International Journal of Science Education
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tsed20

Quantitative Reasoning in Environmental Science: A learning progression

Robert Lee Mayes [a], Jennifer Harris Forrester [b], Jennifer Schuttlefield Christus [c], Franziska Isabel Peterson [d], Rachel Bonilla [e] & Nissa Yestness [f]

[a] Department of Teaching and Learning, Georgia Southern University, PO Box 8134, Statesboro 30460, GA, USA

[b] Elementary and Early Childhood Education, University of Wyoming, Laramie, WY, USA

[c] Department of Chemistry, University of Wisconsin Oshkosh, Oshkosh, WI, USA

[d] Secondary Education, University of Wyoming, Laramie, WY, USA

[e] Department of Mathematics, Georgia Southern University, Statesboro, GA, USA

[f] NREL, Colorado State University, Fort Collins, CO, USA

Published online: 29 Jul 2013.

To cite this article: International Journal of Science Education (2013): Quantitative Reasoning in Environmental Science: A learning progression, International Journal of Science Education, DOI: 10.1080/09500693.2013.819534

To link to this article: http://dx.doi.org/10.1080/09500693.2013.819534

International Journal of Science Education, 2013
http://dx.doi.org/10.1080/09500693.2013.819534

Quantitative Reasoning in Environmental Science: A learning progression

Robert Lee Mayes[a*], Jennifer Harris Forrester[b], Jennifer Schuttlefield Christus[c], Franziska Isabel Peterson[d], Rachel Bonilla[e] and Nissa Yestness[f]

[a]Department of Teaching and Learning, Georgia Southern University, PO Box 8134, Statesboro 30460, GA, USA; [b]Elementary and Early Childhood Education, University of Wyoming, Laramie, WY, USA; [c]Department of Chemistry, University of Wisconsin Oshkosh, Oshkosh, WI, USA; [d]Secondary Education, University of Wyoming, Laramie, WY, USA; [e]Department of Mathematics, Georgia Southern University, Statesboro, GA, USA; [f]NREL, Colorado State University, Fort Collins, CO, USA

The ability of middle and high school students to reason quantitatively within the context of environmental science was investigated. A quantitative reasoning (QR) learning progression was created with three progress variables: quantification act, quantitative interpretation, and quantitative modeling. An iterative research design was used as it is the standard method for the development of learning progressions. The learning progression was informed by interviews of 39 middle and high school students from 5 schools in the Western USA using QR assessments. To inform the lower anchor, intermediate levels, and upper anchor of achievement for the QR learning progression, an extensive review of the literature on QR was conducted. A learning progression framework was then hypothesized. To confirm the framework, three QR assessments within the context of environmental literacy were constructed. The interviews were conducted using these QR assessments. The results indicated that students do not actively engage in quantitative discourse without prompting and display a low level of QR ability. There were no consistent increases on the QR learning progression either across grade levels or across scales of micro/atomic, macro, and landscape.

Keywords: *Environmental education; Model-based learning; Qualitative research*

*Corresponding author: Department of Teaching and Learning, Georgia Southern University, PO Box 8134, Statesboro 30460, GA, USA. Email: rmayes@georgiasouthern.edu

Introduction

The Science, Technology, Engineering, Mathematics (STEM) challenge for our education system is twofold. First, the education system must increase and then sustain a STEM pipeline of students to serve as the next generation of scientists, engineers, and mathematicians, who will help research and solve grand challenges in areas such as environment (National Research Council, 2001) and energy. These challenges include biogeochemical cycles of carbon, biological diversity and ecosystem functioning, and hydrological forecasting of freshwater resources. Second, the education system needs to produce scientifically literate citizens that can make informed decisions about grand challenges. Economic, policy, and social issues will converge around the grand challenges forcing citizens to make decisions that will impact the future of their resources. These two desired outcomes will require that STEM education for all students have deeper learning experiences facilitating the understanding and use of key scientific concepts to interpret, evaluate, and solve real-world problems.

Current codified abstract sequences in school science, such as biology, chemistry, earth science and physics, do not capture the complexity or interdisciplinary nature of reasoning about grand challenges in environment and energy. However, perspectives such as teaching science as modeling (Duschl, Schweingruber, & Shouse, 2007) promote a more viable pedagogy for educating citizens who can reason about grand challenges. This experience will require students who can reason both qualitatively and quantitatively. The ability to reason quantitatively is essential for a citizen of a democracy, allowing them to make informed data-based decisions at home, in the workplace, and on complicated national and international issues that impact their local community. If quantitative reasoning (QR) is to serve as a trigger for interdisciplinary problem-based pedagogies in STEM teaching, then more needs to be known about the trajectory or progression of the QR development in STEM.

The development of learning progressions provides an iterative research design that explicates trajectories of learning over long periods of time. The purpose of this study was to develop a learning progression for QR with environmental sciences as a context. The development included creating QR assessments and interview protocols addressing questions of the QR development. The theoretical implications of the study are a learning progression that informs the development of QR and the role of QR in interdisciplinary problem-based learning. The QR assessments are not presented in the article due to length considerations, but are available for review upon request from the authors.

Literature Review

The capacity to reason quantitatively within real-world contexts which impact one's life has many names, including numeracy, number sense, deductive reasoning, mathematical literacy, quantitative literacy, problem solving, contextualized mathematics, mathematical modeling, and QR. As a result, a clear definition of QR and its components is needed for this study.

First, there is consideration of the quantification act (QA) itself, the mathematical process by which one moves from context to quantity and back to context. Thompson (2011) defines quantification as the process of conceptualizing an object and an attribute of it so that the attribute has a unit measure, and the attribute's measure entails a proportional relationship (linear, bi-linear, or multi-linear) with its unit. Quantification requires conceptualization and reconceptualization in relation to each other of the object being quantified, the attributes of that object, and the measure of the attribute. Quantification is known to be a significant component in modeling and has been found to be difficult for students (Thompson, 2011). Part of the conceptualization process of the QA is the ability to conceive of the problem mentally through an image. Research indicates that the QA involves the development of mental images (Moore, Carlson, & Oehrtman, 2009) and correlational reasoning (Carlson, Jacobs, Coe, Larsen, & Hsu, 2002; Lobato & Siebert, 2002; Thompson, 1994).

