	1	1%	0	0%	10	7%	0	0%	1	1%	19	14%	7	5%	3	2%
Application	31	1	1%	5	4%	7	5%	6	4%	6	4%	0	0%	1	1%	4	3%	1	1%
Analysis	3	0	0%	0	0%	0	0%	0	0%	1	1%	0	0%	2	1%	0	0%	0	0%
Synthesis	0	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
Evaluation	0	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
Total*	139	65	47%	6	5%	7	5%	16	11%	7	5%	1	1%	22	16%	11	8%	4	3%

*Note: The rounding of individual figures created a total percentage greater than 100%.

Table 8: Summary of statistical tests of student performance by question type (α=0.05, significant difference if $p < 0.05$).

	Bretz/Smith/Nakhleh	
Instructor	F	p
Professor A	0.654	0.687
Professor B	1.400	0.216
Professor C	1.497	0.167

In subsequent semesters, some of the professors used more 'buddy' questions during lectures. The data from the fall 2006 study demonstrated that the students benefitted from discussion of their response with a peer, thus it seemed sensible to modify classroom practices in subsequent semesters to encourage more interaction between students.

The type of analysis used in this study encourages faculty to reflect and engage in data-driven practices, and to make changes that move classroom practices in a desired direction. Fellow faculty may also wish to make use of such methodologies to analyze and subsequently modify their own classroom practices.

Acknowledgements

We would like to thank Professors Cisneros, Robinson, Weaver and Wenthold for giving us access to their CPS (classroom performance system) database of questions and responses.

References

Bangert-Drowns R. L., Kulik C. C., Kulik J. A. and Morgan M., (1991), The instructional effect of feedback in test-like events, *Rev. Educ. Res.*, **61**, 213-238.

Bloom B. S., (1956), *Taxonomy of educational objectives, handbook I: the cognitive domain*, New York: David McKay Co. Inc.

Bretz S. L., Smith K. C. and Nakhleh M., (2004), *Analysis of the ACS blended general chemistry exams using a new coding framework*, 227th American Chemical Society National Meeting, Anaheim, CA, March 28, 2004.

Bunce D. M., VandenPlas J. R. and Havanki K. L., (2006), Comparing the effectiveness on student achievement of a student response system versus online WebCT quizzes, *J. Chem. Educ.*, **83**, 488-493.

Burnstein R. A. and Lederman L. M., (2001), Using wireless keypads in lecture classes, *Phys. Teach.*, **39**, 8–11.

Caldwell J. E., (2007), Clickers in the large classroom: current research and best-practice tips, *CBE Life Sci. Educ.*, **6**, 9-20.

Crossgrove K. and Curran K. L., (2008), Using clickers in nonmajors- and majors-level biology courses: student opinion, learning, and long-term retention of course material, *CBE Life Sci. Educ.*, **7**, 146-154.

Crouch C. H. and Mazur E., (2001), Peer Instruction: ten years of experience and results, *Am. J. Phys.*, **69**, 970-977.

Ebert-May D., Brewer C. and Allred S., (1997), Innovation in large lectures – teaching for active learning, *Bioscience*, **47**, 601-607.

Fagen A. P., Crouch C. H. and Mazur E., (2002), Peer Instruction: results from a range of classrooms, *Phys. Teach.*, **40**, 206-209.

Freeman S., O'Connor E., Parks J. W., Cunningham M., Hurley D., Haak D., Dirks C. and Wenderoth M. P., (2007), Prescribed active learning increases performance in introductory biology, *CBE Life Sci. Educ.*, **6**, 132-139.

James M. C., (2006), The effect of grading incentive on student discourse in Peer Instruction, *Am. J. Phys.*, **74**, 689-691.

Judson E. and Sawada D., (2002), Learning from past and present: electronic response systems in college lecture halls, *J. Comput. Math. Sci. Teach.*, **21**, 167–181.

King D. B. and Joshi S., (2007), *Quantitative measures of personal response device effectiveness*, Presented at the 232nd National Meeting of the American Chemical Society. San Francisco, CA. accessed online at http://hdl.handle.net/1860/1269 on June 22, 2008.

Klionsky D. J., (2001), Knowledge in the lecture hall: a quiz-based, group learning approach to introductory biology, *J. Coll. Sci. Teach.*, **31**, 246-251.

Knight J. K. and Wood W. B., (2005), Teaching more by lecturing less, *Cell Biol. Educ.*, **4**, 298-310.

Kulik J. A. and Kulik C. C., (1988), Timing of feedback and verbal learning, *Rev. Educ. Res.*, **58**, 79-97.

MacArthur J. R. and Jones L. L., (2008), A review of literature reports of clickers applicable to college chemistry classrooms, *Chem. Educ. Res. Pract.*, **9**, 187-195.

Meltzer D. E. and Manivannan K., (1996), Promoting interactivity in physics lecture classes, *Phys. Teach.*, **34**, 72-76.

Nurrenbern S. C. and Robinson W. R., (1998), Conceptual questions and challenge problems, *J Chem. Educ.*, **75**, 1502-1503.

Robinson W. R. and Nurrenbern S. C., (2006), Conceptual problems and challenge problems, *J. Chem. Educ. Online*, Retrieved from http://jchemed.chem.wisc.edu:8000/JCEDLib/QBank/collection/CQandChP/index.html on May 5, 2007.

Tanner K. and Allen D., (2005), Approaches to biology teaching and learning: understanding the wrong answers – teaching toward conceptual change, *Cell Biol. Educ.*, **4**, 112-117.