International Journal of Science Education
Vol. 30, No. 13, 17 October 2008, pp. 1725–1751

Routledge
Taylor & Francis Group

RESEARCH REPORT

Use of First-hand and Second-hand Data in Science: Does data type influence classroom conversations?

Barbara Hug[a]* and Katherine L. McNeill[b]
[a]*University of Illinois Urbana-Champaign, USA;* [b]*Boston College, USA*

In this paper, we examine how students discuss and interpret data and whether these actions vary depending on the type of data they analyse. More specifically, we are interested in whether students perform differently when analysing first-hand data, which they collect themselves, compared with second-hand data provided to them. Our data analysis focused on two classrooms using two different curriculum units, chemistry in Grade 7 and biology in Grade 8, collected during the 2002/03 school year from a Mid-western urban middle school in the USA. We analysed classroom videotape associated with lessons in which students discussed first-hand and second-hand data both in small group settings and full class discussions. We found the two types of data offer different benefits and limitations, suggesting that both types of data are important for students to work with as they develop skills in scientific inquiry practices. We discuss the characteristics of classroom discussions around different data sources as well as implications for the design of curriculum materials, instructional environments, and student learning in science.

Introduction

Students develop scientific knowledge by actively engaging in scientific ways of knowing, such as asking questions, designing experiments, and analysing data (Driver, Asoko, Leach, Mortimer & Scott, 1994). Research literature (Krajcik, Blumenfeld, Marx, & Soloway, 2000; Metz, 2000; White & Frederiksen, 1998) and national standards (American Association for the Advancement of Science, 1993; National Research Council, 1996) call for students to engage in a range of scientific inquiry practices. One critical scientific practice is data analysis, an inquiry practice key to making meaning out of the investigations in which students engage within the

*Corresponding author. Department of Curriculum and Instruction, College of Education, University of Illinois Urbana-Champaign, 1310 South Sixth Street, Champaign, IL 61820, USA. Email: bhug@uiuc.edu

ISSN 0950-0693 (print)/ISSN 1464-5289 (online)/08/131725–27

DOI: 10.1080/09500690701506945

science classroom and essential for the issues that face them in their everyday lives (Duschl, 1990; Germann & Aram, 1996; Lehrer & Schauble, 2002; Wu & Krajcik, 2006).

When analysing data, students can analyse first-hand data that they have collected, second-hand data that were gathered first-hand by other individual(s), or a range of secondary data sources. The *National Science Education Standards* acknowledge the role of these different data sources in inquiry: 'some activities provide a basis for observation, data collection, reflection and analysis of firsthand events and phenomena. Other activities encourage the critical analysis of secondary sources—including media, books and journals in a library' (National Research Council, 1996, p. 33). Our study investigates potential benefits and limitations of having middle school students engage in analysing these different types of data.

Defining First-hand versus Second-hand Data Experiences

We are interested in examining how students discuss and interpret data, and whether these actions vary depending on the type of data they are using. More specifically, we are interested in whether students respond differently to the contrasting opportunities presented by using first-hand or second-hand data. Similar to Palincsar, Magnusson, and colleagues (Magnusson, Palincsar, Hapgood, & Lomangino, 2004; Palincsar & Magnusson, 2001), we define first-hand data experiences as those that occur when students investigate phenomena using various hands-on inquiry practices. Consequently, first-hand data analysis results when students analyse data that they collect through their own investigation.

Magnusson et al. (2004) define second-hand data experiences as when students 'learn about and evaluate other's investigations of the same or similar phenomena' (p. 318). Specifically, they study how elementary students use notebook entries documenting scientific investigations created by a fictitious scientist named Leslie Park. We use a slightly broader definition of second-hand data experiences, as students in K–12 science classrooms often cannot collect data for a variety of reasons. Some phenomena are too dangerous (e.g., explosive chemical reactions), too slow (e.g., natural selection), or too expensive (e.g., DNA sequencing), for classroom investigation. Thus, our definition of second-hand data experiences includes having students evaluate not only others' actual investigations, but also evaluating data collected and possibly reported by other individuals. We expanded our view because analysis of second-hand data experiences occurs every day when individuals interpret data presented in the newspaper, in magazines, or on television. Furthermore, the analysis of second-hand data has recently become increasingly important in science at all levels, as large data-sets (e.g., genomic databases and weather databases) are readily accessible to a range of people from practicing scientists to K–12 students. This access to a range of data types allows a deeper scientific understanding of the phenomena under investigation to be developed (Duschl, 1990).

Although we define first-hand and second-hand data experiences in this manner, data type can be seen as a continuum between first-hand and second-hand with no clear boundary between the two. For example, if a class collects and uses a set of data for a water-quality unit, this would be classified as first-hand data. But if a different class later in the day or in a subsequent year used these data, is it first-hand or second-hand data for these new students? We would argue they are now using second-hand data, but data that have connections and personal relevance to the students, making it potentially different in character from a second-hand water-quality data-set about a river in a neighbouring state or region.

An example of first-hand data that might require clarification is data collected by different groups of students in a single class, pooled and used collectively to answer a research question. We consider such a data set to be first-hand data even though not all of the students collected all of the data. In this study, we identify examples of first-hand and second-hand data as a starting point for this type of analysis. We believe that future research can build on this study to develop our understanding of the many possible ways that students generate, use, and derive meaning from different types of data.

Benefits and Limitations of First-hand versus Second-hand Data

Students may develop scientific reasoning skills differently when they use first-hand data compared with second-hand data. Kanari and Millar (2004) argue students reason differently when they use data they collect instead of data collected by another person. They showed that having students understand the role of data uncertainty in data collection is critical for the scientific reasoning that follows. However, overall there appears to be little empirical work comparing the benefits and limitations of having students engage in using these two different types of data. In this next section we discuss the research that examines the use of first-hand and second-hand data separately, which informed the codes we developed to analyse classroom talk around these two different types of data.

First-hand data. In terms of first-hand data experiences, there have been a number of studies that have looked at students' collection and use of data (Metz, 2000; Petrosino, Lehrer, & Schauble, 2003). These studies and others have provided varying results about students' successes and challenges in analysing data and supporting their conclusions with evidence. For example, Kanari and Millar (2004) found that when students collect their own data, they analyse data more easily when two variables show clear co-variation. In terms of benefits from analysing their own data, Hug and Krajcik (2002) found that students better appreciate the need to carefully design investigations (e.g., importance of using multiple trials), and they express ownership in the data if allowed to collect and work with first-hand data that has personal relevance. In addition, studies examining students' use of micro-based laboratories have illustrated how manipulating first-hand data in real time can positively influence the understanding students develop (Schecker, 1998).