Second, the process of modeling and interpreting models is gaining prominence and should be given consideration as a component of QR. *Taking science to school* (Duschl et al., 2007) makes the call to move learning toward literacy and modeling practices in the sciences. *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas* (National Research Council, 2011) also promotes modeling in science. Science as model-building is defined as learning science as a process of building theories and models using evidence, checking them for internal consistency and coherence, and testing them empirically (Duschl et al., 2007). The seminal research done by Schwarz et al. (2009) in the Modeling Designs for Learning Science (MoDeLS) project defined scientific modeling as elements of practice including constructing, using, evaluating, and revising scientific models, and the metaknowledge that guides and motivates the practice. Their learning progression for scientific modeling has two dimensions: (1) scientific models as tools for predicting and explaining, and (2) models change as understanding improves. MoDeLS provides a scientific qualitative account of modeling, this needs to be expanded to include the quantitative science account of modeling across grades 6–12. Lesh, Middleton, Caylor, and Gupta (2008) view modeling from a mathematical perspective, using model-eliciting activities to reveal students' difficulties with the complexity of real-world systems, conceptual systems, and understanding how they develop models.

For a more detailed discussion of QR in the sciences, see Mayes, Peterson, and Bonilla (2012).

Theoretical framework. The definition of QR used in this study, derived from the literature above as well as work conducted by the researchers on QR in STEM through the National Science Foundation (NSF) Pathways project is

> Quantitative reasoning is mathematics and statistics applied in real-life, authentic situations that impact an individual's life as a constructive, concerned, and reflective citizen. QR problems are context dependent, interdisciplinary, open-ended tasks that require critical thinking and the capacity to communicate a course of action.

We propose that QR has four fundamental components:

- *Quantification act (QA)*: mathematical process of conceptualizing an object and an attribute of it so that the attribute has a unit measure, and the attribute's measure entails a proportional relationship (linear, bi-linear, or multi-linear) with its unit
- *Quantitative literacy (QL)*: use of fundamental mathematical concepts in sophisticated ways
- *Quantitative interpretation (QI)*: ability to use models to make predictions and discover trends, which is central to a person being a citizen scientist
- *Quantitative modeling (QM)*: ability to create representations to explain a phenomena

These components interact within a QR cycle when students engage in the process of science as model-building. First, the individual engages in the QA by identifying objects, their attributes, and assigning measures. This provides variables that can be operated on mathematically or statistically. Second, depending on both the query of interest to the individual and the data they access, they engage in QR through the three components of quantitative literacy, QI, or QM. These three components are interconnected and typically engaging in one requires elements of another.

Learning progressions are central to the theoretical framework for the study. *Taking science to school* (Duschl et al., 2007) recommends that learning and curriculum designs be organized around learning progressions as a means of supporting learners' development toward attaining the four proficiencies in science which are know, use, and interpret scientific explanations of the natural world, generate and evaluate scientific evidence and explanations, understand the nature and development of scientific knowledge, and participate productively in scientific practices and discourse. The Consortium for Policy Research in Education report *Learning progressions in science: An evidence-based approach to reform* (Corcoran, Mosher, & Rogat, 2009) identified learning progressions as a promising model that can advance effective adaptive instruction teaching techniques and thereby change the norms of practice in schools. We hypothesize that QR is essential for data-based and modeling approaches to learning the sciences. We are developing learning performance assessments that inform the progressions and teachers' adaptive instruction strategies.

Creating learning progressions is an iterative research process that involves grounding the lower anchor in domains that for this study are accessible to sixth graders, then identifying intermediate levels of understanding through which they pass on their way to attainment of the upper anchor. The learning progression upper anchor is based on expert views of what QR a scientifically literate citizen should know and be able to do by the 12th grade. While learning progressions in science have incorporated some components of QR (Louca, Zacharia, & Constantinou, 2011; Pluta, Chinn, & Duncan, 2011; Schwarz et al., 2009; Stefani & Tsaparlis, 2009; Taylor & Jones, 2009), there presently is no research in science education on a progression examining either the QA or the trajectory of quantitative literacy, QI, and QM supporting science as model-building. Science progressions that are especially pertinent to our study are those that integrate significant quantitative components such as the modeling in science (Schwarz et al., 2009), data modeling and evolution (Lehrer & Schauble, 2002), and atomic molecular theory (Smith, Wisner, Anderson, & Krajcik, 2006) learning progressions.

Purpose and Rational

The purpose of this study is to establish a learning progression for QR within the context of environmental science for middle and high school students. Establishing such a progression requires collecting and analyzing data that informs the lower anchor as well as the intermediate levels of the progression. The work to this point is *preliminary*, as the iterative nature of learning progression research requires revisions based on student interviews and written assessments of QR. The current study presents qualitative data collected through student interviews that was used to inform the hypothesized learning progression framework presented in this paper.

Research questions. The central research question for this qualitative study is

- How do students develop QR in the context of environmental science across 6th–12th grade?

To study this central question, the following procedural questions were addressed:

- What are the QR progress variables (dimensions of understanding, application, and practice) that support the development of an environmentally literate citizen?
- What level of QR within the context of environmental science do students bring to the discourse at the sixth grade level?
- What are the key QR conceptual stepping stones to moving from a novice to environmentally literate citizen? How do these inform a QR learning progression?
- What are the QR tasks students at a given learning progression level should be capable of performing?

Methods

Participants

Thirty-nine middle and high school science and mathematics students from five schools in the Western USA participated in the study. The sample was 54% male and 46% female and was almost entirely Caucasian, which reflects the demographics of the state in which the research was conducted.

Materials

To inform the lower anchor, intermediate levels, and upper anchor of achievement for the QR learning progression, an extensive review of the literature on QR was conducted. A learning progression framework was then hypothesized. To confirm the framework, three QR assessments within the context of environmental literacy were constructed. The interviews were conducted using these QR assessments, each one based on a key conceptual strand identified by the NSF Pathways[1] science teams: QR Carbon Cycle, QR Water Cycle, and QR Biodiversity. Each assessment was organized across three levels of scale: beginning with the macro-scale (personal experience of the world, what can be seen with the eye), followed by a question on the landscape scale (global generalizations, what could be seen with a telescope or larger), and finishing with amicro/atomic scale question (hidden mechanisms, what could be seen with a microscope or smaller).

Why scale? Researchers in the current study hypothesized that as students moved from the macro-scale to landscape scale or micro/atomic scale; there may be a greater need to engage in QR. Tretter, Jones, and Minogue (2006) found students had difficulty in reasoning beyond the limit of visibility. Their understanding of scale dropped significantly at the threshold of the microscopic scale, which is a major concern since the hidden mechanisms behind macro-observations are in this realm. While a similar well-defined barrier was not found at the landscape scale, large scale accuracy declined in a smooth, uniform fashion as scale increased. The research indicated the importance of context as well as experiences when moving to the macro- or landscape-scale. Additionally, Tretter, Jones, Andre, Negishi, and Minogue (2006) discussed the need to support qualitative differences in scale with quantitative differences. This included conceptions of a universal referent which we explore as an element of measure and proportional reasoning as prerequisite to what they refer to as unitizing to understand scale.

