Evaluating the Efficacy of Questioning Strategies in Lecture-Based Classroom Environments: Are We Asking the Right Questions?

Lincoln R. Larson
Matthew D. Lovelace
University of Georgia

Classroom questions represent a potentially powerful tool for interacting with students and stimulating critical thinking. However, the efficacy of the question-and-answer approach to teaching is influenced by many factors, such as the type and cognitive level of questions asked as well as students’ response time. To examine these factors and identify optimal strategies for asking questions in college classrooms, the authors assessed the use of questions in lecture-style science classes at a large public research university in the Southeastern United States. Results revealed several guidelines that could help instructors refine their questioning strategies to enhance student learning.

Introduction

Teachers on all levels have used questions for thousands of years to stimulate students’ thinking and learning. Socrates, for instance, embraced the value of inquiry in the development of thought, using questioning to teach students to reason from specific cases to general principles (Boghosian, 2006). Although advances in technology and pedagogical methods are changing the ways in which teachers teach, basic question-and-answer approaches continue to be a fundamentally important aspect of instruction in modern college classrooms (Nilson, 2010; Prince & Felder, 2006). Questioning is a commonly recommended method for stimulating students’

Although questioning can be a valuable pedagogical tool across all disciplines, questioning strategies can be especially effective in science courses. Science itself is driven by the formation of logical questions and the subsequent quest for answers. Unfortunately, university science courses often are characterized as content-rich but interaction-poor lectures in cavernous auditoriums. However, an increasing number of college instructors in science, technology, engineering, and mathematics (STEM) fields are incorporating innovative instructional methods and technology into their teaching, including group-based learning strategies and classroom response system technology (for instance, Gauci, Dantas, Williams, & Kemm, 2009; Prezler, 2009). As more and more instructors adapt the traditional lecture-based paradigm to be more interactive, effective questioning practices likely will play an important role in the transition.

Although questioning is commonly endorsed as a method for encouraging students to practice and develop higher-order thinking skills (Duron et al., 2006; Edwards & Bowman, 1996; Solomon, Rosenberg, & Bezdek, 1964), research has shown that college instructors tend to focus on fact-based inquiries and rarely pose questions that engage higher-order processes such as application and interpretation (Barnes, 1983; Wilen, 1991). Because questions requiring higher-order thinking can lead students to higher levels of understanding, instructors should reconsider the types of questions they are asking (Dillon, 1988; Edwards & Bowman, 1996; Solomon et al., 1964; Tsui, 2002). Classification systems that have been used to characterize questions include the Taxonomy of Educational Objectives (Bloom, 1956) and the Operations of Intellect System (Gallagher & Aschner, 1963). Organizational frameworks like these tend to categorize questions based on the characteristics of the cognitive processes involved in responding to the question. For example, lower-order question prompts commonly elicit recitation of facts (for example, “What is the name of this species?” or “What are the steps of the scientific method?”), whereas higher-order question prompts stimulate more advanced thinking (for example, “How does speciation occur?” or “Why is the scientific method a useful problem-solving strategy?”). Organizational frameworks focused on cognitive objectives are, therefore, useful for examining the types of questions teachers use during instruction.

Another critical factor to consider is the optimal amount of time to wait for a response after asking a question. This factor is often referred to as teacher “wait time” (Duell, 1994; Rowe, 1974). Researchers on faculty
development commonly encourage instructors to adopt a wait time of several seconds to allow students to consider the question and formulate a response (Sachdeva, 1996; Sanders, 1990; Svinicki & McKeachie, 2011). For instance, Svinicki and McKeachie’s (2011) Teaching Tips recommends waiting at least five seconds for a reply, and even longer for questions that require higher-order processes. Wait time guidelines like this often are based on a combination of teacher experience and research. Nearly 40 years ago, Rowe (1974) first observed that, after asking a question, high school teachers waited, on average, less than one second for students to initiate a reply. Subsequent research has shown that extending wait time by a few seconds (that is, three seconds or more) can improve the frequency, length, and quality of student responses (Gall, 1984; Rowe, 1986). Although several studies of wait time have been conducted in higher education settings (Duell, 1994; Duell, Lynch, Ellsworth, & Moore, 1992; Schneider, Sherman, Prystowsky, Schindler, & Darosa, 2004), the vast majority of studies have examined teacher-student dynamics in primary- and secondary-level settings. In other words, assumptions based on observations from these studies rarely have been tested in college classrooms. While some research supports the value of extended wait time at the college level (for instance, Ellsworth, Duell, & Velotta, 1991), others have found evidence to the contrary (Duell, 1994; Schneider et al., 2004). For example, Duell (1994) found no evidence that extending wait time from 3-6 seconds enhances the quality of either lower-level or higher-level responses by undergraduate students. Furthermore, investigations of wait time at the college level have typically involved class sizes of 25 to 30 students—rarely accounting for the unique dynamics of large lecture settings.

