Electronic Journal of Science Education Vol. 18, No. 5 (2014)

What kinds of questions do future elementary teachers ask in a university science course? Results from an online question-ranking tool

Matthew A. d'Alessio
California State University, Northridge

Abstract

Asking questions is one of the key external indicators of curiosity and a cornerstone of scientific inquiry. To understand the types of questions students ask in an introductory college science class for preservice elementary teachers, students used an online question-ranking tool (Google Moderator) as part of the daily routine. This study compiled 2,494 student questions from 128 students in four sections of the same course taught by the same instructor. After grouping questions into nine categories based on cognitive level, two independent regimes of questioning emerged – knowledge-rich and exploration-rich. When students had more specific knowledge about a topic, they asked a greater percentage of questions that demonstrate knowledge, fewer testable scientific questions, and fewer questions about course logistics. This knowledge-rich regime is more active in students with high quiz scores. Lower scoring students tended to be more active in an exploration-rich regime. They asked more scientifically testable questions that reflect a desire to explore cause-and-effect relationships. During an online question-ranking process, students decided which questions they would like to discuss in class by voting for their preferred questions. They submitted 13,301 student votes on the questions in this study. Despite differences in question-asking, students at all performance levels voted for similar questions in this voting process. Voting patterns revealed that students slightly preferred higher order questions that demonstrate knowledge. Questions that are both scientifically testable and demonstrate knowledge received, on average, the highest number of votes per question (about 30% higher than logistical questions about course expectations).

Correspondence concerning this article should be addressed to: Matthew d'Alessio, matthew.dalessio@csun.edu, California State University Northridge, 18111 Nordhoff St, Northridge, CA 91330-8266

Keywords: curiosity, questioning, preservice teachers, technology

Introduction

Humans are born curious, which is an inherent part of our ability to adapt and thrive (Gopnik, Meltzoff, & Kuhl, 2000). While the internal psychological construct of curiosity may be difficult to measure, studies that quantify external expressions of curiosity show that it is virtually universal: curiosity is independent of gender (Johnson & Beer, 1992; Smith, 2010), overall academic achievement (Day, 1968; Smith, 2010), and even age (Engel, 2011; Peterson, 1979; Smith, 2010). As we grow, we formalize this curiosity about the natural world into the process of science, finding answers to our questions using methods of scientific inquiry (University of California Museum of Paleontology, 2013). Because curiosity is a driving force behind scientific progress, curiosity might be considered critical to the success of science education (Osborne, Collins, Ratcliffe, Millar, & Duschl, 2003). Despite curiosity's foundational role, it is conspicuously absent from our elementary classrooms (Engel, 2009, 2011; Susskind, 1979). As science educators strive to bring back curiosity, we need to better understand what it is, how to cultivate it, and how to harness it to advance scientific discovery.

Theoretical framework

Because curiosity is innate, cultivating curiosity itself may be less important in science education than helping students express their curiosity as productively as possible. Questions are the vehicle by which curiosity is expressed and pursued. While curiosity may be universal, the ability to ask questions that guide scientific investigations is not. Expert scientists ask more questions that allow them to better plan and guide their investigations than novice science students (Hackling & Garnett, 1992). Fundamental to this study is the assumption that question-asking is an expert skill that can be cultivated to empower novice students to express their curiosity. So how does one develop an expertise in question-asking?

Previous researchers have probed instructional approaches for promoting and improving question-asking at the University level (Donohue-Smith, 2006; Keeling, Polacek, & Ingram, 2009; Kowalski & Kowalski, 2012; Marbach-Ad & Sokolove, 2000) down to the elementary level (Jirout, 2011; King, 1994; Susskind, 1979), where an entire issue was devoted to the topic in the NSTA elementary school journal (Froschauer, 2010). The Next Generation Science Standards identify question-asking as a key skill in the scientific process and define progressive steps for developing that skill in K-12 education (National Research Council, 2012). The standards begin with asking questions based on observations at the early primary level, proceed to being able to identify scientifically testable questions in upper elementary grades, expand to generating questions that can actually be investigated in a classroom during middle school, and culminate in students' ability to frame detailed questions that clarify evidence and its relationship to scientific arguments by the end of high school (NGSS Lead States, 2013). The progression is designed to take young novice questioners and bring them closer to expert scientists.

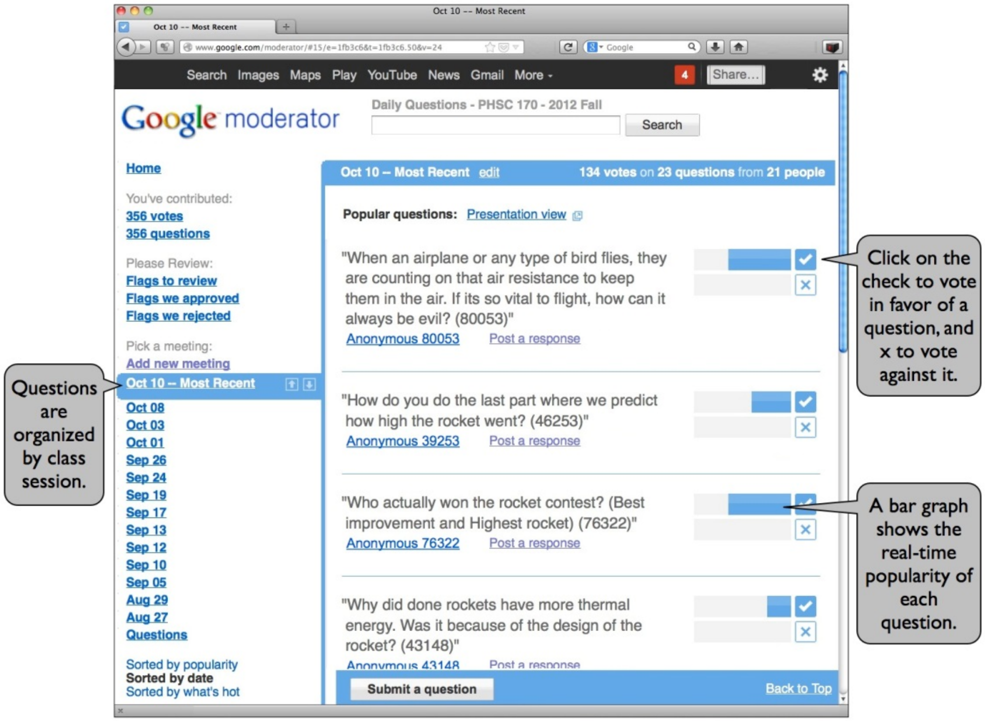
Unanswered questions about question-asking

Despite these attempts to cultivate question-asking, there are substantial gaps in the existing theories about why people ask questions. While experts differ from novices in many ways, the expert's deep library of background knowledge is particularly relevant for understanding their questioning abilities (Bransford, Brown, & Cocking, 2000, Chapter 2). There is competing evidence over the role of pre-existing knowledge in question-asking. One model suggests that people ask questions because they are missing information, an "information gap" (Loewenstein, 1994). Stated another way, questions represent "knowledge goals" (Ram, 1991); when information gaps become sufficiently large, people are motivated to ask questions in order to meet their knowledge goals (Jirout, 2011). Using this paradigm, some have argued that students cannot be expected to formulate meaningful questions without context and ample time for exploration (Arnone, 2003; Froschauer, 2010) – these activities ensure that students have sufficient background knowledge from which to identify gaps. Experts clearly have large bases of knowledge from which they can identify gaps. However, investigations by Scardamalia and

Bereiter (1992) raise interesting questions about how novices use preexisting knowledge during questioning. Their study prompted students to generate as many questions on a topic as they could. One group read a text with background information prior to generating questions, while the other group was simply given a blank piece of paper for writing the questions with no prior introduction to the topic. Reading background information seemed to *inhibit* question-asking; students without the text generated more questions, and those questions had greater educational value. What exactly is the role of preexisting knowledge in question asking?