Procedure

Interview protocols were created based on each of the assessments and shared with the four persons conducting interviews to improve consistency. Science and mathematics teachers in the participating school districts were asked to select students whom they considered of moderate and high level ability to participate in the interviews. Selection of participants was at the discretion of the teachers, so selection bias is not controlled. No baseline or pre-testing was conducted to establish sample distributions. The purpose of this exploratory study was to establish the current level of QR ability under existing curriculum and teaching, there was no treatment. Environmental science was selected as the context for QR to parallel the on-going scientific research in the NSF Pathways project. Table 1 provides the sample distribution across grades and the means for overall QR and the three components of QA, QI, and QM.

Students were randomly assigned to participate in one of the three assessment interviews water, carbon, or biodiversity. The 30–40 min audio-taped interviews were conducted in the students' school building. All 39 interviews were transcribed and coded, with NVivo employed as a qualitative research tool to identify themes. A Grounded Theory research design was employed, with multiple coders working to reach consensus. An analysis of code distribution was conducted to inform revisions in the QR assessments as well as the QR learning progression.

To attempt to determine the trajectory of QR development within each science content strand (carbon, water, and biodiversity), a qualitative analysis of interviews

Table 1. QR by grade

Grade	Number of students	Total QR mean	QA mean	QI mean	QM mean
6	3	1.75	1.78	1.78	1.50
7	7	1.77	1.90	1.67	1.00
8	12	2.01	2.08	2.06	1.33
9	4	2.22	2.25	2.17	2.33
10	3	2.08	2.11	2.00	–
11	7	1.96	1.90	2.05	1.75
12	3	1.67	1.78	1.78	1.00
Cor[a]		0.076	0.012	0.294	0.026

[a]Correlation between grade level and QR processes.

for a selected track of students was conducted. These students were selected from the full sample of students completing interviews. For each of the three science content strand assessments, one student was randomly selected from each grade Levels 6–12, providing a representative trajectory of students across which the QR development was tracked. For example, we selected one student from each grade level who completed the carbon interview. For some science content strands, we did not have students at a given grade level who completed an interview. A qualitative case analysis for each of the tracks was developed and used to revise the learning progression that was originally theoretical and literature based.

Each student in the track was then rated on QR ability using the revised learning progression (Appendix 1). Two researchers rated each student on the three scales (macro, landscape, micro/atomic) at three QR component levels (quantification/literacy, interpretation, and modeling), compared the nine scores per student, then came to consensus on the scores for each student. After building consensus on how to rank students using the revised learning progression, all remaining student interviews were scored by one researcher for each strand. Descriptive statistics were employed to provide a preliminary picture of trends across grade levels, science scales, and QR process levels both within strands and across strands. Results from both the qualitative and quantitative analysis of the interviews were used to inform further revisions of both the QR assessments and QR learning progression.

Results

Development of learning progressions is an iterative design-based research process, where consensus on the progression is built by testing proposed progressions against student data. The research presented herein is early in the iterative cycle, so the results say more about verifying the assessments and learning progression than they do about the actual level of student QR ability. The following sections provide preliminary quantitative data on students' QR ability, and then move to qualitative data supporting assessment and learning progression development.

Learning Progression Ranking

Overall, 39 students were ranked using the learning progression, providing 9 scores per student for a total of 351 rankings. Rankings ranged from Level 1 (lower anchor-novice) to Level 4 (upper anchor-expert) and included a no evidence/not asked ranking which was assigned by raters when no question was asked by the interviewer in that category. The overall distribution of rankings is given in Table 2. The no evidence (NE)/not applicable (NA) distribution indicates that overall the assessment interviews failed to elicit responses on QM (16% response rate). This may be due to limiting the interviews to 30 min, which resulted in the interviews focusing more on QL and QI interview questions which required less response time. In addition, the QM questions were the last questions on the interview. Another trend is that no student response was ranked at Level 4. While this is concerning, the iterative nature of the assessments will tease out if the expectations for the levels need adjusting or if students are not reaching the desired level. Only 14% of rankings were at Level 3, indicating that these students did not display a sophisticated level of QR abilities. The predominate ranking was 2 (66%), which may signify that the learning progression needs to have items at this level distributed to other levels to provide more sensitivity to student differences.

As a student moves from grade 6 to grade 12, they are exposed to more and deeper science and mathematics conceptions. How does this impact the student's QR abilities? Correlations between grade level and the QR processes indicate a similar trend for overall QR, QA, and QI across all grade levels: there is a small increase from grade 6 to grade 9, followed by a small decline from grade 10 to grade 12 (Table 1). Though only a small population of students was sampled, a trend was apparent. In fact, all the values round to Level 2 except for those in QM, which are suspect due to the small amount of data in this area.

Issues of scale in science are paramount. How does QR vary across scales? Analysis was restricted to QA, QI, and overall QR (due to minimal data on QM). The research team was also interested in examining whether or not there were differences across QR processes within scales. Mean values of all 39 students' rankings on QA, QI, and combined QA and QI (overall QR) across science scales

Table 2. Overall learning progression rankings

Ranking	Macro-scale			Landscape scale			Micro/atomic scale			
QR level	QA	QI	QM	QA	QI	QM	QA	QI	QM	Total
1	5	9	7	6	5	3	7	9	0	51
2	28	27	6	28	24	2	26	26	0	167
3	6	3	1	5	10	0	6	4	0	35
4	0	0	0	0	0	0	0	0	0	0
NE/NA	0	0	25	0	0	34	0	0	39	98

QA, quantification act; QI, quantitative interpretation; QM, quantitative modeling.

were examined. A 3×3 analysis of variance (ANOVA) was performed and there were no significant differences across scales or processes (Table 3). One could hypothesize that students would perform best at the macro-scale, since here the comfort of their own convictions potentially reduce cognitive load freeing up capacity to bring quantitative discourse to the table. Additionally, one might theorize that students would perform worse on the micro/atomic scale since they cannot interact directly but must observe through tools. Neither of these hypotheses was supported. From a process perspective, one could hypothesize that QA would be better than QI since it is a more fundamental process. It was not. Perhaps because students practice QI with tables and graphs, even if they do not fully comprehend the quantities they are comparing.

Disaggregation of the data by strand reduces the sample size making statistical analysis untenable, but the use of different assessments for each strand makes examination of trends in each strand imperative. We, therefore, provide a short overview of potential trends by the biodiversity, carbon, and water strands. There is not much deviation between strands in the overall rankings, with a Level 2 ranking being predominate (biodiversity 76%, carbon 64%, and water 60%).

Trends for QR across grade levels vary across strands. For QR biodiversity, there is a positive increasing trend across grade levels, but it is leveling off at 10th grade, and we do not have data on the upper grades to determine if it dips in later grades. If the exceptional sixth grader is removed from the QR carbon data, then both the carbon and water strands exhibit an increase followed by a decrease in QR ability. The failure of QR rankings to correlate positively with the grade level increase could be due to limitations of the assessments. The increasing and then decreasing nature of the grade level trends could be a direct result of a lack of explicit instruction on QR in context in middle and high schools.