Although the value of intentional questioning in college classrooms cannot be disputed, it is clear that additional research is needed to synthesize findings, inform pedagogical practices, and systematically target effective questioning strategies for large lecture settings. In this study, we attempted to address this research gap by documenting patterns in and relationships between instructor questions and student responses in lecture-based science courses at a large public research university. The study was designed to accomplish three major research objectives: (1) to evaluate the frequency and type of questions asked by instructors; (2) to examine the relationships between question complexity and frequency and depth of student responses; and (3) to examine instructor wait time following questions and determine the wait time that optimizes the frequency and detail of student responses. We constructed the project around an evolving framework for improving the scholarship of teaching and learning (for instance, Badley, 2003; Glassick, Huber, & Maeroff,
This model includes clear goals, appropriate methods, significant results, effective presentation, and reflective critique. These components and critical reflections are incorporated throughout the paper.

**Methods**

We focused our data collection efforts on two lecture-based science courses (50-90 students in each 50-minute section) taught by four different instructors (one associate professor, three assistant professors) at a large public research university in the Southeastern United States. To protect privacy and to reduce observer bias, each faculty participant was assured during informed consent procedures that his or her name and course would remain anonymous in this study.

**Question Variables**

During the 2009-2010 academic year, we observed 18 classes (9 for each course) and documented the following variables for 276 instructor-generated questions recorded during the lectures: question type, question cognitive level, student response attempt, student response detail, and wait time.

**Question Type**

Question types were grouped into one of two categories. Clarification questions required generic responses such as “yes” or “no,” nods, or raised hands, and often referred to general comprehension or course logistics (for example, “Does everyone have access to online course materials?” “Do you have any questions?”). Content-specific questions required specific responses based on course material (for example, “What is the major investment material in forestry?” “How do protists differ from monera?”). The classroom observers were not able to determine if or when instructors’ questions were designed to be rhetorical in nature, so all questions (including those with rhetorical intent) were included in the analysis.

**Question Cognitive Level**

Questions were also grouped into six cognitive categories based on Anderson and Krathwohl’s revised taxonomy of education objectives (2001). These categories included remember, understand, apply, analyze, evaluate, and create. To simplify certain analyses, questions were
regrouped into two major categories based on their cognitive level (lower-order thinking = remember, understand; higher-order thinking = apply, analyze, evaluate, create).

**Student Response Attempt**

A dichotomous variable indicated if students initiated any response (either verbal or nonverbal) to the instructor’s questions. Any attempt at a response (including a simultaneous response from multiple students) was coded as “yes.” The absence of a response (generally silence) was coded as “no.”

**Student Response Detail**

This dichotomous variable characterized the nature of the students’ response as either simple (for example, “yes” or “no,” murmurs, nodding) or detailed (complex, articulated answers). Any response that involved words beyond “yes” or “no” was coded as a detailed response. Because observers were not completely familiar with course content and question topics, the accuracy of student responses was not considered in this coding process. However, “response detail” is a viable metric for assessing response quality, especially if one of the goals of questioning is to simply engage students’ cognitive processes.

**Wait Time**

Wait time was defined as the time (in seconds) between the end of an instructor’s question and either a student response (wait time with response) or a decision by the instructor to move on to another topic (wait time without response). If an instructor interrupted the silence to repeat the same question, wait time did not reset. If an instructor interjected to rephrase the question, wait time did reset, and the new phrase was considered a separate question.