In a Delphi questionnaire of science experts about important ideas in science to include in science classrooms, the analysis and interpretation of data emerged as one of nine themes (Osborne, Collins, Ratcliffe, Millar, & Duschl, 2003). A number of the experts in this study hypothesized that experiences with first-hand data were particularly important because students would be more interested in the data and it would provide an opportunity to discuss uncertainty in data and different interpretations of data. Although these hypotheses may be correct, there is little empirical evidence to support whether analysing first-hand data provides different opportunities for students when compared with second-hand data.

Second-hand data. Palincsar and Magnusson's (2001) work argues that experiences with second-hand data in textual form may offer students a number of unique benefits, including preparing students for first-hand data experiences, extending first-hand data experiences, and providing a common inquiry to promote students' conceptual understanding. Engaging students with second-hand data within texts can offer students a model of how to organize and represent data when they engage in similar first-hand experiences. This experience may better prepare students to collect their own data.

Experiences with text-based second-hand data can be used to extend experiences with first-hand phenomena. For example, students cannot conduct first-hand investigations of the sun or an atom in their classroom. Yet they can conduct other first-hand experiences with light or elements, which can then be extended through their experiences with second-hand data embedded within text about these concepts. In addition, second-hand data experiences can extend first-hand data experiences when there has not been consensus with student results from an experiment that they have conducted in class. The experience of using a common second-hand data set can provide a common inquiry or shared context that can further discussion and student understanding.

A goal of science education is to help students become competent members of a science and technology infused world where they are able to critically evaluate science (McGinn & Roth, 1999). Specifically, it is through the texts of science or media accounts that citizens frequently interact with and consider the implications of science (Osborne, 2002). This provides an argument for using second-hand data in their investigations, because students need to be able to evaluate and interpret scientific information in books, periodicals, and other resources (National Research Council, 1996). Second-hand data have also become more prevalent in science as scientists collect and share large data-sets, such as data on climate change and ecological diversity. Advances in technology enable the collection or creation of new data-sets that allow scientists to engage in their own second-hand data investigations to identify key patterns and relationships in the data, which allows for new or modified explanations of phenomena, models, or theories to be developed (Duschl, 1990). Educational researchers have investigated students' use of such large second-hand data-sets (Edelson, 2001; Reiser et al., 2001; Songer, 1996), and have found that students actively engage in the task they are asked to participate in and show gains in understanding both the science content and the scientific practices assessed. However, these studies identified challenging areas for students that require additional supports or scaffolds in order to successfully engage in using the data-sets. None of these studies explicitly focused on the issue of what students are able to do with first-hand versus second-hand data.

Student Difficulties with Analysing Data

When analysing data, regardless of whether they are first-hand or second-hand, individuals need to evaluate what data to use as evidence and then examine the evidence for patterns in order to draw conclusions and make decisions (Aikenhead, 2004). There are multiple steps in this process that are frequently difficult for students. In this next section, we provide a review of the literature in terms of common student difficulties. We used these common student difficulties to develop codes for the analysis of classroom conversation around both first-hand and second-hand data. We were interested in determining whether these difficulties arose more frequently with one type of data.

Data as evidence. In terms of determining what data can be used as evidence, individuals often do not engage in the types of reflection and evaluation necessary to make this decision (Sadler, 2004). When students are confronted with both appropriate and inappropriate data, they have difficulty selecting the appropriate data to construct their claims (McNeill & Krajcik, 2007). Students frequently use their own intuitive beliefs, values, or experiences as evidence instead of scientific data (Hogan & Maglienti, 2001). They appear to have difficulty understanding what counts as appropriate evidence to solve a scientific problem. Students also have difficulty understanding that some degree of variation will always be present in data (Germann & Aram, 1996). When pressed, students can come up with a variety of reasons for why differences between repeated observations may occur. For example, Masnick and Klahr (2003) discuss multiple types of error or limitations of data that students can recognize, including design error, measurement error, limitation error, and interpretation error—all of which can lead to uncertainty in data and impact the usefulness of the data as evidence. Kanari and Millar (2004) have shown students have difficulty in understanding the role of experimental data and measurement in investigations when the variables investigated did not co-vary. Their study illustrates the importance of the context within which students collect and analyse their data, and suggests the necessity of instructional support for students as they carry out an investigation.

Identifying patterns or trends in data. Students often have difficulty extracting accurate or representative patterns from both first-hand and second-hand data (Schauble, Glaser, Duschl, Schulz, & John, 1995). One particular extensively researched topic is the difficulty that students have in identifying patterns from different graphical forms. Studies have indicated there are three areas that can lead to students' difficulty in understanding graphs (reviewed in Shah & Hoeffner, 2002). These areas are connected to the students' own understanding of graphs, the type of graph that is used (i.e., use of a bar or line graph), and the students' content knowledge about what the graph is depicting (McDermott, Rosenquist, & van Zee, 1987). All of these factors can influence students' success in correctly interpreting data presented in tables and graphs.

Use of personal experiences to explain data. Students use their personal experiences to understand scientific phenomena and to provide them with a unique perspective on the phenomena (Warren, Ballenger, Ogonowski, Rosebery & Hudicourt-Barnes, 2001). In doing so, students access different 'funds of knowledge' generated from experiences throughout their lives (Gee, 1996; Moll, Amanti, Neff, & González, 1992). These 'funds of knowledge' represent the knowledge students have as a result of home culture, extra-curricular, or school activities. In negotiating these different spheres, students draw from and build different knowledge resources that can be used to develop an understanding about science concepts and practices. Understanding how these spheres interact becomes important when students bring the everyday context to the science classroom and use it to explain scientific data or concepts. Moje, Collazo, Carrillo, and Marx (2001), in analysing a project-based science classroom in which students and teachers engaged in a range of inquiry practices, documented that students often draw on different funds of knowledge to explain issues in science class, but do not develop an understanding about the scientific content or practices. Warren et al. (2001) argue that students' everyday ways of talking and thinking allow them to access and understand scientific phenomena. This suggests that, depending on the particular context, students' personal experiences can be either a resource or a barrier to their learning of science.