Moderated Questions

To encourage students to ask questions and provide a forum for them to evaluate questions of their peers, instructors can employ online question-ranking tools like "Google Moderator" (http://google.com/moderator, Figure 1). In a classroom, question ranking tools can allow students to submit questions anonymously and then members of the class can vote for the questions they would like to see addressed. Instructors can devote instructional time to answering the most popular questions. While Google promotes the service in its training for educators (Google, n.d.), there is minimal research on its efficacy in the classroom, nor that of similar online question-ranking tools. The following sections describe the theoretical motivation for using online question-ranking tools in a classroom setting.



Oct 10 -- Most Recent

www.google.com/moderator/#15/e=1fb3c6&t=1fb3c6.50&v=24

Daily Questions - PHSC 170 - 2012 Fall

Oct 10 – Most Recent edit 134 votes on 23 questions from 21 people

You've contributed:
356 votes
356 questions

Popular questions: Presentation view

"When an airplane or any type of bird flies, they are counting on that air resistance to keep them in the air. If its so vital to flight, how can it always be evil? (80053)"
Anonymous 80053 Post a response

"How do you do the last part where we predict how high the rocket went? (46253)"
Anonymous 39253 Post a response

"Who actually won the rocket contest? (Best improvement and Highest rocket) (76322)"
Anonymous 76322 Post a response

"Why did done rockets have more thermal energy. Was it because of the design of the rocket? (43148)"
Anonymous 43148 Post a response

Questions are organized by class session.

Click on the check to vote in favor of a question, and x to vote against it.

A bar graph shows the real-time popularity of each question.

Figure 1. Annotated screenshot of Google Moderator, an online question-ranking tool.

Student ownership of learning experience stimulates curiosity

Educators can activate curiosity more effectively when they give students choices that allow them to shape the experience into something that has greater personal relevance (Arnone & Small, 2011). Student questions regularly relate course material to personal experience and personal interests (e.g., questions about skateboarding during a unit on kinetic energy). Online question tools therefore allow students to direct at least some portion of the instructional time towards their interests. Students gain greater control over the zone of proximal development, which can lead to greater knowledge construction (Scardamalia & Bereiter, 1991).

Curiosity involves a range of feelings including 1) the frustration of not-knowing and 2) the pleasure and satisfaction derived from satisfying curiosity (Loewenstein, 1994). Students may avoid curiosity-arousing situations to avoid the frustration of not-knowing or the demoralization of being wrong. However, if they are confident that their curiosity will be resolved then they are willing to endure the temporary frustration. Increasing the probability that a question interesting to a student will be answered could help activate curiosity. Using online question-ranking tools, students express curiosity through asking a question and again by voting on the questions submitted by peers.

Metacognitive value of voting

Online question ranking has the potential to provide a forum for student self-reflection on course content. As students review and vote on the questions, they have the opportunity to ask themselves, "Do I already know the answer to this question? Do I want to know it?" The online question-ranking tool provides a forum for students to reflect on the boundaries of their knowledge in a quick but structured process.

Low risk questioning environment

Students are more likely to fully embrace the scientific enterprise when the classroom climate is considered "low risk," meaning that everyone is encouraged to participate without feeling judged (Arnone, 2003; Lederman, 1992). Google Moderator includes support for students to respond directly to questions by their classmates, but this feature may actually work against the goal of encouraging questioning. While peer instruction can be valuable for learning content, a primary goal of this activity is to encourage students to express their curiosity and model question-asking. Scardamalia and Bereiter (1991) showed that holding students accountable in some way for seeking answers to questions they ask reduces both the number and complexity of questions they ask. They suggest that students strive to minimize risks of failure or overwork. Preventing student responses keeps the focus of online question tools largely on question-asking.

Method

Research Questions

Ram (1991) argues that a theory of question generation will help us develop more effective theories of instruction. In particular, such an understanding will provide guidance about how to help students progress from novice question-askers to experts at scientific inquiry. To provide observational evidence that may help distinguish between competing theories of question generation, this study investigated the following questions with particular emphasis on the role of prior knowledge in question-asking:

1. What is the relationship between the knowledge students have and the types of questions they ask?
2. To what extent can students recognize and value questions at different cognitive levels (as identified by question-ranking using an online tool)?

Data Collection

Students submitted 2,494 student questions and 13,301 votes using Google Moderator during four semester-long sections taught by the same instructor of a general-education physical science course (a full list of questions and votes is available in a supplementary file available upon request from the author). The course taught science content to preservice elementary teachers at a large, regionally-focused state university. The questions included in this study represent every question submitted by every enrolled student in the course during the study period. The 128 students enrolled ranged from freshman through senior status (*Freshman*=11%, *Sophomore*=31%, *Junior*=37%, *Senior*=21%) and the majority (70%) were officially declared in a teacher preparation B.A. program. This course is required for the teacher preparation program, so it forms a somewhat captive audience of students that may or may not wish to be there. Many of the students had negative experiences with math and science in the past. In surveys given at the beginning of the course, 62% of the students identified their level of comfort with math as "it terrifies me" (17%) or "I am hesitant about doing math" (45%). They offered similar anecdotal comments about science. The teacher preparation program requires students to declare a subject matter concentration; only 9% of the study participants specified science as their concentration and none chose math.

Classroom Procedure

Students asked questions after every class session (29 total during each semester) using the online tool. Students completed an online form called the “Daily Report” within 24 hours of each class session. Among other prompts promoting reflection, it required students to ask one question related to the day’s class session. Students did not have the opportunity to see others' questions before writing their own. After 24 hours, all questions were automatically posted to the online question-ranking site (Google Moderator), with questions shown anonymously. Students had until the beginning of the next class period to vote on which of their classmates’ questions they would like to see addressed during class. Students could vote for up to five questions and the online tool ensures that students can’t vote more than once for the same question. Completion of the daily report counted as 10% of the semester grade. Students also earned a bonus point each time their question was the top vote recipient for the day. Each class period, the instructor devoted 10-15 minutes to answering the top few questions. Appendix A provides further detail for instructors considering implementing this system in their own classroom.

Results and Analysis

Student questions ranged from mundane logistical inquiries to sophisticated testable scenarios. Questions were blinded and randomized so that they could be analyzed for trends and patterns that provide insight into students’ questioning abilities. .

Table 1 outlines nine thematic categories that emerged during open coding of the questions and Table 2 provides additional detail, including their relative frequency. Categories 2 and 3 are procedural – either about the course itself or about mathematical procedures. The majority of questions relate to science content and were differentiated by cognitive levels based on Bloom’s taxonomy (Krathwohl, 2002). They ranged from recall-style questions at the lowest level, questions about explanations (categories 5 & 6), and scientifically testable questions at the highest level (categories 7 & 8). This progression is similar to the structure of categories used by Marbach-Ad and Sokolove (2000) in a university biology course, but this study further differentiates the categories to indicate the amount of synthesis of knowledge from the course (see Table 1). Since most of the students were preservice elementary teachers, about 6% of the questions probed pedagogical concerns such as how course topics could be taught to young children (category 9).