In addition to these five variables, researchers also recorded the number of students present each day in class. Observers took great care to ensure data collection did not interfere with normal classroom dynamics. Although instructors realized that outside observers were present in the audience, they were not aware of the focus of the study. Instructors were told that the guest students (the observers) were graduate students interested in learning about new scientific concepts. The data should, therefore, reflect authentic classroom proceedings.
Data Analysis

We used analysis of variance (ANOVA) to compare the means of different instructors on the quantitative (discrete and continuous) variables of interest (for example, number of questions asked, wait time). We used Pearson correlation coefficients (r) to assess the strength of the relationship between these quantitative variables (for example, number of questions asked, number of students, student response rate). Pearson’s chi-square tests were used to compare observed and expected frequencies and examine relationships among descriptive categorical variables (for example, types of questions, cognitive level of questions, student response attempts, student response detail). When expected cell counts were small (that is, less than five), we calculated a Fisher’s exact p-value. Because substantial differences between results of chi-square and Fisher’s tests of independence were not observed, only chi-square results are presented in this paper. Spearman’s rank correlation coefficient (rho), a non-parametric measure of statistical dependence that compares the ranked values of two variables, was used to examine the strength of association between certain quantitative and qualitative metrics.

Limitations

Upon reflection, we acknowledge that our inferences in this study were limited by several factors. First, the small subset of courses and instructors sampled may not adequately represent the larger population of large lecture science courses. For example, only one associate professor was included in the analysis, and the questioning approach of this single individual may not reflect the teaching style of other senior professors. The use of professor rank (assistant, associate) as a proxy for teaching experience is also limiting, because it does not account for other potentially influential factors such as previous courses taught, involvement in professional development programs, and experience teaching in non-academic environments. Unfortunately, these data were not available for this study.

Although this study focused on several key variables affecting the question-and-answer dynamic, other factors influencing student (for example, age, gender, academic experience, mean GPA) and instructor (for example, confidence with subject matter, number of times teaching course) behavior were not considered and could affect these dynamic classroom processes. Furthermore, the researchers’ coarse categorization of responses did not allow for fine scale assessments of student answers. For example, although the “response detail” metric captured levels of student engagement and critical thinking, it did not convey the accuracy
of student responses. Because observers were not completely familiar with course content, they were not able to assess and record these data. Future studies could endeavor to capture this element of detail, which represents an important component of formative assessment.

Finally, the concept of wait time is subject to multiple interpretations. In this study, researchers considered a repeated question to be an identical question and a rephrased question to be a new question. Other interpretations could yield slightly different results. Despite these limitations, results of this exploratory study provide insights that could identify problematic patterns and help instructors to refine their questioning techniques in large lecture courses.

**Results**

**Instructor Questions**

Collectively, the four instructors asked an average of 15.3 (SD = 11.3) questions per 50-minute class period, or approximately one question every three minutes. The number of questions asked per class ranged from 5 to 37 and, as expected, varied significantly among the instructors ($F_{3, 14} = 13.1; p < 0.001$). The three assistant professors asked significantly more questions per class period than the associate professor ($F_{1, 16} = 8.4; p = 0.011$). After controlling for instructor differences, the relationship between class size and the number of questions asked by instructors was not statistically significant ($r = 0.269; p = 0.296$).

Most questions asked by instructors (89.9%) were based on course content. Basic clarification questions that did not target a particular subject were asked much more infrequently than content-related questions (8.3% of total), and questions completely unrelated to course content (1.8%) were the least common. About one fourth (23.2%) of the content-related questions required a simple response such as “yes” or “no,” and 66.7% required a more complex and detailed answer. Significant differences in the types of question asked were observed among the four instructors, but these differences were not related to professor rank ($\chi^2_{3, 276} = 37.4; p < 0.001$; range = 0.8% to 43.8% clarification questions). Cognitive level of questions trended toward the lower end of Bloom’s Revised Taxonomy (Anderson & Krathwohl, 2001), with 78.2% of the questions related to remembering or understanding (see Figure 1). Inquiries targeting Bloom’s higher levels of evaluation and creation were very rare (4.7% of total). When questions were regrouped into two major categories based on their cognitive level (targeting lower-order and higher-order thinking skills), differences between
instructors were evident ($\chi^2_{3, 276} = 8.2; p = 0.042$). The associate professor asked lower-order questions at a higher rate than the assistant professors.