Conclusions from evidence. Another important aspect of analysing data is drawing conclusions from evidence and justifying those conclusions. This idea is often discussed in the science education community both in terms of scientific explanation (Sandoval, 2003) and argumentation (Erduran, Simon, & Osborne, 2004; Jiménez-Aleixandre, Bugallo Rodríguez, & Duschl, 2000). When drawing conclusions from evidence, students tend to make claims or conclusions without adequate justification (Sadler, 2004). Classroom discussions are often dominated by claims with little backing for those claims (Jiménez-Aleixandre et al., 2000). Students rarely support their conclusions with specific data (Germann & Aram, 1996); instead, they will often use personal views when drawing conclusions (Hogan & Maglienti, 2001). Although students can discuss multiple limitations of data, they often do not consider measurement error when expressing confidence in their conclusions; rather, they draw from prior knowledge or current experimental data to justify their confidence (Masnick & Klahr, 2003). Students do not appear to think about error or limitations in data when analysing data or drawing conclusions. Koslowski (1996) found the conclusions people draw from data depend on whether they can imagine an underlying mechanism for a pattern. Consequently, the conclusions they draw depend on both their prior knowledge of a possible mechanism and on the patterns in the actual data. For middle school students, Koslowski identified that they were more likely to rely on implausible mechanisms than older individuals, which she attributed to the younger students having less background knowledge. There appears to be a positive relationship between students' conceptual understanding and their ability to reason in science (Sadler, 2004). Middle school students with stronger content knowledge are more likely to provide appropriate claims and evidence for those claims (McNeill, Lizotte, Krajcik, & Marx, 2006). This suggests that, when middle school students analyse data, their understanding of evidence and the relevant scientific principles influences their ability to analyse data.

Summary

We are interested in extending these studies to see whether students' ability to analyse and use data depend on whether the data are first-hand or second-hand data. Previous research suggests this might be the case as the two data types have different identified benefits. For example, first-hand data may provide students with ownership, resulting in greater engagement with use of the data, while second-hand data may provide the entire class with a common experience to focus data discussion. In this paper, we investigate what are the potential benefits and limitations of having middle school students engage in first-hand and second-hand data analysis in a classroom learning environment. In order to examine this, our data analysis focused on the classroom conversations resulting from lessons based on these two types of data.

Method

Our study is a design experiment in which we examined the classroom discussions of two middle school classrooms around the use of data. Design experiments combine the design and implementation of innovative learning environments and study them in the complexity of real classrooms (Brown, 1992; Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). Since context is fundamental to educational research, its consideration is essential. If we want to understand learning and instruction in classroom contexts, we need to study them in their naturalistic settings (Brown, 1992). However, this creates the challenge of drawing valuable conclusions from the complexity and messiness of a design-based research study (Barab & Squire, 2004). We try to balance these tensions in our reporting of our findings from this exploratory observational study in which we investigate classroom discussion of data using two different contexts.

Instructional Context

Our analysis focused on the enactment of two curriculum units, chemistry in Grade 7 (McNeill et al., 2003) and biology in Grade 8 (Reiser et al., 2003) developed as part of an National Science Foundation-funded instructional materials development project (Krajcik, McNeill & Reiser, in press). By examining two different content areas we hoped to connect our findings to the type of data used, and not simply to the content studied. We felt that looking across content areas and grades better allowed us to address the question of data type. The 6–8-week chemistry unit 'How can I make new stuff from old stuff?' (referred to as 'chemistry') engages students in the study of substances and properties, the nature of chemical reactions, and the conservation of matter. The 8-week biology unit 'Who will survive?' (referred to as 'biology') engages students in the study of species' interactions in ecosystems, including structure/function, variation, competition, and natural selection. A major design feature of these units is the systematic and scaffolded use of various scientific inquiry practices, particularly with respect to data analysis and interpretation, the creation and/or use of varied representations, and constructing evidence-based explanations.

The units were enacted during the 2002/03 school year in one K–8 magnet school located in a large Mid-western city in the USA. In the city school district, over 90% of the students are African-American and come from lower-middle-income to middle-income families. The percentage of students that qualify for free and reduced lunch was 15.2%, which is much lower than the typical school in this district. However, the school has many of the issues often seen in US urban schools (i.e., overcrowding and lack of materials) (Haberman, 1991; Lynch, 2000). There were approximately 37 students in each classroom that participated in this study. When students worked in groups, there were typically six students in each group because of limitations in classroom space (six laboratory tables in each classroom) and materials (i.e., lack of sufficient hotplates and computers).

The study took place in two different classrooms in this same school. Ms Carter enacted the chemistry unit with her seventh-grade students. Mr Davis enacted the biology unit with his eighth-grade students. Mr Davis also had a student teacher, Ms Stevens, who frequently taught the students.

Data Sources

For this observational study, we examined three lessons from the chemistry unit. In two of these lessons, students used data they generated (first-hand data experience); and in the other lesson, students used data that was given to them (second-hand data experience) (Table 1).

During the first 3 weeks of the chemistry unit, students explored two unknown substances (soap and fat) to introduce the concepts of substance and property. In order to develop students' understanding of properties, they investigated solubility, melting point, and density for soap and fat through a series of lessons that included both first-hand and second-hand data. In one lesson focusing on first-hand data,

Table 1. First-hand or second-hand experiences used during analysis of data

Lesson	Discipline	Topic	First-hand or second-hand	Days
Lesson 2	Chemistry	Solubility	First-hand	2
Lesson 4	Chemistry	Density	First-hand	2
Lesson 5	Chemistry	Properties	Second-hand	1
Lesson 13	Biology	Variation	First-hand	3
Lesson 14	Biology	Sexual selection	Second-hand	2
Lesson 16	Biology	Natural selection	Second-hand	5

students conducted an experiment to determine the solubility of both soap and lard (a type of fat) in oil and water. The teacher modelled a procedure for testing solubility using butter as the solute and using water and oil as the two solvents. Using this technique, students conducted their own investigations of solubility for lard and soap.

In a subsequent lesson in which students used first-hand data, students collected data to calculate the density of both soap and lard as a way of developing a concrete understanding of density. This lesson began with a discussion of ways to investigate whether two things that appear to be the same are actually the same. Students determined the mass of three metal blocks—an aluminium block, an iron block, and an unknown metal block. The teacher used this activity to introduce the idea that mass is not a property. The students continued to develop this idea by massing the same volume of water, rubbing alcohol, and unknown clear liquid. The teacher used the results of this first-hand investigation to introduce the concept of mass per unit volume or density. Using the techniques they learned through these investigations, the students designed and executed their own density experiment using samples of soap and fat. Using the results of this first-hand investigation, the teacher led the students in a discussion about how density is another piece of evidence for whether fat and soap are the same substance or different substances.