Table 1. Each question was coded into one of nine categories.

	#	**Category**	**Details**			**Common phrases**
🟦	**1**	**No meaningful question**	Blank or non-questions			“I don’t have a question”
🟦	**2**	**Course logistics**	Calendar, grading policy, etc.			“When will we…”
🟦	**3**	**Calculations**	Equations, formulas, quantitative procedures, graph reading			“How do you calculate…” “Is there a formula for…” “Can you go over how to…”
			Cognitive level		**Knowledge required to ask**	
🟩	**4**	**Science Content Questions**	Short Answer/ Factoid/ Definition / Terminology	OR	Vague question, Little knowledge required to ask	“Is __ true?” “How fast do bullets move?” “What is the best way to…” “Is __ a __ or a __?” (terminology)
🟩	**5**		Explanation / How / Why?	AND	Little knowledge required to ask	“Can you explain…” “Why does…” “How does __ work?” “[How] Do ___ [work]?”
🟨	**6**		Explanation / How / Why?		Specific knowledge required to ask	
🟧	**7**		Testable question		Little knowledge required to ask	“What would happen if…” “Could we try…” “Would…”
🟥	**8**		Testable Question		Specific knowledge required to ask	
🟫	**9**	**Application to Education**	Pedagogical questions or how to adapt our learning to the elementary classroom.			“How would you teach this?” “*Why* did you teach ____ this way?”

Table 2. Details about each of the nine categories of questions. Percent indicates what fraction of the total number of questions are in each category.

	#	Category	%	Comments
🟦	**1**	**No meaningful question**	1.6	Asking a question was one item on a daily reflection. Students received credit if they submitted the reflection regardless of whether or not it included a question. Students chose to ask a question more than 98% of the time.
🟦	**2**	**Course logistics**	7.0	Even though students were encouraged to ask questions about scientific processes, about 7% of all questions pertained to course logistics such as due dates, grading policies, etc.
🟦	**3**	**Calculations**	9.4	Since many of the students struggle with mathematics, procedural questions related to calculations were common following lessons where we used mathematical problem solving.
🟩	**4**	**Factoid / Definition**	32	About a third of all questions related to simple facts or clarifications about definitions. These questions represent the lowest level of Bloom's Taxonomy, "remembering."
🟩	**5**	**Explanation**	19	Students seek an explanation of a process or phenomenon that they observed or experienced in everyday life. These explanations largely represent the second tier in Bloom's taxonomy, "understanding."
🟨	**6**	**Explanation demonstrating knowledge**	14	As in category 5, but the question explicitly demonstrates specific knowledge about course topics or asks a question that could only be posed once the student understands the course material. These questions also represent "understanding," but with a higher order of synthesis than category 5.
🟧	**7**	**Testable question**	7.8	Students pose scientifically testable questions about scenarios they have not directly observed. A form of hypothesis generation, these questions represent the top tier of Bloom's taxonomy, "creating."
🟥	**8**	**Testable question demonstrating knowledge**	2.6	As in category 7, but the question explicitly demonstrates specific knowledge about course topics or asks a question that could only be posed once the student understands the course material. These questions also represent "creating," but with a higher order of synthesis than category 7.
🟫	**9**	**Application to Education**	5.9	The course is designed for preservice elementary teachers, so about 6% of the questions probed pedagogical concerns such as how course topics could be taught to young children. Questions from this category do not fit neatly into a level of Bloom's taxonomy. While they often investigated an applications of knowledge to an educational setting ("applying"), they often touched on topics outside of the cognitive domain such as student behavior or attitude towards science.

To minimize complexity, Figure 2 and 3 show results for 18 randomly selected students generated by sorting students from a single course section by quiz score and selecting odd-numbered students from the sorted list. The complete data set, including every question and its assigned category, is included in supplementary files. For all parts of the manuscript except Figures 2 and 3, analyses include all 128 participating students. In Figure 2, each square represents a single question asked by a student after a single class session. The horizontal axis shows the progression of time throughout the semester (29 sessions total). The students are sorted on the vertical axis by their overall quiz score during the semester. White squares in the plot reveal days when the student failed to submit a question, and students with higher quiz scores do not necessarily participate at a higher rate than students with lower quiz scores (see Appendix A for participation data).

Figure 2's multi-colored patchwork indicates that students at all performance levels asked a wide range of questions. More than half the students asked questions from at least six of the categories and 98% asked questions from at least 4 categories. Some students typically asked questions from a single category, while others had questions more evenly distributed between categories. For a typical student, about 40% of his or her questions fell into a single category (38±11%), and category 4 (science content questions at the lowest level of Bloom's taxonomy) was that category for more than half the students (53%).

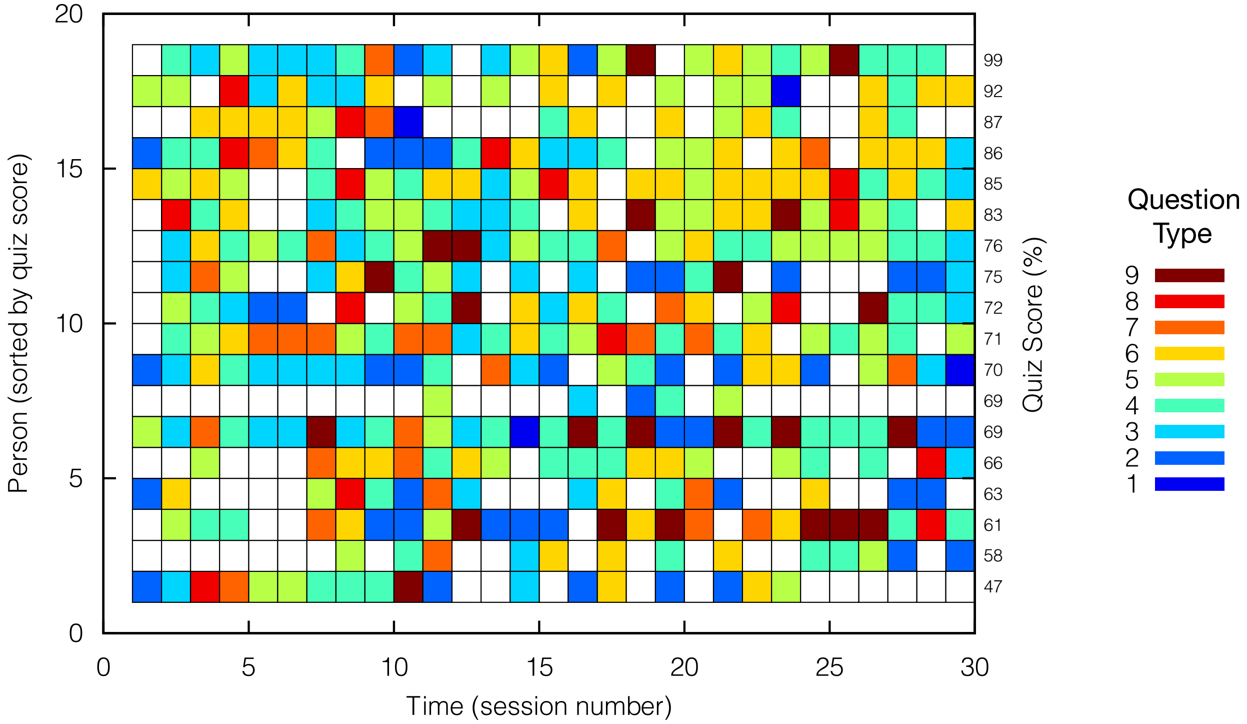


Figure 2. Category of question vs. time for individual students. Colors represent the different categories (see Table 1 for definitions). Each row represents one of 18 randomly selected students (out of the 128 total). Each column represents one class session during the semester. White cells indicate days when the student did not submit a question. Numbers on the right-hand y-axis show the students' average quiz score for the semester, with lowest average quiz score at bottom and highest at top.