Trends across scales and QR processes by science strand vary as well. Biodiversity QR shows variation at the macro-level with QA rated higher than QI, but the processes are evenly ranked at the landscape and micro/atomic scales. Carbon QR shows the most variation at the micro level with QA rated higher than QI, and overall QR rated higher at the landscape scale. Water QR shows variation at the landscape scale with QI rated higher than QA, but the processes are evenly ranked at the macro- and micro-level.

Table 3. ANOVA on science scale versus QR process

Source of variance	SS	df	MS	F	p
QR processes	0.003	2.000	0.001	0.180	0.842
Science scales	0.030	2.000	0.015	1.964	0.255
Error	0.030	4.000	0.008		
Total	0.063	8.000			

Alpha level 0.05.

Tracking Trajectories

We now turn to the qualitative analysis of a track of students for the water strand. We focus on the water strand here, due to limited space the biodiversity and carbon strands will be reported in other papers. The themes that arose from the qualitative analysis of the QR water interviews can be grouped into the QA and literacy, interpretation of models, and QR in the context of environmental science. We do not report findings on modeling due to the minimal responses provided by students on the assessment.

Quantification Act

QR requires as an initial step that students quantify objects within a context, resulting in variables on which they can then reason in context. The QR water assessment provided descriptive information on how students initially approach QR within environmental science contexts through a variety of representations: tables, graphs, equations, and science models. Three different 'acts' were observed by students during the assessment: avoidance or engagement of quantitative information, the ability to identify variables, and the use of covariation. Avoidance or engagement with quantities meant they could choose to ignore the quantitative information and provide a strictly qualitative account or use the quantitative information to support their qualitative account. For example, when students were asked a question using a table and a pie graph on the macro-scale, only the 11th grader avoided the use of quantitative information to answer the questions until prompted by the interviewer, but the 12th grader avoided calculations with the quantitative data to support their qualitative arguments.

In another example of avoidance, a science box model showing the amount of fresh water in reservoirs on the landscape scale was compared to a traditional table using percentages and numbers to describe the amount of fresh water in reservoirs, all students but the seventh grader avoided using quantitative data without prompting and all avoided computation to support qualitative accounts.

It was also observed from the student assessments that scale along with pictorial representation had no impact on whether a student avoided the quantitative tools available to them. A pictorial science model was used to describe evaporation on the microscopic/atomic scale. Equations were also used in the problem to describe the process of evaporation in relation to kinetic energy of molecules. None of the students avoided using the graph, but the equations were avoided by the 6th, 9th, and 12th graders even when prompted to use them by the interviewer and only the 11th grader used the pictorial science model to support her account of kinetic energy and evaporation. Overall, when data were provided in tables or graphs, students were more open to using the quantitative information to make an argument. Equations were the last representation selected by the students and quantities embedded in science models were ignored until the students were prompted to use it. The students avoided computation, choosing instead to provide often vague qualitative accounts of the computational process they would use.

The QA was also observed to include identification of variables, their attributes, and measure, which is essential in interpreting models. When presented with a model, the students were first asked what they thought the model was telling them, which provided an opportunity for the student to identify variables, discuss attributes, and relate them to measures. For example, on the macro-scale, students were presented with a pie chart describing the proportion of different surfaces (e.g. concrete, roof, grass, etc.) of a generic school yard and a traditional table describing the amount of rain that falls on the school yard during a given year. They were asked to identify variables, assign the variables attributes, and relate measures to the variables. All of the students were able to correctly identify variables, assign attributes, and relate measures, except the seventh grader who on the pie chart did not use the percent quantitative measures.

When more complex water cycle tables with overlapping categories (e.g. total percentage of fresh water and percentage of fresh water stored in lakes) and part/whole percentages were presented to students, it became more problematic for them to correctly identify variables. The three middle school students were confused when asked to focus on a single case in the table, only the eighth grader was able to provide attributes for the variable and associate it with measures in the tables. The high school students interpreted the tables more holistically, identifying multiple variables correctly, but two of the three erred in assigning attributes to the variable and in relating measures to the variables. Only the 11th grader successfully identified variables in the more complex table, was able to assign attributes to the variables, and sort out the partial and whole percentages as measures of the variables.

Students were also provided a landscape scale water cycle box model with embedded quantitative data describing the amount of water stored in specific reservoirs. The box model was rendered using a diagram of the world. While students identified some of the variables in the water cycle box model, only after prompting did they attribute quantitative values of flow of water between reservoirs to the variables. The only exception was a seventh grader, who related the quantitative measures in the box model to the variables.

In another example, a traditional two-dimensional graph of vapor pressure (Torr) versus temperature (°C) was used to describe the process of vaporization at the microscopic level. Variables were identified correctly by all of the students without prompting, however, only the ninth grader correctly quantified the two axes of the graph. The other students were able to correctly identify temperature on the horizontal axis, but ignored or misinterpreted the vertical axis even while discussing correlations and trends of the two variables. The students correctly identified an increasing trend in the graph, but when attempting to relate the x-axis variable of temperature to the y-axis, none of them understood that axis represented Torr—a measure of vapor pressure. Only the ninth grader asked for assistance to understand Torr, whereas the interviewer had to prompt the other students with information on Torr as a measure of vapor pressure. They were all then able to discuss the positive correlation underlying the covariation of temperature and vapor pressure, but only the 11th and 12th graders attempted to relate this to the variable of vaporization. While temperature and vapor pressure were explicit variables represented by the axes of the graph, the variable of vaporization was embedded in the model—the curve itself represented the temperature and vapor pressure at which water would vaporize. This embedding of an additional variable in the graph was difficult for the students in the lower grade levels and distracted many of them from relating the measure of temperature or vapor pressure to the associated variables. In the case of vaporization, the students were asked to identify any variables that impacted evaporation before being shown the graphical model. This was to assess their QA ability. While there are multiple variables influencing evaporation, five of the six students named only one variable (four heat, one surface area) and only the 12th grader named two variables. For the variables identified, the students did not relate how attributes of the variable impacted evaporation or what measures were used.

The idea of covariation is common in science and understanding covariance is an essential part of quantification, a necessary aspect of explicating trends and relationships in models. As mentioned above, in the graphical model of vaporization, all students had a fundamental notion of covariance between temperature and some variables represented by Torr on the *y*-axis, indicating an increasing trend relationship. However, as was discussed above, the embedded variable of evaporation was confused with the *y*-axis pressure variable by all of the students, so covariance between temperature and vapor pressure impacting evaporation was not comprehended.

In another example, when students were presented with equation models describing the relationship between vaporization, temperature, and kinetic energy of molecules at the microscopic/atomic scale, only the 6th and 11th graders addressed the issue of covariance. The sixth grader interpreted a covariate relationship between energy and the velocity of water molecules, but erroneously included a related change in mass. The 11th grader related the equation for vaporization to the graph for this relationship and discussed covariance through the graph, avoiding the equation completely.