For the second-hand data experience in the chemistry unit, students analysed two data-tables. The initial, less complex data-table focused on the properties of colour and density for three unknowns. After analysing this data-table, students examined a second, more complex, data-table that included information about an additional property, melting point, for the three unknowns. They used these tables to determine whether the unknowns were the same or different substances.

We analysed three lessons from the biology unit, one lesson in which students used data they generated (first-hand data experience) and two lessons in which students used data that were given to them (second-hand data experience) (Table 1). The biology unit is divided into two parts. The first part addresses ecosystems, structure/function, species interaction, and food webs in the context of a sea lamprey invasion in the Great Lakes. The second part of the unit focuses on understanding competition, environmental change, and differential survival in the context of a crisis in the Galapagos Islands. All three lessons analysed for this project were taken from the second part of the unit.

For the first-hand data experience in the biology unit, students measured the leg length of preserved grasshoppers to develop an understanding of variation in a population. Students constructed and analysed histograms of the data they collected in order to make meaning of the numbers. Subsequently, students predicted what would happen to the population if a forest fire occurred. The goal in this activity was for students to understand how environmental pressures can become selective pressures when there is variation in a trait within the population.

The two second-hand data experiences focused on differential survival: one based on sexual selection, and the other on natural selection. In one lesson, students examined scientific data collected by a female scientist about variation in tail eye-spot number and size in male peacocks. Students were asked to make predictions about the relationship between the males that were successful in mating and the physical characteristics of the offspring in the next generation. In the second lesson, students used a software application (BGuILE) to manipulate data collected by scientists about variation in a finch population in order to answer questions about why certain finches survived and others did not (Reiser et al., 2001). Students examined the age, weight, wingspan, sex, and beak shape of surviving finches as well as a range of environmental factors. They analysed these second-hand data to develop explanations about what might have influenced the survival of the finches. The product of this investigation was then presented to the class.

Lessons in each unit lasted between 1 and 5 days. All lessons were videotaped for a total of 15 days of instruction. During full class discussions, the camera focused on the individual, either the student or teacher, capturing what individuals said by using a boom microphone. During group work, the camera focused on one group of students, using a table microphone to capture student talk. All videos were transcribed and then analysed.

Data Analysis

We coded the transcripts from each lesson to characterize students' discussions of scientific data. We developed the coding schemes from the theoretical framework presented in this article and an iterative analysis of the data (Miles & Huberman, 1994). The coding scheme focuses on common student difficulties with data analysis, including codes for data measurement, limitations of data, data source, data manipulation, patterns/trends, conclusions from data, and consideration of content knowledge (Table 2).

We selected these codes because they are important in examining how students work with different types of data and have been shown in the literature to be areas of difficulty with students' use of data. One of these codes (data measurement) was only used in the first-hand data experiences as our definition of second-hand data excluded this code for second-hand data experiences. All our codes had multiple criteria to capture a complete picture of how students discussed the data. These

Table 2. Description of codes

Code	Evaluation criteria	Description
Data measurement—only first-hand experiences	Measure data	Students' data were evaluated for the presence or absence of necessary measurements.
	Accuracy	Students' discussions of data were examined for accurate and inaccurate measurements.
	Agreement	Data measurements and discussions were evaluated for agreement or consensus.
Limitations of data	Design limitation	Students do/do not discuss any limitations in the data.
	Execution limitation	Students do/do not discuss any difficulty with completing the experiment.
	Measurement limitation	Students do/do not discuss any difficulty with measuring.
	Representation limitation	Students do/do not discuss any difficulty with representing data.
Data source	Identify	Students identified where the data came from either vaguely or explicitly.
	Explicitly identify	Students explicitly identified data sources by claiming ownership of the data or by stating who collected the data.
Data manipulation	Manipulate data	Students' representations/calculations using data were evaluated for presence.
	Appropriateness and completeness of manipulation	Students' representations/calculations using data were evaluated for appropriateness and completeness.
Patterns/inferences	Identify patterns or inferences	Students explicitly identified patterns or inferences from the data.
	Accurate patterns or inferences	Students were evaluated for identifying accurate patterns or inferences.
Conclusions	Conclusions created	Students created conclusions from the data.
	Accuracy	Students' conclusions were examined to determine whether they were all accurate.
	Data-based	Students' conclusions were evaluated to determine whether they were all based in relevant data.
	Use of personal experience	Students' conclusions were explicitly connected to personal experiences.
Consideration of content knowledge	Draw on content knowledge	Students' content knowledge was explicitly used in data discussion.
	Accuracy of content knowledge	If students used content knowledge, it was evaluated for accuracy and appropriateness.
Use of everyday examples or analogies	Everyday examples and analogies	Students' use of everyday examples and analogies to help clarify the data being discussed were examined for presence and accuracy.

criteria included: presence or absence of the code, characteristics of the discussion, and an evaluation of quality. For example, in the data measurement code, we first evaluated whether or not students measured data (presence or absence). Then we had two other criteria, agreement (characteristic of discussion) and accuracy (evaluation of quality). The agreement criterion measured whether, at the end of the episode, students agreed on the data measurements or whether there was a class consensus about the data. The accuracy criterion captured whether all students' measurements were appropriate for the content addressed.

In order to code the transcripts, we split each day of a lesson into a series of episodes. An episode is a period of time that we identified by the teacher and student interactions and the focus of the class. When the classroom interactions shifted from full class to small group or from small group to full class, we classified this as a separate episode. We also identified a section as an episode if the focus of the conversation shifted. For example, during a full class discussion in the solubility lesson, the teacher first demonstrated how to test the solubility of a substance and then shifted the conversation to a discussion of students' predictions for their solubility experiment. We classified this full class discussion as two separate episodes—one where the teacher demonstrated how to test solubility, and one in which students predicted what would happen in their own investigation. We then analysed the discussion in each episode using the eight different codes (see Table 2). Each day consisted of between one and four episodes, for a total of 45 episodes. The authors each coded the episodes. We selected 20% of the episodes (nine episodes) to be coded independently by both authors. The selected episodes included at least one episode in each lesson and included both full class and small group discussions. Our estimates of inter-rater reliability were calculated by percentage agreement. Agreement was 92% with all disagreements resolved through discussion.