To highlight the overall differences between students, Figure 3 shows a single stacked bar with the distribution of categories of questions for the semester for the subset of students shown in Figure 2. There is quite a bit of variation between students. For example, students 3 and 6 asked more than 20% of their questions about applications to education while students 14-17 asked none in that category. At first glance, individual students with similar quiz scores seem to have asked very different types of questions (e.g., note the differences in colors between students 8 and 9 with nearly identical quiz scores of 70% and 71%, respectively).

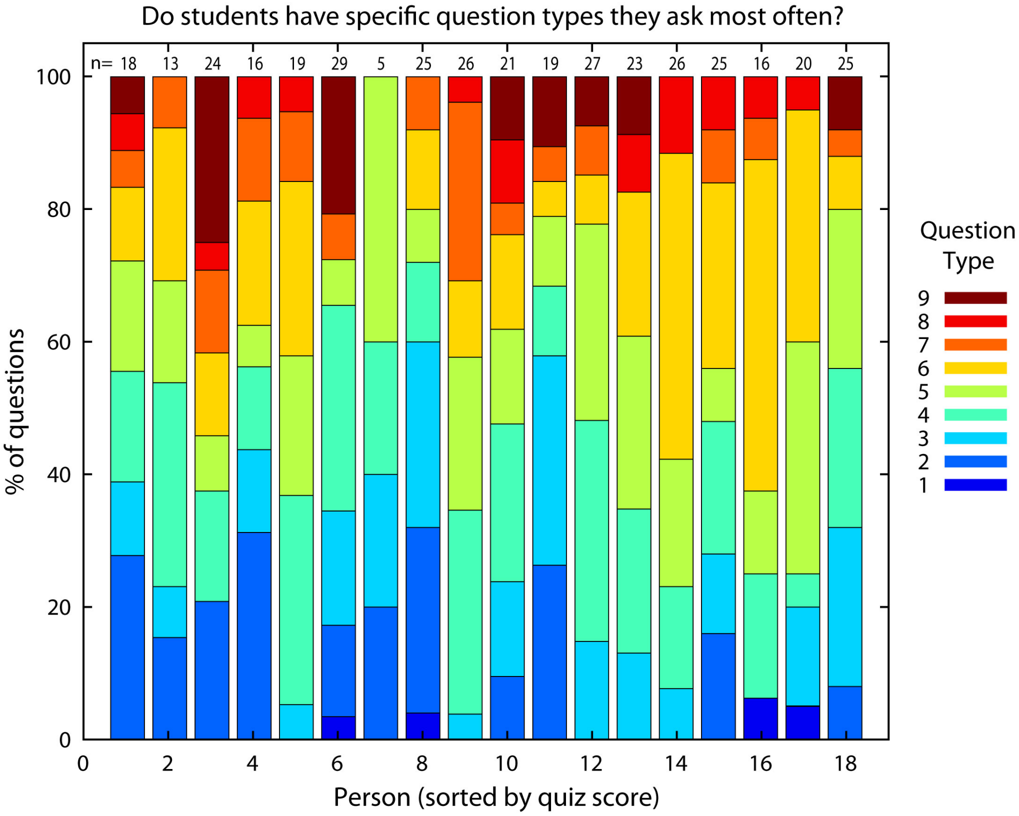


Figure 3. Distribution of question-type by individual students. Each stacked bar represents a single student (same students as Figure 2) sorted by ascending quiz score, and colors represent the relative proportion of different question categories (defined in Table 1). "n=" along the top of each bar indicates the number of questions the student asked during the semester.

In order to investigate trends and patterns related to this study's research questions, combinations of these nine categories that follow two themes were calculated:

***Knowledge* category = category 6 + category 8**. While all questions drew upon prior knowledge to some extent, some questions required specific knowledge of the topic and clearly demonstrated understanding of key concepts discussed in class. To illustrate the difference, consider two questions posed by students following an in class demonstration. One student asked a typical category 7 question, "Would the balloon have popped sooner if we had a bigger flame?" That question probably could have been asked by any college-age student who had seen the demonstration. A second student posed a different testable question that reflects knowledge of the phase transformations discussed in class: "If we had turned off the heat before the balloon popped, would the air that had evaporated [be] turned back into water?" (category 8). Questions in the *Knowledge* category demonstrated and required specific knowledge. *Knowledge* questions were 17% of all submissions.

***Testable* category = category 7 + category 8**. The two questions in the previous paragraph were both scientifically testable – an experiment could be designed to answer them. To emphasize the difference between testable and non-testable questions, consider a third student's question after the same classroom demonstration described in the previous paragraph: "What exactly caused the balloon to pop?" (category 5, not *Testable*). *Testable* questions require a form of creativity where the student must invent a possible scenario that might lead to a different outcome. These questions could expand knowledge by allowing for the testing of a hypothesis (e.g., to answer the questions in the previous paragraph, students could repeat the experiment with a bigger flame, or turn off the heat before the balloon pops and look inside the balloon to see if there is liquid water). It was therefore worthwhile to differentiate the skill where students ask their own scientifically testable questions (*Testable*) from the skill of seeking a scientific fact or explanation. *Testable* questions were 10% of all submissions.

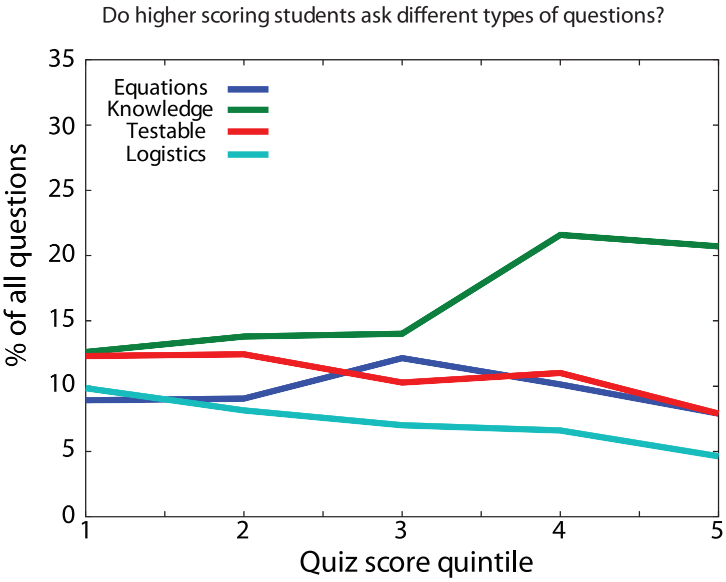


Figure 4. Percentage of questions in several categories vs. Quiz score quintile. Students are grouped into quintiles based on their performance on quizzes. Higher performing students asked fewer logistical and testable scientific questions but more questions that demonstrated knowledge.