**Student Responses**

Students attempted to answer the instructors’ questions 51.1% of the time overall. Many of these responses (21.6%) were simple “yes” or “no” answers. Statistically significant differences in response rates in classes taught by the different instructors were not observed ($\chi^2_{3, 276} = 6.7; p = 0.081$). However, when the percent of total questions per class eliciting a response was compared between the different professor ranks, data showed that students generally responded to the associate professor’s questions more often than to the assistant professors’ questions ($F_{1, 16} = 7.3; p = 0.016$).

Class size was positively associated with the likelihood of a student response to instructor questions ($r = 0.604; p = 0.008$). The number of questions asked per minute of instruction was not related to the frequency of student responses ($r = -0.078; p = 0.759$), but the detail of student responses
seemed to increase as the number of questions asked per minute decreased (Spearman’s rho = -0.473; p = 0.011).

The frequency of student response attempts did not depend on the cognitive level of the question that was asked ($\chi^2_{5, 276} = 7.6; p = 0.179$; see Table 1). This pattern held even when questions were regrouped into two major categories based on their cognitive level: lower-order thinking = remember, understand; higher-order thinking = apply, analyze, evaluate, create ($\chi^2_{1, 276} = 0.1; p = 0.884$). However, the frequency of student response attempts did depend on the type of question that was asked ($\chi^2_{1, 276} = 6.5, p = 0.011$). In general, students responded to content-specific questions at a much higher rate (53.8%) than basic clarification or logistical questions (26.1%).

Detail of student responses was not significantly related to the cognitive level of the question asked ($\chi^2_{4, 142} = 8.5; p = 0.073$; see Table 1). When questions were regrouped into two major cognitive categories (lower-order and higher-order thinking), a similar result was obtained ($\chi^2_{1, 142} = 3.5; p = 0.068$). Although the relationship between cognitive level of question and student response detail was not statistically significant, higher-level questions—particularly those in the apply and analyze categories—were generally more likely than lower-level questions to elicit a detailed or complex student response.

**Wait Time**

The mean wait time before a student response varied significantly across instructors ($F_{3, 138} = 7.6; p < 0.001$), and ranged from 1.8 to 5.0 seconds per instructor (overall mean = 3.75 seconds; overall median = 3.0 seconds). On average, students responded to the assistant professor’s questions faster than they responded to the associate professors’ questions (2.3 vs. 4.5 seconds; $F_{1, 140} = 14.0; p < 0.001$). The mean wait time before instructors moved on to another topic without a response also varied across instructors ($F_{3, 130} = 7.1; p < 0.001$), ranging from 2.1 to 6.2 seconds per instructor (overall mean = 4.56 seconds; overall median = 3.0 seconds). Assistant professors waited significantly longer (5.3 vs. 2.1 seconds; $F_{1, 132} = 13.3; p < 0.001$) than the associate professor before moving on without a student response. In many cases, all of the instructors waited one second or less (11.9% of all unanswered questions), three seconds or less (50.0%), or five seconds or less (71.6%) before moving on. Wait times exceeding 10 seconds were observed for only 8.2% of questions.

Examination of wait time prior to student response showed that response rates generally increased rapidly for approximately five sec-
onds and then began to wane (see Figure 2). Mean wait time without a response did not vary significantly according to the cognitive level of the questions ($F_{5, 128} = 1.0; p = 0.395$). When questions were regrouped into two major cognitive categories (lower-order and higher-order thinking), a similar result was obtained ($F_{1, 132} = 1.3; p = 0.264$). Mean wait times without responses for questions on most cognitive levels were below the five-second threshold (see Figure 3), with the exception of higher-order questions in the evaluation and creation categories.

Wait time did not significantly influence the detail of student responses ($F_{1, 140} = 0.1; p = 0.764$). Similarly, wait times before student responses to lower-order and higher-order questions were not significantly different ($F_{1, 140} = 0.1; p = 0.740$). Class size was inversely related to mean wait time with ($r = -0.195; p = 0.001$) or without ($r = -0.195; p = 0.024$) a response. This pattern showed that more students in the class often equated to shorter wait times for both the students answering the questions and the professors asking them.