There were 21 first-hand data episodes and 24 second-hand data episodes. We calculated the percentage of occurrence for each criterion by dividing the number of episodes where the specific criterion was identified by either 21 if a first-hand data lesson or by 24 if a second-hand data lesson. We subsequently calculated the percentages of specific qualities or characteristics by examining the number of episodes where the criterion was initially identified. These percentages do not add up to 100% due to the range of sightings for each criterion. We present these percentages in our findings in order to describe trends that we identified from the different data points. The data reduction done by calculating percentages for each criterion in either the first-hand or second-hand data discussions allowed us to identify patterns across the different contexts.

Findings

Our study addresses the following research question: What are the characteristics and the quality of classroom conversations when students use first-hand data compared with when students use second-hand data? Our findings are presented below for each of the nine different codes we analysed. In examining our data at the episode level, it appeared certain discussions occurred more frequently when the class was engaged in experiences with first-hand data while other discussions were more likely to occur during experiences with second-hand data. We were interested in identifying not only *if* the different aspects of data analysis were discussed with the different data types, but also the *quality* of the discussion. In examining our findings, we did not view the frequency as being the sole indication of a benefit or a limitation of a particular data type; rather, we examined the accuracy and completeness of that discussion surrounding the data.

Data Measurement

In examining students' data measurement, we examined first-hand experiences only, as students did not engage in data collection during the second-hand lessons. Students engaged in data measurement in 57% of the first-hand episodes, suggesting that a large percentage of the classroom time is spent in the data collection phase of the investigations (Table 3).

When students engaged in data measurement, the majority of measurements made by the students were accurate. Furthermore, during the majority of the episodes, students came to consensus regarding the quality of the actual measurements. Because of this, we see experiences with first-hand data as providing a benefit that cannot occur in experiences with second-hand data. In the first-hand data experiments, this process of data measurement allowed students to focus on how data should be measured and organized.

Limitations of Data

Discussion of limitations of the data occurred in both first-hand and second-hand data experiences, but more frequently in the first-hand experiences. In the first-hand data experiences, students discussed limitations in 67% of the episodes, as compared with 42% in second-hand data experiences (Table 4). This suggests that experiences with first-hand data provide more opportunities for students to discuss data limitation or error. Yet, this distinction became more complicated when we examined the type of limitation.

The types of limitations discussed were different across the two types of data being analysed by the students. In first-hand experiences, there was a more equal distribution across the four types of possible error (design, execution, measurement, and

Table 3. Data measurement

Criterion	First-hand experiences (%)	Second-hand experiences (%)
Occurrence of data measurement	57	NA
Percentage accurate among occurrences	67	NA
Percentage of occurrences with agreement	75	NA

Table 4. Limitations of data

Criterion	First-hand experience (%)	Second-hand experiences (%)
Occurrence of discussions about data limitation	67	42
Discussions acknowledging design limitations	12	27
Discussions acknowledging execution limitations	35	18
Discussions acknowledging measurement limitations	35	0
Discussions acknowledging representation limitations	24	55

representation). However, in the second-hand discussions, we saw no mention of limitations due to issues surrounding data measurement. The majority of the discussion focused on possible problems with the data representation. In both cases, we identified instances in which students discussed how the design of the investigation might have impacted the data collected.

These variations are illustrated below in examples taken from classroom transcripts for each of the units. In examining measurement limitation, we looked to see whether either the students or teacher acknowledged that the data could have been measured or collected incorrectly. In the first-hand data lessons, both whole-class and small-group episodes had instances of discussions focusing on possible incorrect data measurement for each of the two content areas. In the chemistry lesson, where students determined the density of two substances, students repeatedly mentioned problems with measuring the lard when they were trying to determine the density. For example, students commented 'This scale is off' and 'The scale's broken. That's probably why we got these measurements'. Students also acknowledged that human error could be part of the issue with their data collection as they commented on how the teacher cut the lard. They said 'If she cut it evenly …', suggesting that, since the lard was not in cubes, this could have impacted the data they collected. In the second-hand data experiences, students never discussed that the data could have been measured incorrectly.

However, the second-hand data experiences provided opportunities for students and teachers to discuss problems with their data representations. This is illustrated in the group discussion below taken from the biology unit. In this discussion, the students discussed the limitations of the graphical representations of the data under investigation.

Molly: But you can't tell the difference between the ones that died and the ones that survived!

Student 2: Yes, I can because they're two different graphs.

Molly: I wanna know why they don't make their stuff simpler, because it's too confusing. You know how on the line graphs, they had two different things on there and it was color coordinated. You know, like blue might be survived and red might be died. ...

Student 3: I told you, it's all on one line, Molly.
Student 2: Can you do it in a chart for this stuff?
Student 3: Just do them different and compare.
Student 2: I know, but I can't see it unless I got a chart.

In this excerpt, Molly critiqued the graph as being too complex and another student points out that she needs to see it in chart form in order to understand the data. These comments suggest possible limitations of the current representations. This type of discussion occurred less frequently in the experiences focusing on first-hand data. One reason for this might have been the complexity of the biology second-hand data-sets, which created a need to construct multiple representations. In the first-hand data experiences, the data the students collected did not include as many cases of complex variables. This suggests that second-hand data may provide students with an opportunity to look at more complex data and multiple representations of data.

Data Source

We examined whether the class discussed the data as being first-hand (e.g., we collected them) versus second-hand (e.g., someone else collected them). In both first-hand and second-hand data experiences, the class identified the source of the data using either explicit labels or vaguely claiming ownership (Table 5).

However, students were more likely to identify the data source in the first-hand data experiences. For example, an explicit discussion of where data came from was seen in the chemistry lesson where students determined whether or not fat and soap were soluble in water or oil. Students referred to 'our results' and 'our difference', indicating that they claimed ownership of the data. Later in this same lesson, the chemistry teacher mentioned the ownership of the data when talking to a student in a whole class discussion.

> Right, the experiment that you did. DeShawn, you're going to get information from the experiment that you did. So, you should actually in your evidence refer to what you observed. You did the experiments—the investigations. You all obtained the results based upon your verbal communications to my questions and your answers. You guys came up with the correct results .... (Ms Carter)

Table 5. Data source

Criterion	First-hand experiences (%)	Second-hand experiences (%)
Occurrence of discussions identifying data source	81	54
Discussions acknowledging source explicitly	65	54

The teacher repeatedly discussed how the students collected the data. Interestingly, in second-hand data experiences the class was less likely to discuss that someone else collected the data; rather, they were more likely not to mention where the data came from.