Quantitative Literacy

Quantitative literacy includes the arithmetic skills to manipulate, compare, and reason with the variables that result from the QA to address the question of interest. The most prevalent QL skills used by the students were proportional reasoning, numeracy, and measurement. Proportional reasoning is broadly defined here to include using ratios, fractions, rates, percents, and proportions. For the pie chart question on the macro-scale, students at all grade levels were able to interpret percent, relate percent to area of school yard, and perform calculations with percent. However, the 6th–9th graders struggled with interpreting percentages of fresh water sources versus percentages of overall fresh water in the reservoir table at the landscape scale, while the 11th and 12th graders reasoned with both part and whole percentages to discuss concerns about the amount of fresh water on Earth. Additionally, at the landscape scale, only the 6th and 12th graders interpreted the proportional relationship of arrow size to water amount exchanged. When examining the effect of temperature on kinetic energy of molecules in water, an equation was provided which showed that the kinetic energy of a molecule is proportional to temperature. The sixth grader avoided discussion of what proportional meant, whereas all other students interpreted proportional as a change in the same direction (i.e. if kinetic energy increases, so does velocity or even mass), but none discussed a fixed rate of change as an element of being proportional.

Numeracy in the assessments appeared in the form of numeric calculations or comparing magnitudes. At the macro-scale for the pie chart students engaged in calculation with prompting, all but the 7th, 8th, and 11th graders made errors in their calculations. At the landscape scale with both science and table models, all but the 12th grader avoided calculation, preferring qualitative accounts. The data did not indicate that calculation improved across grade levels, but this may be due as much to us encouraging students to verbalize but not work out the details of their processes as it is to student ability.

Measure was assessed through responses on amount of rain, using appropriate measurement units, conversion factors, and using referents (such as how many bathtubs of water fall on school yard). At the macro-scale, the 6th, 11th, and 12th graders misinterpreted the measure of amount of rain as an area, ignoring accumulated millimeters of rain and focusing on surface area. The seventh, eighth, and ninth graders believed the correct measure was volume, but did not provide a quantitative or qualitative account supporting their conviction. Both metric and English measures were included in the rain fall table, with the 7th, 9th, 11th, and 12th graders showing a preference for inches versus millimeters and the other students not explicitly discussing the unit. A conversion factor from cubic feet to gallons was provided for the calculation of amount of rain on school yard. The conversion factor was avoided by the eighth grader, correctly interpreted but not used in calculation by the sixth and seventh graders, and was misinterpreted by the high school students. Finally, the students were asked to provide a referent for the total amount of rain falling on the school yard, such as how many swimming pools of water would be filled by the total amount of rain water? Only the ninth grader provided a referent without prompting, comparing the amount of rain to a full water tower.

There was an element of measure in the comparison of the water cycle model and the flow table. Students either on their own or with prompting noticed a difference in the amount of precipitation falling in the ocean between the two models. When asked to explain this, none of the students assessed provided an argument based on estimation error or precision of measure. They all provided arguments that did not question the authority or accuracy of either model, including the models representing different geographic places (6th, 7th, 9th graders), temporal or time change (8th, 9th, 12th graders), and a change in processes between the models (11th grader). In addition, the sixth, seventh, and eighth graders referred to volume of water as having a linear unit of measure.

Quantitative Interpretation

The QR water assessment explored 6th–12th grade students' ability to interpret environmental science models represented as tables, graphs, analytic equations, and science models. At the lowest level, QI is the ability to explain a selected case in context, with other evidence of QI being the ability to determine trends in a model, make predictions of future events, and translate between different models of the same phenomena. We also examined student preference for a given representation of a model.

Tabular Model

The most basic representation of data is a table. If students use a table to determine trends or make predictions, then we view the table as more than just a representation of the data, but as a model of the phenomena on which data were collected. An average rainfall table with a focus on identifying when maximum rainfall levels occurred was used as a representation at the macro-scale. All of the students except the eighth grader immediately interpreted the average rainfall table. However, only the 12th grader related the table to overall rainfall and attempted to use it to determine the total yearly volume of rain falling on the school yard. However, even she failed to provide a correct quantitative account for volume of rain. Students were also asked to identify variables in a table showing percentages of fresh water in multiple reservoirs at the landscape scale. Both partial and total percents were used to describe the storage of fresh water. Middle grade students were able to identify some of the variables representing fresh water sources, but only the seventh grader was not confused by partial percents and total percents in the table, and therefore was able to correctly interpret the total percentage of fresh water. However, the student failed to provide a meaningful quantitative account of why a small amount of fresh water is a societal concern. Students were then asked to compare the traditional table to a pictorial science box model that described the storage of water in reservoirs. The middle grade students were confused when asked to compare the table with a box model and had to be prompted to begin to compare the data sets. Students were asked to select one specific case from the flow table to compare with the box model. In contrast, the high school students correctly identified the percent of fresh water in the traditional table, and made comparisons between the table and box model without prompting. While the high school students provided quantitative descriptions of trends in the table and box model, only the 11th grader was able to provide a quantitative argument for why the percent of fresh water might be a concern.

Graphical Model

Students in the 6th–12th grades should be familiar with a graphical representation of data in a table and have skill in creating a graph from a table. However, science graphs often embed more than two variables in a graph and as discussed in the Quantification

Act section, the students assessed in this study appeared to have difficulty understanding this process. For example, when students were asked to interpret a pie chart that describes the proportion of different surfaces in a school yard at the macro-scale, all of the students correctly identified that grass had the maximum surface area for the problem. Students were then asked what would happen to the areas of the pie chart if a gym was to be built next to the school. All students were able to describe a modified pie chart which accounted for dynamically changing the pie chart regions to represent changes in surface areas, though the seventh and eighth graders did make calculation and interpretation errors.

In the vaporization model question, where students were presented with a traditional two-dimensional graph of vapor pressure (Torr) versus temperature, students struggled with interpreting the axes of the vaporization graph (see Act of Quantification for details). However, even when they were not sure what the y-axis variable represented, they correctly interpreted the increasing trend in the graph. The trend was explained incorrectly in terms of the science where the trend was described either in terms of temperature only or as temperature versus vapor pressure without reference to vaporization. All students assessed with the exception of the 12th grader were able to interpret a point on the graph in terms of specific numerical values for temperature and pressure, but not one of the students was able to provide an interpretation of the graph with respect to all three variables of temperature, pressure, and vaporization.