**Discussion**

Results of this study revealed that in the lecture-based science classes
we observed, instructors attempted to engage their students by regularly asking questions throughout the class period. The less-experienced professors (at the assistant level) generally asked more questions, but the questions asked by the more experienced professor (at the associate level) were answered at a higher rate. Theory and research on how professors develop as instructors support this trend, reflecting a movement toward quality over quantity in questioning approaches as teaching experience is gained (Kane, Sandretto, & Heath, 2002; Kugel, 1993).

Most questions asked by instructors did not require higher-order thinking skills to develop a response, and a majority of the questions were rooted in the remembering and understanding levels of Bloom’s Revised Taxonomy (Anderson & Krathwohl, 2001). In fact, the highest-ranking professor in this study (associate) asked the most lower-order questions. Surprisingly, data also showed that the cognitive level of the question did not influence students’ response rate. This finding suggests that

---

**Figure 2**

*Wait Time Until Student Response Following Instructor Questions (n = 142 Questions With Responses)*

*Note.* Optimal wait time of five seconds is marked.
students may be willing and able to respond with higher-order thinking if professors make a concerted effort to ask higher-order questions. Additional research is needed to better understand the complexity and quality of student responses to these higher-order questions. This study also revealed an inverse relationship between the number of questions asked per class period and the detail of student responses. Data showed that the teacher who carefully crafted a few pointed questions at critical junctures throughout the lecture appeared to be more effective than those who bombarded the students with questions to stimulate superficial engagement. Future studies could incorporate a broader spectrum of instructors and experience levels to examine this pattern in more detail.

Although previous research has highlighted the communication challenges associated with teaching in a spacious, impersonal, lecture-based context (Gleason, 1986), associations between class size and student

Figure 3
Mean Wait Time (With 95% CI) Without Student Response Following Instructor Questions at Various Cognitive Levels Based on Anderson & Krathwohl’s (2001) Revised Taxonomy ($n=134$ Questions Without Response)

Note. Optimal wait time of five seconds is marked.
responses in this study showed that questions asked in larger classes were more likely to generate responses and more likely to generate quick responses than questions asked in smaller classes. These patterns might be expected given an increase in the number of potential respondents, and they suggest that questioning strategies may be equally effective in classes of all sizes. In other words, even if the use of questions in cavernous auditoriums may seem daunting from the instructors’ perspective, asking questions in such large lecture settings may, indeed, be a worthwhile endeavor.

Despite this encouraging relationship between student responses to questions and class size, the overall student response rate observed in this study was surprisingly low (approximately one half of all questions went unanswered) and remains a major concern. Previous research has indicated that students frequently are embarrassed or reluctant to admit confusion in front of their peers (Svinicki & McKeachie, 2011). Instructors should, therefore, account for this sensitivity and strive to develop mechanisms for clarifying course content that minimize the risk of public humiliation. For example, instructors could cultivate a welcoming atmosphere for interjections during lecture, encourage small-group consultations before calling on students, ask some questions that do not have wrong answers, or enhance communication outside of class (Svinicki & McKeachie, 2011). To interpret the low response rate in this study, it might also help to consider how instructor questions were phrased. In our study, we observed that student responses to clarification questions were minimal. Instructors who attempted to clarify difficult concepts with quick queries such as “Any questions?” or “Does everyone understand?” generally were greeted with silence. Dillon (1988) recommends using alternative ways of questioning, including rephrasing prompts into imperative forms. “Tell me what questions you have” is an example of an alternate way to prompt for clarification questions from students that would likely be more effective.

Wait time is another important issue instructors should consider when asking questions in large courses. Many teachers, when confronted by awkward silence, are inclined to cut in and move on to another topic. This phenomenon can be exacerbated by the characteristics of larger courses. Increased distances between students and the instructor, theater-style seating, and greater numbers of students—most of whom do not know each other—can lead to an impersonal atmosphere that inhibits communication (Gleason, 1986; Trees & Jackson, 2007). In such an environment, the perceived weight of an intimidating silence may seem even greater. Our study supported these assumptions, revealing that larger class sizes
were correlated with shorter instructor wait times for questions that were not answered. Surprisingly, the highest-ranked instructor in this study demonstrated the shortest average wait time without a response. Overall, instructors often waited three seconds or less for an answer before moving on. Additionally, most instructors did not appear to adjust their wait time to account for question complexity.