Data Manipulation

Data manipulation referred to students either manipulating their data in a calculation (e.g., calculating a mean) or by creating a representation (e.g., a line graph). Students were more likely to manipulate data in second-hand data experiences, but they made more appropriate and complete manipulations during the lessons that focused on first-hand data (Table 6).

For example, the transcript below is from the chemistry lesson in which students calculated the volume and then the density using the first-hand data they collected.

> Student 6: What was the height?
> Student 2: One!
> Student 6: And what was the width?
> Student 5: It's 3.5! No, not the width, no the width. The width is one. The height is one, the length is …
> Student 2: So the volume …so the volume, the volume will be 3.5 right?
> Student 5: Right. You don't need a calculator for this one …

After calculating the volume, students then used their calculators to divide mass by the volume and come up with the correct density. This example illustrates that the students appropriately and completely manipulated the data as part of a group discussion.

This ability to manipulate data was also seen in the second-hand data experiences. In the group example below, students and teacher created graphs using a software program to determine why some finches died and others survived. This exchange was identified during a sequence of episodes in which students are just beginning the process of manipulating data.

Table 6. Data manipulation

Criterion	First-hand experiences (%)	Second-hand experiences (%)
Occurrence of discussions about data manipulation	38	54
Percentage appropriate and complete	50	8

> Mr Davis: Ok, but you're looking at—you see the numbers here? That's the number of finches. Remember on the island, when the Grants were talking about all the birds that they had? … So when you have that graph, the graph represented each individual bird, ok? So, if you notice that, let's say, birds 4, 20, 19 were alive in the wet season of '77 and you look at another graph, and you don't see those birds, what does that tell you?

Student 2: That they died.
Mr Davis: Right.
Student 2: But Mr. Davis, look at this. I did the wet and dry season of '76 and looked at the ones that survived.
Student 3: No, that's the wet '73 and dry '76.
Mr Davis: So you have to be careful when looking at the axis. You have to read the title of the graph and also the axis before you even attempt to analyze the data. Ok, some of you are just looking at the information given inside the graph but you're not looking at the x and y-axis, nor are you looking at the title of the graph.
Student 2: I see that, but I'm saying. Ok, so this is the wet season of 1977, so they all died before the dry season of 1976. Or did they grow up?
Mr Davis: But you're not comparing '76. You have to look at '77.
Student 2: Yeah, but you gotta look to the past to see the future. But what I'm saying is, did they all die off, or did they grow up?

In this discussion, students manipulated the data, but they also had difficulty interpreting the graphs that they created. This is an example of an incomplete second-hand data manipulation by the students because they did not create all of the graphs they needed in order to understand the data and answer the question. In our analysis, data manipulation occurred more frequently in the experiences focusing on second-hand data than in the first-hand data experiences. Yet the students also had more difficulty creating complete and appropriate representations in the second-hand data experiences. Similar to the discussions about the limitations of the representations discussed above, this difference could stem from the complexity of the data-sets. The second-hand data-sets tended to be more complicated, which made the representations and calculations more difficult.

Patterns or Inferences

Although students identified patterns or trends in their data for experiences where they used either first-hand and second-hand data, this practice occurred more frequently in second-hand data experience (Table 7). However, similar to the data manipulation code, it was appropriate and complete more often in first-hand data experiences.

In the biology lesson, which used a second-hand data-set specific to why a group of finches died, the patterns and inferences made by the students were not always accurate or complete. This is illustrated in the excerpt below, in which the teacher

Table 7. Patterns or inferences

Criterion	First-hand experiences (%)	Second-hand experiences (%)
Occurrence of discussions identifying patterns or inferences	29	75
Percentage appropriate and complete	60	39

asked the class about the characteristics of the finches that died compared with those that survived.

Mr Davis: What types of traits did they have compared to those that died?
Student 1: The bigger finches had the smaller beaks and the smaller finches had the bigger beaks so the smaller finches died.
Student 2: How do you know that?
Student 1: (Commotion between students). Isn't that true??? That's true though.

One student identified an incorrect pattern that the small finches with the bigger beaks were the ones that died. In this second-hand data lesson, students struggled with both creating and interpreting their representations.

In the chemistry investigation described below, in which students collected their own first-hand data, they correctly constructed inferences from their data about solubility during a small group discussion. Students used their observations to infer whether a substance was soluble.

Student 1: The fat is not soluble in water. It just clogged up the tube.
Student 2: Fat is not soluble in water, but it is in oil. I am so smart.
Student 3: Ok, it's not soluble in water, but it's soluble in oil.
Student 4: Alright, we're done. That was easy.

The students' comments illustrate how easy it was for them to make inferences in this context, unlike the second-hand data biology investigation about the finches. The patterns seen in the biology second-hand data experience were more complex than the chemistry inferences from the first-hand data experience. Again, this is similar to what we observed with data manipulation and discussions about the limitations of the representations of the data. The complexity of the data-set becomes both a benefit and a limitation that needs to be addressed. We return to this point in our discussion.

Conclusions Based on Data

In all of the lessons we examined, students frequently drew conclusions with the majority of the conclusions based on data (Table 8). Such conclusions occurred more frequently in second-hand data experiences, but they were accurate and data-based more often in first-hand data experience.

Table 8. Conclusions

Criterion	First-hand experiences (%)	Second-hand experiences (%)
Occurrence of discussions creating conclusions	48	71
Percentage accurate	60	38
Percentage data-based	90	72
Percentage using personal experience	0	13

In the exchange below, the students discuss the importance of using different pieces of evidence to make conclusions during a first-hand data experience in the chemistry unit. In their discussion, students make reference to their prior experiences with data earlier in the unit and what is necessary for them to make a conclusion. The teacher asks the students if they have enough evidence to make their conclusions, and the students respond yes and continue by explaining:

Student 1:	Because in this experiment and in the last experiment, we proved they don't have the same melting point or density.
Student 2:	But the density by itself doesn't, because the densities are 0.9 and 0.8. The density by itself don't. Cause they're so close together and like error. But from past experience, plus the density, it does. Cause by itself it doesn't because 0.9 and 0.8 are pretty close together. Ya'll hear what I'm saying.