Relationship between *Knowledge* questions, *Testable* questions, and quiz score. The average percentage of a student's submissions that were rated in the *Knowledge* category was 16±13%, which means that the majority of questions could probably have been asked by any college student who showed up to the course for the first time that day. There was high variation, and 4% of students had more than half their questions in the *Knowledge* category. Using quiz scores as a proxy for existing knowledge, there was a correlation between a student's knowledge and the type of questions he or she asked (Figure 4). High performing students were nearly twice as likely to ask a *Knowledge* category question than lower performers (21% v. 13%). The correlation between quiz score and percentage of questions in the *Knowledge* category was significant at the 99% confidence level (Table 3), though quiz scores only explained about 6% of the variance between students.

Meanwhile, students with lower quiz scores asked more *Testable* questions than high performers (12% v. 8%). While the difference for *Testable* questions was not as pronounced as *Knowledge* questions, quiz scores were negatively correlated to the percentage of category 7 questions (scientifically testable without demonstrating knowledge) at the 99% confidence level and explained 2.6% of the variance between students.

A better illustration of the negative correlation between the *Knowledge* and *Testable* categories is the ratio between them. High scoring students ask 2.5 times more *Knowledge* questions than *Testable* questions the lowest scoring quintile ask approximately equal numbers in the two categories (*Knowledge*:*Testable* = 1.1:1). Statistically, quiz score explains a full 18% of the variance in the *Knowledge*:*Testable* ratio, and the correlation is significant at the 99% confidence level. In summary, high performing students were more likely to try to fill gaps in their existing knowledge by asking questions that helped them construct explanations for scientific phenomenon (category 6) while low performing students were more likely to ask testable "what if" questions that did not require specific knowledge to ask (category 7).

Table 3. Did a student's quiz score predict the percentage of questions they asked in a specific category? A positive correlation indicates that students with higher quiz scores asked more questions in the given category while dashes indicate that there was no statistically significant correlation. Quiz score only explained a small fraction of the variance in any category.

Question category	% Variance explained (r^2)	Correlation
1 – No meaningful question	2.9*	Negative
2 – Logistics	4.3**	Negative
3 – Equations	--	--
4 – Factoid	2.3*	Positive
5 – Explanation	--	--
6 – Explanation w/ knowledge	6.2***	Positive
7 – Testable question	5.7***	Negative
8 – Testable w/ knowledge	--	--
9 – Pedagogy	--	--
Knowledge (6 + 8)	6.4***	Positive
Testable (7 + 8)	2.6*	Negative
Ratio of *Knowledge*:*Testable*	18.2***	Positive

Asterisks indicate level of statistical significance (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Dashes indicate that $p > 0.1$.

Relationship between questions asked and votes. One interesting aspect of online question-ranking tools is that students engage in the questioning task twice: once to ask questions and a second time to vote on which of their peers' questions they would like to have addressed. This voting can be considered another expression of curiosity. What types of questions interest students?

Figure 5 shows the average number of votes received by questions in each category. A flat curve would indicate that students express no preference for questions based upon their category. Instead, students preferred questions demonstrating knowledge on behalf of the questioner (categories 6 & 8). One instructional concern is that logistical questions might dominate both questions and voting in an online question-ranking tool. While logistical questions were certainly present, testable scientific questions that required knowledge on the part of the questioner (category 8) consistently received about 30% more votes per question than logistical questions. Questions that were too vague or whose answers could be easily looked up online (category 4) also received fewer votes than higher order questions.

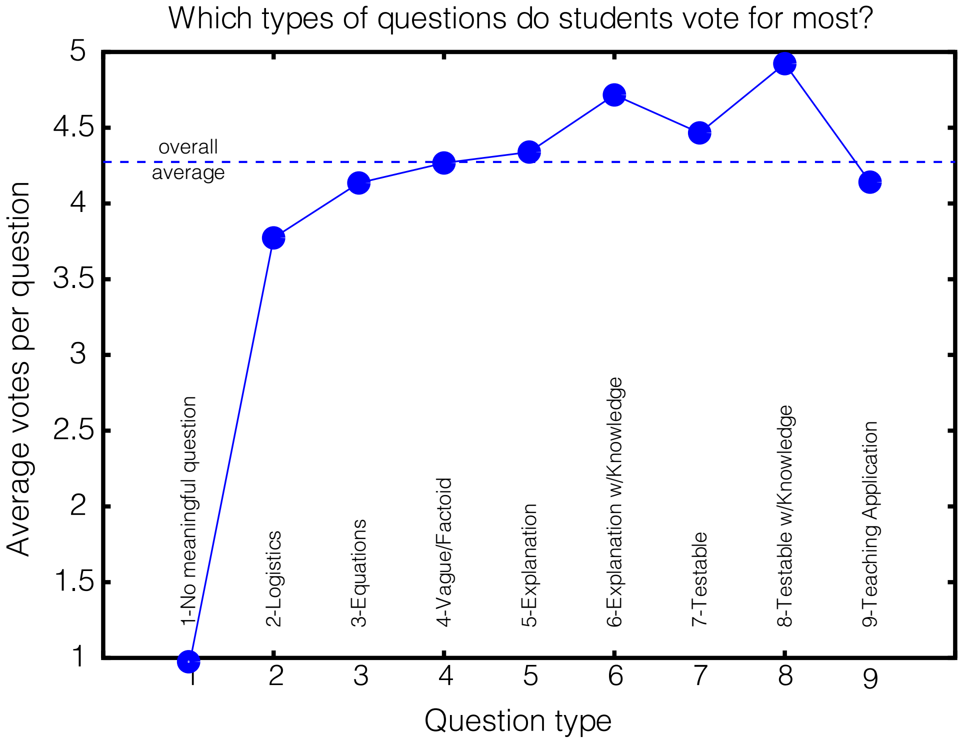


Figure 5. Average number of votes per question for each category of question. Vertical words indicate the categories (see Table 1 for definitions).

Are there differences in the voting preferences of high performing versus low performing students, like demonstrated for question-asking? Individual votes could only be associated to individual students during one course section with 34 students (Google Moderator only gives users access to vote totals. A custom software application was written to track individual voting behavior). The question-asking behavior of this subset of students (Figure 6a) is comparable to that of the entire study (Figure 4); higher performing students asked about twice as many *Knowledge* questions as lower performing students. The distribution of votes (Figure 6b), however, shows much less variation between students. Even though students with low scores asked fewer questions that required knowledge, they recognized and voted for those questions at the same rate as high performing students. Conversely, students with high scores asked fewer logistical questions, but they still wanted to know the answers to those logistical questions at about the same rate as lower performing students.