Analyses of the answered questions in this study demonstrated rapid increases in student response rates until about five seconds, at which time responses began to taper off. Hence, this study suggested an optimal wait time of 4-6 seconds for generating student responses—a recommendation in line with previous studies (Duell, 1994). Wait times before student responses were not significantly related to the cognitive level of questions asked or the detail of responses; however, these unexpected results should be cautiously interpreted due to the coarse nature of the response quality metric (that is, the dichotomous grouping into simple or detailed responses without consideration of response accuracy). In other words, students answering complex questions quickly (demonstrating a short wait time) may not have provided accurate or appropriate responses. Collectively, results suggest that instructors asking questions should consciously monitor wait times and allow a sufficient period for students to respond, especially if the question targets higher-order thinking skills.

For questions that are particularly complex, simply waiting a bit longer for responses may not be enough. A structured way to give students more time to think and potentially to formulate a higher-quality response is to give them a minute or two to outline their response to a question first using short, low-stakes writing prompts (Angelo & Cross, 1993). Additionally, drawing on strategies from the cooperative learning literature (Johnson, Johnson, & Smith, 2007; Slavin, 2011), instructors could consider incorporating informal, structured, group learning strategies into their instructional repertoire. For instance, in an activity commonly known as Think-Pair-Share (see Cottell, 2010), instructors first give students a moment to reflect on a question, then have each student discuss his or her response with a neighbor for a minute or two, and, finally, invite individuals to share their responses with the entire class. Each of these supplemental strategies could increase the efficacy of classroom questions and stimulate students’ critical-thinking skills.

Conclusions and Recommendations

In summary, this study supports the idea that questions—if executed appropriately—represent a useful tool for engaging students and stimulating
higher-order cognitive processes in lecture-based classroom environments. In response to the results of this study and similar findings from previous research, we offer several suggestions that could help teachers improve the way in which they ask questions in lecture-style college classes:

1. **Generate questions that target higher-order thinking skills.** Unfortunately, most questions asked in large lectures tend to focus on basic knowledge and comprehension. Though these questions are appropriate in many situations, instructors should make a concerted effort to challenge students with more difficult questions. Students will still attempt to answer questions that stimulate critical thinking at higher cognitive levels.

2. **Make questions count.** Questions undoubtedly facilitate teacher-student interaction, but excessive questioning can detract from learning and diminish the quality of student responses. Although students may continue answering questions asked at high frequencies, the detail and quality of student responses to these rapid-fire questions typically declines. Every question should have a purpose. In other words, there is such a thing as a bad question.

3. **Don’t be afraid to wait for an answer.** Instructors are often hesitant to wait through prolonged silence and often provide students with inadequate time to formulate responses. Students need some time to think (preferably 4-6 seconds); however, waiting too long may not appreciably improve the frequency or depth of responses and could waste class time. If the question is very complex, give students a chance to write down an answer before responding. Develop strategies to make the inevitable (and essential) wait time more productive.

Although this study was based on a small sample of four instructors across two courses within a science-oriented curriculum, these recommendations likely are applicable in cross-disciplinary contexts due to the generally homogenous structure of the college lecture environment. As scholars expand the use of empirical approaches to inform question-and-answer processes and practices, future investigations could examine student response quality in more detail and incorporate a larger, more diverse sample of instructors and classes. This and similar studies
should, ultimately, allow researchers to identify and assess optimal strategies for asking questions in a variety of college teaching environments.

References


Rowe, M. B. (1986). Wait time: Slowing down may be a way of speeding up!

Acknowledgments

The authors wish to thank all of the anonymous faculty and students who participated in the study.

Lincoln R. Larson recently received his doctorate from the Warnell School of Forestry & Natural Resources at the University of Georgia, and he currently works as a postdoctoral research associate at Cornell University. He has earned degrees in wildlife biology and natural resources recreation and has worked as an environmental educator in both formal and non-formal settings. His current teaching and research interests focus on social science issues in natural resource management. Matthew D. Lovelace is a doctoral candidate in the department of educational psychology at the University of Georgia and a statistics consultant for a National Science Foundation grant-funded study on best practices in teaching biology at large universities. In addition to scholarship of teaching and learning research, he studies factors that influence—and are influenced by—student engagement, college readiness, and graduation.