In this example, students discuss why they believe they have enough evidence to make a conclusion about why the two substances they are investigating are in fact different substances. Using density by itself is not sufficient, but if they also use their other data (i.e., colour, hardness, solubility, and melting point) they can conclude that they are two different substances. We coded the students' conclusion that they are different substances as accurate, data based, and not using personal experience. This is contrasted to an example in the second-hand data-set shown below, in which the teacher used a personal experience to ground the conclusion that students were working on in the biology unit. Here, the teacher used an example of a school dance and the ratio of boys to girls impacting the dynamics of the dance to clarify issues of sexual selection.

Ms Stevens:	Well, think about what happens if everybody wants a dance, and you've got five females and 20 males.
Students:	[Laughter].
Ms Stevens:	Now, you're all thinking about it. I love it. Two things are happening: 1) the females might become more choosey and 2) the males have to work harder to get the attention of the females. Cause now he's gotta compete. He went from a 1:1 ratio to a 1:4. There are four guys for every one of our girls, right? Same thing is happening here. The data show the behavior. The females are being choosey, and who are they choosing?

The class used this personal example to draw a conclusion about how sexual selection impacted which birds mated. Using personal experiences to draw conclusions did occur more frequently in second-hand data experiences. Yet, overall, students were more likely to use data than personal experiences to draw conclusions.

Consideration of Content Knowledge

In both first-hand and second-hand data experiences, students and teachers used content knowledge to help frame the discussion (Table 9). However, content knowledge was used more frequently and at a higher level of accuracy in the second-hand data experiences.

Table 9. Consideration of content knowledge

Criterion	First-hand experiences (%)	Second-hand experiences (%)
Occurrence of discussions drawing on content knowledge	38	63
Percentage accurate	38	64

In a first-hand data chemistry lesson focusing on solubility, teachers and students used content knowledge in their whole class discussion about solubility, but the content knowledge was not always accurate. The discussion below illustrates the difficulty both the teacher and student had with discussing the concept of solubility.

Ms Carter: ...Now, what was the evidence that fat was soluble in oil? What was the evidence?
Student: That some of 'em disappeared, and like...
Ms Carter: That ... what's it?
Student: The lard.
Ms Carter: Disappeared?
Student: Into the oil.
Ms Carter: Into the oil. Which would be an indication, the indication, that fat is soluble in oil.

Here, the student expressed a common misunderstanding in that he talked about the fat disappearing when it dissolved. We coded this as an inaccurate discussion in terms of the content, because the teacher did not address this misconception.

In an example taken from the chemistry second-hand data lesson, students discuss whether the data they have been given in a table shows that the substances are different or the same. One student comments:

> Otherwise, they both do not have—they both—it depends if they are the same substance if they would show other properties. All of its properties have to be the same for it to be the same substance, but they ain't showing all the properties of it. So it depends—to me. But they are showing in the chart—I said it is a substance. I said it was the same substance from the chart though. (Student 3)

In his comment, the student references the importance of knowing multiple properties, which are not shown in the table, to determine whether the substances are the same or not. The students continue to draw from content knowledge of substances and properties throughout their discussion about the data table.

Student 3: Scientists, I think, I think—they use um several different properties because one or two you ain't going to know if it is the same substance. Say like color and hardness, you don't know if it is the same substance if they both—...
Student 2: He is using it as an example.
Student 3: I am saying if color and hardness is the same you wouldn't know if it is the same substance because they ain't list all of the properties. They are only the same substance if all of the properties are the same. You would have to know the density, the melting point, the solubility—everything.

Student 2: Ok. So using more than one property tells you—I mean
Student 3: Using one or two properties doesn't tell you, doesn't tell you —
Student 2: Doesn't make sure. I mean.
Student 3: Doesn't tell you—
Student 2: xactly—
Student 3: Exactly if they are the same substance.

In this whole-class discussion, the students demonstrate they understand that, in order to determine whether substances are they same, you have to compare multiple properties. They show they understand what a property of a substance is and what makes a substance the same or different from another substance in their discussion. Overall, the accuracy of the science content was higher in the second-hand data conversations compared with the first-hand data conversations.

Use of Everyday Examples or Analogies

The use of everyday examples or analogies occurred in both first-hand and second-hand data experiences, although this was not a common occurrence (Table 10). The examples teachers and students used in the second-hand data experiences tended to be more accurate, although we would be more confident in this finding if we had observed more instances of these examples being used.

In the first-hand example shown below, the teacher draws from everyday experiences in order to clarify the rationale of an activity for the whole class.

Ms Stevens: Ok, the length of their legs. Ok, what we're going to look at today is the length of their femurs. When you have longer legs, it's easier to run faster, right? How many of you have, when you were little, it was hard to keep up with your mom and dad because they had such longer legs. So the longer the legs, typically, the better, quicker, the longer the hops and the faster they are, ok. So, yeah...
Student: So why do short people usually run faster than tall people then?
Ms Stevens: Well, now you're getting into muscle. But in general, the length of your leg can help you be a faster runner, cause you've got a longer stride for your step. But you're getting into muscle, that's it. So what we're going to do today though, most of the grasshoppers are pretty much the same except for the fact that some of them. Well, do you think they're all going to have the same length?

The purpose of the classroom activity that the class was discussing was to develop an understanding of why there is variation of physical traits (i.e., leg length) in populations that change over time. But the science behind this example is actually quite complex

Table 10. Use of everyday examples or analogies

Criterion	First-hand experiences (%)	Second-hand experiences (%)
Occurrence of discussions using everyday	19	25
Percentage accurate	25	83

and served to raise additional questions from the students. The teacher-generated example caused a student to ask why can shorter individuals run faster than tall individuals, thereby pulling the focus of the lesson away from the purpose. This illustrates the importance of having carefully considered the personal examples that might be used in the classroom so as not to confuse or complicate the situation. In our analysis, this example was scored as inaccurate because we felt it complicated and confused the task by bringing in an everyday example that did not help to clarify what the students were measuring or why the length of leg was important.

Discussion

We examined a series of inquiry experiences using either first-hand or second-hand data embedded within two extended project-based science units and identified a series of patterns in students' discussions dependent on the use of first-hand or second-hand data in these learning experiences. Our results suggest that both types of experiences allowed teachers and students opportunities to engage in rich discussions around data in different instructional contexts. However, the strengths and weaknesses of these opportunities differed.