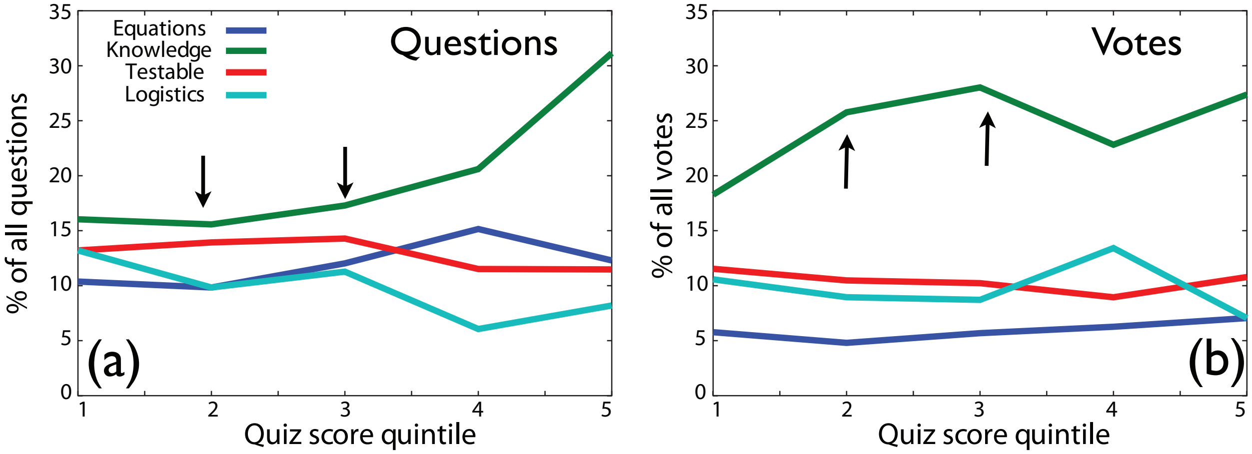


Figure 6. Did students vote for the same type of questions they asked? (a) Types of questions asked by the 34 students for whom votes were tracked (equivalent to Fig. 4, but for only this subset of 34 students) (b) types of questions students voted for. The relatively flatter lines of (b) show that students all voted for similar types of questions regardless of quiz score. The arrows emphasize this effect for questions requiring knowledge: Students in quintiles 2 and 3 for quiz score asked fewer knowledge-based questions, but they voted for them just as often as high-scoring students.

Discussion

The introduction posed an unanswered question about the role of prior knowledge in question-asking. Is curiosity driven by gaps in a store of existing knowledge, or does existing knowledge constrain question-asking and inhibit curiosity? This study's data suggest that both perspectives can be valid in our novice students as different contexts activate either a knowledge-rich or exploration-rich question-asking regime. When students lacked specific prior knowledge of a subject, they asked more questions that probed cause and effect in a very general way (exploration-rich). The questions were either vague (i.e., "why did it do that?") or specific attempts to investigate the cause and effect relationships through testable questions (i.e., "What would happen if…?"). We saw this question-asking regime active in students with low quiz scores (that presumably have not assimilated sufficient background information about the topic). When students were in this exploration-rich but knowledge-poor regime, they also asked a greater percentage of questions about course logistics (i.e., expectations, deadlines). When specific knowledge was present, students asked questions that looked very different. Their questions tended to utilize their existing specific knowledge to fill in gaps (as in Loewenstein, 1994). This study's data suggest that a focus on knowledge gaps may stifle investigative questions – the two are inversely correlated. When students had a strong foundation of existing fact-based knowledge, they sought new information. They asked fewer testable questions that would test or extend their existing mental models of cause and effect. Perhaps they were more focused on clarifying their emerging understanding of explanations based on course material, rather than exploring new realms or testing their existing knowledge. This may be an appropriate way to assimilate new material, but it's also interesting that it shifted the focus away from exploration. Logistical questions about the course dropped substantially in this regime, perhaps because students with higher background knowledge are the ones with the most academic awareness. We saw this question-asking regime in high performing students (who have apparently constructed larger knowledge bases).

There are limits to how far this apparent inverse relationship between questions that demonstrate knowledge and those that are scientifically testable can be extrapolated. Expert scientists are both knowledge-rich and deeply invested in the practice of asking testable questions. They define a third regime. Only 2.6% of the questions in this study were expert-like (both scientifically testable and demonstrating specific knowledge; category 8). It is possible that this study did not see strong evidence of this expert regime because it requires knowledge and experience beyond the level of general education students with limited science backgrounds. The knowledge-rich versus exploration-rich dichotomy applies to novice science students. Experts (presumably) incorporate both knowledge and exploration simultaneously. Understanding how students progress towards the expert-level questioning regime requires research at a broader range of academic levels, but the next section discusses evidence that even novice students are at least aware of the expert regime and are able to recognize it.

Models of higher-order questions

Teachers that ask higher-order questions foster curiosity (Arnone, 2003), have students that ask more questions (Susskind, 1979), and produce students with a better understanding of the nature of science (Kleinman, 1965; Lederman, 1992). In studies of young children, children's ability to ask meaningful questions improves after they are exposed to an adult asking appropriate model questions (Lempers & Miletic, 1983; Zimmerman & Pike, 1972). In one study, students asked six times more questions after hearing a model than a control group exposed to the same situation but with no teacher model of an effective question (Zimmerman & Pike, 1972). With online question tools, the models come from peers rather than the instructor. Results from this study show that high performing students asked more questions that demonstrated knowledge while low performing students asked more scientifically testable questions. Each group could therefore benefit from seeing model questions from the full range of students. Voting data show that all groups were equally skilled at recognizing questions with greater educational value and discriminating those from vague and less educationally valuable questions. In fact, questions that reflect the expert regime (knowledge-rich testable questions, category 8) were the highest voted. Peer questions can therefore serve as models of higher order questions, even when such questions are intermixed with lower order questions in an online question-ranking tool. Future investigations are needed to test whether or not online question-ranking alters student question-asking over time.

Student reaction to question ranking

The introduction of this study described some possible benefits of online question ranking tools based on theoretical perspectives on learning. Insight into how those theories bear out in practice comes from a student reflection activity completed about a third of the way through each semester (see Appendix A). Students were asked to reflect on the motivation for using the online question ranking tool. Students recognized the metacognitive value of online question ranking; nearly every group cited “reflection” or “review” as a reason to vote on questions. Other student comments imply that online question ranking was a low risk experience that may have contributed to a culture where questions were welcomed (e.g., “I have gained the ability to ask questions without the fear of classmates’ laughter”). Students also described how they liked the fact that their interests were addressed when classroom time was devoted to answering the top questions. Voting patterns reveal that, on average, 75% of the class either voted for or authored one of the top five questions every day, meaning that the majority of the class was able to experience the satisfaction of having one of their "preferred" questions being answered. Future work could include a more systematic investigation of student perceptions.

Table 4. Comparison between key question types in this study and previous studies.

	***Knowledge* questions**		***Testable* questions**		***Ratio***
	%	Category	%	Category	*Knowledge:Testable*
This study, Preservice elementary teachers	17%	6+8 ("Knowledge")	10%	7+8 ("Testable")	1.7:1
Marbach-Ad & Sokolove (2000), Introductory biology	22%[b]	4 ("Thoughtful questions")	9%[b]	6 ("Research-like questions")	2.4:1
Keeling et al. (2009), Advanced cell-biology	39%	3 ("Connection or Application")	7%	5 ("Hypothesis or prediction")	7.8:1

[a] All three studies investigate changes in the types of questions over time. For consistency, this table includes the mean over the entire semester in all three courses.

[b] Marbach investigates two experimental groups, a traditional course and an active-learning course. The active-learning course is reported in this table because it is most comparable to the pedagogy used in the present study. For the traditional course, they found 3% in their category 6 and 10% in their category 4.

How generalizable are the findings in this study?