Additionally, students were asked to use the graph to make a prediction about the temperature at which water boils if the vapor pressure was 1600 Torr, which happened to be a value that was just outside the range of the graph provided. The students were provided the graphic model and also an analytic equation model which could be used to correctly determine the pressure. Only the 8th and 11th graders attempted to predict a value of temperature for a given pressure by extrapolating from the trend in the graph, neither providing a detailed quantitative account of the method used. When asked to interpret the vaporization equation, the sixth and ninth graders did not make an attempt and the 12th grader misinterpreted the variables in the equation. The 7th, 8th, and 11th graders were able to identify the variables in the vaporization equation and relate them to the vaporization graph.

Analytic Model

The analytic equation models described in the previous sections for vaporization and kinetic energy on the micro/atomic scale were not well received or interpreted by the students. Each analytic model was related to another representation of the phenomena: vaporization to a graph of temperature versus vapor pressure at which water boils and kinetic energy to a pictorial science model. As discussed above, some students could identify variables correctly, but it was mostly misinterpreted or ignored. Another instance where this was observed was in the model for evaporation, where the kinetic energy equation was used to describe the relationship between the temperature and evaporation of water molecules. Only three of the students attempted to interpret the equation. The 6th, 7th, and 11th graders discussed the covariation between kinetic energy and velocity of molecules, but all erroneously included a change in mass. However, the 11th grader was more sophisticated in her qualitative explanation, discussing a change in phase as spreading out the mass. While this is an isolated case in this study sample, this evidence can be used to inform the learning progression by serving as an exemplar. None of the students assessed discussed the proportional constant in the kinetic equation. Overall, the interpretation of analytic equations was the weakest area for students in the study.

Science Model

The students were provided two science models as well: on the landscape scale, a box model of the water cycle and on the microscopic scale, a pictorial systems model of kinetic energy. The landscape water cycle model had embedded within it specific quantitative data for volume of reservoirs and volume of water transported between reservoirs. When asked to interpret what the model told them about the water cycle, all students in the track except the seventh grader completely ignored the quantitative data, choosing to give a qualitative account of flow between reservoirs. The seventh grader made a quantitative reference to the amount of water being transported between reservoirs. Upon prompting to compare the processes that return water to the atmosphere with precipitation all of the students used the quantitative data in the box model. The 6th, 7th, 9th, and 12th graders identified quantities returning water to the atmosphere (evaporation and evapotranspiration) discussing values individually, but not providing a quantitative account supporting the question of balance between evaporation and precipitation. The 8th and 11th graders concluded that the flow between the surface and atmosphere was 'about' equal, but avoided calculation to verify the relationship. When asked about the impact of deforestation on the water cycle, only the eighth grader provided a response indicating she viewed the water cycle as closed, but she did not provide a quantitative account using data in the model to support her conclusion. The kinetic energy science model is a pictorial representation of increased kinetic energy and its impact on evaporation. When provided the model and told that kinetic energy is proportional to temperature, the 6th, 8th, 9th, and 12th graders related an increase in energy to an increase in evaporation. It was difficult to discern the level of use of the pictorial model versus the use of the proportional relationship given.

Translation Between Models

Students were given a choice of model representations on the landscape scale (science model versus table), the macro-scale (table versus graph), and microscopic scale (graph versus equation; science model versus equation). When a preference for one representation type over another was observed, we found the following: equations were the least preferred representation with five of six students choosing a graph over an equation and two of three choosing a science model over an equation. In fact, only the seventh grader selected an equation over any other representation. When offered a table and a graph, the students were equally split on which they preferred.

QR Across Science Scales

The assessment included questions at the landscape scale, macro-scale, and micro/atomic scale to determine how students responded quantitatively at each scale. The primary conclusion is that students widely preferred qualitative accounts, providing them without prompting, to quantitative accounts which often required prompting, regardless of scale. There were some exceptions, including a more balanced qualitative and quantitative bent for the 7th grader across all three scales, the 8th grader on the macro-scale, and the 12th grader at the macro-scale. While the data are very preliminary, it does provide a couple of trends to be further assessed. First, students may be scale independent, that is, they will prefer qualitative accounts (6th and 11th graders) or quantitative accounts (7th grader) regardless of the scale. For others (12th grader), the scale may influence their choice of qualitative versus quantitative account.

Discussion

The quantitative and qualitative data indicate a number of trends in the QR development for middle and high school students within an environmental context. The first were trajectory issues. We hypothesized that students would increase in their understanding and the use of QR across grade levels in rank on the learning progression, specifically at the macro-scale. In fact, this trend was not observed; there was no consistent increase in learning progression levels across grade levels.

Second were scaling issues. We hypothesized that students would be more proficient in QR at the macro-scale where they can draw on personal experience. There were no consistent differences in QR use on the micro/atomic, macro-and landscape-scale. Personal experience led to qualitative accounts, with quantitative accounts occurring primarily through prompting by the interviewer. We also hypothesized that moving to the landscape scale would engage students in the inherently quantitative task of generalizing or seeking trends and that moving to the microscopic-atomic scale would engage students in QR accounts about hidden mechanisms explained with quantitative sciences such as physics and chemistry. Yet on both the landscape and micro-atomic scales, the students' preference appeared to be almost entirely qualitative. Therefore, does the students lived experience or lack of experience at a scale have the opposite impact we hypothesized? Does their lived experience at the macro-scale provide a level of comfort that supports taking the quantitative point of view? Does their lack of experience at the micro/atomic scale inhibit their willingness to incorporate quantitative accounts, even though the accounts at this scale are inherently quantitative in nature?

Third, there were tool implementation issues. Students often failed to select the appropriate mathematical or statistical tool from their toolbox, and even when the correct tool was selected they failed to use QR to apply the tool within the science context. QR requires knowledge in the science context as well as in mathematics/statistics, but even if a student possesses both, it does not guarantee that QR will occur. It was evident that a lack of understanding of the context results in meaningless manipulation of quantitative information.

The primary focus of this research is on the development of a QR learning progression (Appendix 1). However, there are implications for teaching. Students need to engage in real-world problem-based learning to become environmentally literate. QR can serve as a major support or barrier to this development. To enhance QR, teachers must require students to provide quantitative as well as qualitative support for their arguments. Students should be provided multiple quantitative representations (tables, graphs, equations, science models) within a science context and use QR to provide data-based informed decisions about critical issues that impact their place. Students should engage in building their own QM representing these issues, then test and refine those models. The QR learning progression provides a view of how QR develops in the context of environmental science. The learning progression can be used to determine QR strengths and areas in need of development. Developing QR will require a change in both content and pedagogical practice; content needs to be within a real-world context and QR is inherently interdisciplinary requiring a change in teacher practice.

The QR learning progression framework (Appendix 1) is based on findings from this study. The first column focuses on the QA progress variable, which is the trigger for QR. A student first quantifies an object within a context, allowing them to operate on that quantity using the arithmetic processes within quantitative literacy. The second column is QI, the ability to interpret a model provided to the student. Column three provides a framework for a modeling progression which adds a quantitative focus to that of the MoDeLs science modeling progression.