The preservice elementary teachers in this study declared low interest in science and math, but their question-asking behavior has some similarities to other populations in similar science courses. In addition to this study, studies categorizing student questions have been conducted in an introductory biology course (Marbach-Ad & Sokolove, 2000) and a senior-level undergraduate cell biology course (Keeling et al., 2009). While the three studies each developed unique categorization schemes, they share a category for scientifically testable questions and identify questions that demonstrate connections to existing knowledge (Table 4). The distribution of questions from the introductory biology students is statistically comparable to that of preservice teachers in this study, though the preservice teachers asked a slightly lower percentage of *Knowledge* questions than their general education counterparts. Despite the fact that the advanced cell biology students had a declared major in a scientific field and several years of specialized training in methods of scientific inquiry, they consistently asked fewer scientifically testable questions than students in this study or the introductory biology course. Keeling et al. (2009) also noted that introductory students asked more testable questions than advanced ones and was concerned that different investigators may have used different degrees of rigor in their assignment of this category. While a valid concern, the trend's consistency with the internal findings of this study is suggestive of a pattern. The inverse relationship between *Knowledge* and *Testable* questions identified within the preservice teachers in this study may in fact be amplified as students gain more knowledge in the academic setting (see "Ratio" column in Table 4). The connection to how these students become expert questioners is not apparent in the current data because additional knowledge alone, at least up to the senior undergraduate level, does not seem sufficient to activate the expert regime of questioning.

Implications for instructional design

A common learning objective for general education science classes is to help students understand the methods of scientific inquiry and think like expert scientists. This includes asking questions about the world around them. A consistent finding of the studies described in this paper is that relatively few of the questions students ask are scientifically testable.

Keeling et al. (2009) proposed an intervention to explicitly teach students about the different categories of questions and their relationship to scientific inquiry. Keeling showed that this intervention produced statistically significant shifts towards more exploration-rich questions. While promising, there was no non-intervention comparison case and the course material before and after the intervention was different. Figure 2 illustrates how variable the questions were by the preservice teachers in this study as the content changed throughout the semester. Any apparent change in question behavior over time needs to be considered in the context of the changing content. Additional investigation is warranted, and online question-ranking tools could be a useful tool for evaluating such interventions because they allow monitoring of both question generation skills and students' ability to recognize higher order questions through voting.

The efficacy of such an intervention may depend on the reason that students doing poorer on assessments ask more exploration-rich questions. Scardamalia and Bereiter (1992) documented how existing knowledge can actually constrain question-asking, but they do not have a complete explanation about why. One possible mechanism is that students focusing on knowledge-rich questions have "learned how to do school well," meaning they know the type of items typically found on assessments target their questions towards information that will advance that agenda (H. Hertzog, personal communications, May 30, 2014). In other words, we have trained our students to stop asking questions that express wonder by the ways in which we assess them. Students that ask exploration-rich questions may not be as skilled at building assessment-ready knowledge by this process and therefore defer to their innate child-like inquisitiveness. The lack of questioning in elementary classrooms described in the introduction is evidence that this conditioning begins very early on.

An effective intervention to promote exploration-rich questions may need to focus on: 1) changing course assessments to allow more opportunities to engage in authentic scientific inquiry at a higher cognitive level; 2) making students aware of the importance of questions in scientific inquiry (as in Keeling et al., 2009); 3) providing models of higher order, expert-like questions for students to compare against (perhaps using online question-ranking tools); and 4) deconditioning students from their habit of always seeking answers to known questions that are likely to appear on assessments. A simple awareness of question types may not be sufficient to decondition students, but instead a more fundamental shift in course structure may be required to bring students closer to asking questions like expert scientists. Such shifts will provide more opportunities for students to express and pursue their innate curiosity through authentic questioning.

Conclusions

An online question-ranking tool provides the forum for students to ask open-ended questions and review examples of other questions by their peers. This study is a systematic investigation of the types of questions students asked using this tool. Student questions were grouped into nine categories based primarily on their cognitive level, but also on how much knowledge they demonstrate on behalf of the questioner. In general, individual students seemed to ask a wide variety of questions regardless of their level of performance in the class, though there were subtle but revealing patterns. There was an inverse relationship between asking testable scientific questions and questions that seek explanations to fill gaps in knowledge. When students had less specific knowledge about a topic, they tended to ask a slightly greater percentage of scientifically testable questions. As a whole, students with higher quiz scores asked more questions that required knowledge about the course content (especially questions that sought explanations for phenomena), but they asked fewer scientifically testable questions than students with lower quiz scores. Students did not need to have high scores to express curiosity and ask valuable questions, but the types of questions a low scoring student asked reflect a different aspect of curiosity than expressed by their high scoring peers.

Expert scientists ask questions from a third regime that includes testable questions that also reflect and utilize specific knowledge, and voting data show that students could recognize and preferred these questions. When evaluating peer questions using an online question-ranking tool, students at all performance levels had a slight preference for the highest order questions that both demonstrated knowledge and were scientifically testable (these received the most votes per question – an average of 30% more than the lowest rated question category of classroom logistics).

Online question-ranking tools could help shift the focus of these classrooms from a culture of “right answers” to a culture of questioning. That shift will enable teachers to truly harness the power of curiosity to improve learning in science.

Acknowledgements

The author enthusiastically thanks H. Hertzog, S. Klein, L. Lundquist, N. O’Rode, and J. Reveles at CSUN for thoughtful discussions and critical review of the manuscript at different stages. Their comments, along with those of four anonymous reviewers greatly improved the manuscript’s focus. This research was funded internally by reassigned time granted by the CSUN College of Science & Mathematics.

Appendix A: Classroom implementation notes

Grading. The daily reports are 10% of the students' course grade. Despite this fact, completion rates were only about 60%, though varied throughout the semester (Figure A1). Students anecdotally cited technology access issues, but many regularly forgot about the 24 hour deadline (the online tools can be easily accessed from a smart phone). Voting was not graded because Google Moderator does not allow tracking of individual votes. I encouraged voting both through frequent discussions of the importance of the process as well as awarding bonus points for 1) the whole class of participation reached 90% (this never happened); and 2) for individuals whose questions were chosen as the "most popular" question of the day (giving students incentive to log in to at least vote for their own question). In order to track voting for research purposes during the final semester of the project, I wrote a web application that duplicates the functionality and appearance of Google Moderator that could also log individual voting practices. About half way through the semester, I began offering a ¼ point bonus to students for each session in which they voted. This incentive had no effect on the number of voting participants (20.85 versus 20.92 average voters per session out of 35 students enrolled).

Meta-training. About a third of the way through the semester, I asked students to reflect in writing on why we ask questions and vote on them. I then picked ten representative responses to this prompt and asked the students to work in teams to categorize them (students were free to construct and name the categories however they wished). Table A1 shows a representative grouping. Figure A1 shows that this intervention may have caused a very slight but insignificant increase in voter participation.

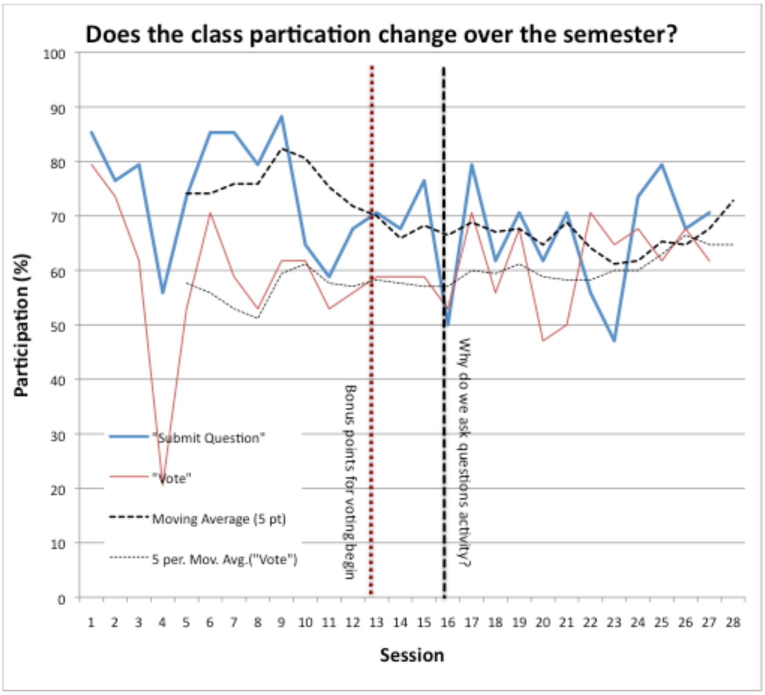


Figure A1. Overall class participation rate for submitting questions and voting

Table A1. Students responses to the prompt, "Why do I make you ask questions and vote on them?