A conundrum which we encountered in the development of a QR learning progression framework is separating QR progression from the mathematics and statistics that underlie the ability to engage in QR. We began by creating progression frameworks addressing QR components, including quantitative literacy frameworks for numeracy, proportional reasoning, change, and measure. We also created frameworks for the multiple representation aspects of QI and an overall framework for QM. The problem with this approach is twofold. First, there is the complication of how so many different frameworks with a focus on mathematics and statistics can be integrated into the project's science learning progressions for the Carbon, Water, and Biodiversity Strands in the Pathways Project. Attempts to cross tabulate the science progressions with the QR frameworks met with little success. In addition, separate progressions for mathematical components of QR are in conflict with the interdisciplinary component of our definition of QR. Second, developing frameworks which focus on mathematical and statistical understandings does not reflect the key aspect of QR as a habit of mind; as the act of using mathematics and statistics within a context. While QR requires the use of mathematics and statistics, it is not the same as mathematics and statistics. QR is the ability to 'see' the mathematics within a context, to choose the appropriate mathematical or statistical tool from a toolbox and apply it within the context, and the ability to move from the mathematical and statistical analysis back to the context to make a decision. Perhaps, this view of QR can be captured more in a meta-level quantitative framework like the one in Appendix 1, rather than a series of detailed frameworks focused on mathematical and statistical tools in the toolbox.

The focus of this study was on the impact that QR may have on understanding environmental science issues. This should not be construed as implying that QR is required to understand all environmental science issues. Certainly, qualitative accounts are important in arguing such issues and often serve as the initial state of developing arguments. However, QR is essential if students are to provide data-based arguments to support their qualitative accounts. The role of QR in environmental science is raised to a higher level as a part of a model-based approach to teaching science as proposed by Taking Science to School (Duschl et al., 2007) and the Framework for K-12 Science Education (NRC, 2011).

Implication for the International Science Education Community

Creating a global community of scientifically literate citizens is imperative for the next generation. Subsequent generations will face environmental issues concerning energy (i.e. depletion of fossil fuels, harnessing alternative energy sources) and environment (i.e. understanding the management of natural resources and climate change) and without the ability to quantitatively reason within these contexts, the decision-making process will potentially be hindered. QR allows an individual to conceptualize, understand, and apply information presented in order to make an informed decision. As science educators, not only should we be researching students learning progressions of QR, but also how QR is taught (or not) in the K-12 classroom. Understanding these components may allow us the opportunity to positively impact students' QR abilities.

Note

1. This project is supported in part by a grant from the NSF: Culturally Relevant Ecology, Learning Progressions, and Environmental Literacy (DUE-0832173), which we refer to as Pathways.

References

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, *33*, 352–378.

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). *Learning progressions in science: An evidence-based aproach to reform*. Philadelphia, PA: Consortium Policy Research in Education.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.

Lehrer, R., & Schauble, L. (2002). *Investigating real data in the classroom: Expanding children's understanding of math and science*. New York: Teachers College Press.

Lesh, R., Middleton, J., Caylor, E., & Gupta, S. (2008). A science need: Designing tasks to engage students in modeling complex data. *Educational Studies in Mathematics*, *68*, 113–130.

Lobato, J., & Siebert, D. (2002). Quantitative reasoning in a reconceived view of transfer. *Journal of Mathematical Behavior*, *21*, 87–116.

Louca, L. T., Zacharia, Z. C., & Constantinou, C. P. (2011). In quest of productive modeling-based learning discourse in elementary school science. *Journal of Research in Science Teaching*, *48*(8), 919–951.

Madison, B. L., & Steen, L. A. (Eds.). (2003). *Quantitative literacy: Why numeracy matters for schools and colleges*. Princeton, NJ: Woodrow Wilson Natl Foundation.

Mayes, R., Peterson, F., & Bonilla, R. (2012). Quantitative reasoning: In context. In R. Mayes & L. Hatfield (Eds.), *WISDOMe: Quantitative reasoning and mathematical modeling: A driver for STEM integrated education and teaching in context* (pp. 7–38). University of Wyoming.

Moore, K. C., Carlson, M. P., & Oehrtman, M. (2009). *The role of quantitative reasoning in solving applied precalculus problems*. Conference on research in undergraduate mathematics education (CRUME). Raleigh, NC.

National Research Council (NRC). (2001). *Grand challenges in environmental sciences*. Washington, DC: National Academy Press.

National Research Council (NRC). (2011). *A framework for K-12 science education: Practices, cross-cutting concepts, and core ideas*. Washington, DC: National Academy Press.

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, *48*, 486–511.

Sadler, T. D., & Zeidler, D. L. (2009). Scientific literacy, PISA, and socioscientific discourse: Assessment for progressive aims of science education. *Journal of Research in Science Teaching*, *46*(8), 909–921.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Archer, A., Fortus, D., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*, 632–654.

Shavelson, R. (2008). Reflections on quantitative reasoning: An assessment perspective. In B. L. Madison & L. A. Steen (Eds.), *Calculation vs. context: Quantitative literacy and its implications for teacher education* (pp. 22–47). Washington, DC: Mathematical Association of America.

Smith, C. L., Wisner, M., Anderson, C. W., & Krajcik, J. (2006). Implication of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic molecular theory. *Measurement: Interdisciplinary Research and Perspective*, *4*, 1–98.

Steen, L. A. (Ed.). (2001). *Mathematics and democracy: The case for quantitative literacy*. Princeton, NJ: The Woodrow Wilson National Fellowship Foundation, The National Council on Education and the Disciplines.

Stefani, C., & Tsaparlis, G. (2009). Students' levels of explanations, models, and misconceptions in basic quantum chemistry: A phenomenographic study. *Journal of Research in Science Teaching*, *46*, 520–536.

Taylor, A. R., & Jones, M. G. (2009). Proportional reasoning ability and concepts of scale: Surface area to volume relationships in science. *International Journal of Science Education*, *31*, 1231–1247.

Thompson, P. W. (1994). The development of the concept of speed and its relationship to the concepts of rate. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 181–234). Albany, NY: SUNY Press.

Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlain, & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education* (pp. 33–57). Laramie, WY: University of Wyoming.

Tretter, T. R., Jones, M. G., Andre, T., Negishi, A., & Minogue, J. (2006). Conceptual boundaries and distances: Students' and experts' concepts of the scale of scientific phenomena. *Journal of Research in Science Teaching*, *43*(3), 282–319.

Tretter, T. R., Jones, M. G., & Minogue, J. (2006). Accuracy of scale conceptions in science: Mental maneuverings across many orders of spatial magnitude. *Journal of Research in Science Teaching*, *43*(10), 1061–1085.

Appendix 1. QR learning progression framework

Achievement level	QR progress variable		
	QA	QI	QM
Level 4 (upper anchor)	4a *Variation*: reasons about covariation of two or more variables; comparing, contrasting, relating variables in the context of problem 4b *Quantitative literacy*: reasons with quantities to explain relationships between variables; proportional reasoning, numerical reasoning; extend to algebraic and higher math reasoning (MAA) 4c *Context*: situative view of QR within a community of practice (Shavelson, 2008); solves ill-defined problems in socio-political contexts using ad-hoc methods; informal reasoning within science context (Madison & Steen, 2003; Sadler & Zeidler, 2009) 4d *Communication*: capacity to communicate quantitative account of solution, decision, course of action within context	4a *Trends*: recognizes and provides quantitative explanations of trends in model representation within context of problem, including linear, power, exponential trends 4b *Predictions*: makes predictions using model with covariation and provides a quantitative account which is applied within context of problem 4c *Translation*: translates between different models, at least categorically (i.e. this graph looks exponential) 4d *Revision*: revise models theoretically without data, evaluate competing models for possible combination (Schwarz et al., 2009) 4e *Authority*: question model by challenging quantitative aspects as estimates or due to measurement error, especially when contrasting models	4a *Create model*: ability to create a model representing a context and trace through model correctly 4b *Refine model*: test and refine a model for internal consistency and coherence to evaluate scientific evidence and explanations; results; extend model to new situation (Duschl et al., 2007) 4c *Model reasoning*: construct and use models spontaneously to assist own thinking, predict behavior in real-world, generate new questions about phenomena (Schwarz et al., 2009) 4d *Methods*: demonstrate ability to use variety of methods to construct model within context; least squares, linearization, normal distribution, logarithmic, logistic growth, multivariate, simulation models 4e *Statistical*: conduct statistical inference to test hypothesis (Duschl et al., 2007)

(*Continued*)

Appendix 1. Continued

Achievement level	QR progress variable		
	QA	QI	QM
Level 3	3a *Variation*: recognizes correlation between two variables but provides a qualitative or isolated case account; lacks covariation 3b *Quantitative literacy*: manipulates quantities to discover relationships; measure, numeracy, proportional, statistical procedures 3c *Context*: display confidence with and cultural appreciation of mathematics within context; number sense, practical computation skills (Steen, 2001) 3d *Communication*: capacity to communicate qualitative account of solution, decision, course of action within context; weak quantitative account 3e *Variable*: mental construct for object within context is identified, conceptualized so that the object has attributes that are measurable (Thompson, 2011); uses variable in context	3a *Trends*: expand recognition of patterns in models of one variable to recognizing linear versus curvilinear growth 3b *Predictions*: interprets models where one variable is categorical, identifying trends and making predictions with strong quantitative accounts; make predictions using model with covariation but only provide qualitative account 3c *Translation*: attempts to translate between models if prompted but fails to relate variables between models 3d *Revision*: revise model to better fit evidence and improve explanatory power (Schwarz et al., 2009) 3e *Authority*: question differences between models, but use erroneous qualitative accounts not error or approximations	3a *Create model*: create simplistic models for covariation situations that lack quantitative accounts; fail to trace model correctly 3b *Refine model*: test and refine model based on supposition about data; extend model without verifying fit to new situation 3c *Model reasoning*: construct and use multiple models to explain phenomena, view models as tools supporting thinking, consider alternatives in constructing models (Schwarz et al., 2009) 3d *Methods*: demonstrate ability to use two different methods to model a situation 3e *Statistical*: use descriptive statistics for central tendency and variation; make informal comparisons to address hypothesis

Level 2	2a *Variation*: sees causation in relationship between two variables, provides only a qualitative account; lacks correlation 2b *Quantitative literacy*: poor QL interferes with manipulation of variables; struggle to compare or operate with variables; ability to manipulate and calculate with one variable to answer questions of change, discover patterns, and draw conclusions; 2c *Context*: lack confidence with or cultural appreciation of math within context; practical computation not related to context 2d *Communication*: provides elements of account, but lacks capacity to communicate solution, decision, course of action within context; weak qualitative account 2e *Variable*: object within context is identified, but not fully conceptualized with attributes that are measurable; object is named creating a variable (Thompson, 2011)	2a *Trends*: identify and explain single case (point) in model within context; recognize increasing/decreasing trends but not relating to change in both variables (covariation lacking) 2b *Predictions*: makes predictions for models with one variable but provides only qualitative arguments 2c *Translation*: indicate preference for one model over another but do not translate between models 2d *Revision*: revise model based on authority rather than evidence, modify to improve clarity not explanatory power (Schwarz et al., 2009) 2e *Authority*: acknowledge quantitative differences in models but does not provide an explanation 2f *Interpret*: identify variables in the model (i.e. graph axes, table headings, equation unknowns); provide qualitative account, avoiding quantities; form correct mental image to conceive problem; difficulty with models that embed variable or have more than two interrelated variables	2a *Create model*: creates visual models to represent single variable data, such as statistical displays (pie charts, histograms) 2b *Refine model*: extends a given model to account for dynamic change but provides only a qualitative account 2c *Model reasoning*: construct and use model to explain phenomena, means of communication rather than support for own thinking (Schwarz et al., 2009) 2d *Methods*: constructs a table or data plot to organization information but does not use as model 2e *Statistical*: calculates descriptive statistics for central tendency and variation but does not use to make informal comparisons to address hypothesis

(*Continued*)

Appendix 1. Continued

Achievement level	QR progress variable		
	QA	QI	QM
Level 1 (lower anchor)	1a *Variation*: does not compare variables 1b *Quantitative literacy*: struggles to manipulate and calculate with even one variable to answer questions of change, discover patterns, and draw conclusions 1c *Context*: does not relate quantities to context 1d *Communication*: discourse is force-dynamic; avoids quantitative account, ignoring quantities providing weak qualitative account 1e *Variable*: objects within context are not identified, no attempt to conceptualize attributes that are measurable	1a *Trends*: do not identify trends 1b *Predictions*: avoid making predictions 1c *Translation*: fail to acknowledge quantitative difference in models 1d *Revision*: view models as fixed, test to see if good or bad replicas of phenomena (Schwarz et al., 2009) 1e *Authority*: does not acknowledge difference in models 1f *Interpret*: fail to relate model to context; avoid using model	1a *Create model*: does not view science as model-building and refining so does not attempt to construct models; forced dynamic or low level school science discourse, expect to receive facts and memorize processes 1b *Refine model*: no model created to refine 1c *Model Reasoning*: construct and use models that are literal illustrations, model demonstrates for others not tool to generate new knowledge (Schwarz et al., 2009) 1d *Methods*: no evidence of knowledge of methods for building models 1e *Statistical*: does not use statistics; no calculation of even descriptive statistics