Reflection	Content review	Social
• “When I see others questions it makes me think more of ‘oh yea why is that?’” • “to reflect on ideas that we did not completely understand.” • “To see what good questions look like, and help to improve our question-asking skills. • “I have learned that some answers to our questions can be found if we reread the packets we do in class.”	• “obviously for them to be answered” • “I didn't learn much initially until we began answering them at the beginning of the class” • “student can think about the question during voting and maybe formulate an answer to it.” • “allows us to look at the main idea of today in a different perspective.”	• “I have gained the ability to ask questions without the fear of classmates laughter etc.” • “You make us vote to encourage us to participate as a class.”

References

Arnone, M. P. (2003). Using Instructional Design Strategies To Foster Curiosity. *ERIC Digest.* Retrieved from http://www.ericdigests.org/2004-3/foster.html

Arnone, M. P., & Small, R. (2011). *Arousing and Sustaining Curiosity: Lessons from the ARCS Model* (SSRN Scholarly Paper No. ID 1782259). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=1782259

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). How People Learn: Brain, Mind, Experience, and School. Expanded Edition. Retrieved from http://eric.ed.gov/?q=%22how+people+learn%22&ff1=autBransford%2c+John+D.%2c+Ed.&id=ED481522

Day, H. (1968). Role of Specific Curiosity in School Achievement. *Journal of Educational Psychology*, *59*(1, Pt.1), 37–43. doi:10.1037/h0025460

Donohue-Smith, M. (2006). Improving the Questions Students Ask. *Education Digest: Essential Readings Condensed for Quick Review*, *72*(3), 41–43.

Engel, S. (2009). Is Curiosity Vanishing? *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(8), 777–779. doi:10.1097/CHI.0b013e3181aa03b0

Engel, S. (2011). Children's Need to Know: Curiosity in Schools. *Harvard Educational Review*, *81*(4), 625–645.

Froschauer, L. (2010). No Dumb Questions, But... *Science and Children*, *48*(4), 6–6.

Google. (n.d.). *lesson 6.2.1b - Google Apps Education Training Center*. Retrieved April 16, 2013, from http://edutraining.googleapps.com/Training-Home/module-6-other-tools/chapter-2/1-2

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (2000). *The Scientist in the Crib: What Early Learning Tells Us About the Mind*. New York: William Morrow Paperbacks.

Hackling, D. M., & Garnett, D. P. (1992). Expert — Novice differences in science investigation skills. *Research in Science Education*, *22*(1), 170–177. doi:10.1007/BF02356893

Jirout, J. J. (2011). Curiosity and the Development of Question Generation Skills (pp. 27–30). Presented at the AAAI Fall Symposium Series. Retrieved from http://www.aaai.org/ocs/index.php/FSS/FSS11/paper/view/4194

Johnson, L., & Beer, J. (1992). Specific and diversive curiosity in gifted elementary students. *Perceptual and Motor Skills*, *75*(2), 463–466.

Keeling, E. L., Polacek, K. M., & Ingram, E. L. (2009). A Statistical Analysis of Student Questions in a Cell Biology Laboratory. *CBE Life Sciences Education*, *8*(2), 131–139. doi:10.1187/cbe.08-09-0054

King, A. (1994). Guiding Knowledge Construction in the Classroom: Effects of Teaching Children How to Question and How to Explain. *American Educational Research Journal*, *31*(2), 338–368. doi:10.3102/00028312031002338

Kleinman, G. S. (1965). Teachers' questions and student understanding of science. *Journal of Research in Science Teaching*, *3*(4), 307–317. doi:10.1002/tea.3660030410

Kowalski, F. V., & Kowalski, S. E. (2012). Enhancing curiosity using interactive simulations combined with real-time formative assessment facilitated by open-format questions on Tablet computers. In *Frontiers in Education Conference (FIE), 2012* (pp. 1–6). doi:10.1109/FIE.2012.6462282

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. doi:10.1207/s15430421tip4104_2

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, *29*(4), 331–359. doi:10.1002/tea.3660290404

Lempers, J. D., & Miletic, G. (1983). The immediate and delayed effects of different modeling strategies on children's question-asking behavior with different kinds of messages. *The Journal of Genetic Psychology: Research and Theory on Human Development*, *142*(1), 121–133.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98. doi:10.1037/0033-2909.116.1.75

Marbach-Ad, G., & Sokolove, P. G. (2000). Can undergraduate biology students learn to ask higher level questions? *Journal of Research in Science Teaching*, *37*(8), 854–870. doi:10.1002/1098-2736(200010)37:8<854::AID-TEA6>3.0.CO;2-5

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, D.C.: The National Academies Press.

NGSS Lead States. (2013). Appendix F: Science and Engineering Practices in the NGSS. In *Next Generation Science Standards: For States, By States*. Washington, D.C.: The National Academies Press.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about-science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, *40*(7), 692–720. doi:10.1002/tea.10105

Peterson, R. W. (1979). Changes in curiosity behavior from childhood to adolescence. *Journal of Research in Science Teaching*, *16*(3), 185–192. doi:10.1002/tea.3660160302

Ram, A. (1991). A Theory of Questions and Question Asking. *The Journal of the Learning Sciences*, *1*, 273–318.

Scardamalia, M., & Bereiter, C. (1991). Higher Levels of Agency for Children in Knowledge Building: A Challenge for the Design of New Knowledge Media. *Journal of the Learning Sciences*, *1*(1), 37–68. doi:10.1207/s15327809jls0101_3

Scardamalia, M., & Bereiter, C. (1992). Text-Based and Knowledge-Based Questioning by Children. *Cognition and Instruction*, *9*(3), 177–199. doi:10.2307/3233513

Smith, K. H. (2010). *Curiosity and Pedagogy: A Mixed-Methods Study of Student Experiences in the Design Studio*. University of Arkansas, Fayetteville. Retrieved from http://udini.proquest.com/view/curiosity-and-pedagogy-a-mixed-goid:305185216/

Susskind, E. (1979). Encouraging teachers to encourage children's curiosity a pivotal competence. *Journal of Clinical Child Psychology*, *8*(2), 101–106. doi:10.1080/15374417909532896

University of California Museum of Paleontology. (2013). *Understanding Science*. Retrieved February 2, 2013, from http://www.understandingscience.org

Zimmerman, B. J., & Pike, E. O. (1972). Effects of Modeling and Reinforcement on the Acquisition and Generalization of Question-Asking Behavior. *Child Development*, *43*(3), 892–907. doi:10.2307/1127640