One intriguing trend that emerged from our study is that first-hand data experiences resulted in a greater focus on students gathering data, discussing the limitations of the data, and identifying the data source. This suggests that use of first-hand data experiences may be particularly useful for discussing the limitations or error in data as well as helping students understand that the origin of the data is important. Students may view data as facts or truth and not understand that it is important to consider how data were collected. One objective of teaching science is to have students understand that scientific knowledge is not an absolute truth, but rather that it is a model upon which the scientific community currently agrees to explain natural phenomena, and that this model is periodically modified based on the available data (Duschl, 1990; Kuhn, 1970). In both types of experiences, we saw students and teachers being critical of the source of data (i.e., questioning where the data came from and whose data they were). However, we saw this occur more frequently in the first-hand data experiences. Our finding is similar to Palincsar and Magnusson's (2001) work, where they observed that students did not assume a critical stance to second-hand data, but rather seemed to view it as authoritative. While we saw students willing to accept second-hand data with little questioning about the source, we observed an unexpected amount of ownership that students voiced with regard to their own first-hand data and their willingness to critique and discuss the limitations around it. If the goal of instruction is to help students understand a complete picture of data and how they are used in science, it is important to engage students in first-hand data experiences where they collect their own data, discuss the limitations of those data, and develop a sense of ownership of the data.

In the second-hand data experiences, we saw students frequently engage in data manipulation, identification of patterns, draw conclusions from data, and consider content knowledge—all necessary in order to make meaning from an investigation and key components of the practice of science. We believe this may have occurred more frequently in second-hand data experiences than in first-hand data experiences due, in part, to the nature and complexity of the data given to the students. Although we did not realize this when we began our analysis of classroom discussions, two of the three lessons we analysed for the second-hand data experiences included the most complex data-sets in our analysis. This level of complexity may have influenced the nature of the discussion we observed around the second-hand data. Yet this is also a benefit of second-hand data, which can give students access to more complicated data-sets than they could collect on their own. These more complex data-sets can encourage certain types of conversations around data (i.e., what variables to manipulate, how the data should be graphed, and finding patterns). This complexity can also give rise to additional difficulties. While data manipulation occurred more frequently in second-hand data experiences, the discussions we coded were more likely to be inappropriate or incomplete. Since these data-sets were complicated, more time was needed for organization of data and students had more difficulty understanding the complex patterns identified in the data. However, this difficulty might become an advantage if students are supported in their conversations (Wu & Krajcik, 2006). This highlights the need for supports to be developed and included in the curriculum materials as one possible way for teachers and students to address some of the difficulties around analysing data.

One difference in our study from the reported literature (Germann & Aram, 1996; Sadler, 2004) is that we saw students making claims or conclusions with adequate justification in first-hand and second-hand data experiences. This type of conversation occurred at a higher frequency in the second-hand data experiences, while the accuracy was greater in first-hand data experiences. The lack of accuracy in using the second-hand data may have been because students were trying to make sense of complicated data-sets and were struggling with the representations.

In our analysis of both first-hand and second-hand data experiences, we rarely saw students and teachers using their everyday experiences to make connections to their data under discussion. If we want students to draw on other 'funds of knowledge', we need to identify ways to scaffold these connections between the everyday knowledge that students bring to class and the scientific knowledge that we want students to develop, regardless of the type of data being used (Gee, 1996; Warren et al., 2001).

Implications for Curriculum Development

We believe our study has a number of implications for curriculum development. One important consideration in thinking about our findings is the impact of the unit of analysis on our findings. If we just focused on frequency, then second-hand data experiences appeared to provide students with more opportunities for some types of discussions, such as drawing conclusions. But when we also considered the quality of those discussions, we found that students had more difficulty in the second-hand experiences. We believe both levels of analysis are important and have implications for curriculum development. We want students to engage in conversations around data; however, we need to be concerned about the quality of the discussions.

Recent empirical work by Magnusson et al. (2004) has suggested there is a complex interplay between the sequence of first-hand and second-hand data investigations that can impact student learning. In their study, they suggest there might be a 'preferred sequence and mode of investigation of contexts to support learning' (Magnusson et al., 2004, p. 318). They argue that more research must be done in order to understand more fully the interplay between learning experiences using first-hand and second-hand data. We believe that further research might suggest that the interaction between these two types of learning experiences depends on content and context.

Our findings suggest that first-hand and second-hand data may be appropriate for different learning goals. If the goal is to explore content that is not easily investigated in classrooms (like evolution) or if the learning goal is a scientific inquiry goal such as working with complex data-sets, then second-hand data-sets might be more appropriate to use. On the other hand, if the goal is to help students understand measurement limitations or to help motivate student engagement in the science content, then first-hand data-sets might be more appropriate. A complex array of factors (i.e., scientific discipline, classroom context, and activity structure) will influence a teacher or curriculum designer to use one type of data over another. We believe that more focused studies are needed in order to understand how these factors interact. For example, although we examined two different content areas (i.e., biology and chemistry) in this study, we were not able to tease out the impact of the content area on students' discussions. To better understand the issue of the content, we need to examine a range of data within a specific content as well as across multiple disciplines. Such research will provide guidance in making design decisions as well as develop the necessary supports for teachers and students in using the different data types.

Besides the design of curriculum materials, it is also critical to consider the role of the teacher in how data are used in the classroom. Teachers vary in how they use curriculum materials, and the ways in which they adapt them will influence what happens in the classroom (Remillard, 2005; Schneider, Krajcik, & Blumenfeld, 2005). Consequently, teachers need to be provided with support through various means such as educative curriculum materials (Ball & Cohen, 1996; Davis & Krajcik, 2005) as well as professional development opportunities (Fishman, Marx, Best & Tal, 2003). Such support can help teachers take advantage of the different benefits of first-hand and second-hand data-sets to encourage productive classroom conversations and promote greater student learning.

Acknowledgements

The research reported here was supported in part by the Investigating and Question Our World through Science and Technology project (IQWST) (NSF-ESI-0101780 or NSF-ESI-0439352) and the Center for Curriculum Materials in Science

(CCMS) (NSF-ESI-0227557). Any opinions expressed in this work are those of the authors and do not necessarily represent either those of the funding agency or the University of Illinois Urbana-Champaign or Boston College. The authors would like to thank all of the researchers involved with IQWST and CCMS, especially Joseph Krajcik and Brian Reiser.

References

Aikenhead, G. (2004). Science based occupations and the science curriculum: Concepts of evidence. *Science Education, 89*(2), 242–275.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Ball, D.L., & Cohen, D.K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6–8, 14.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1–14.

Brown, A.L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141–178.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Davis, E.A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher, 34*(3), 3–14.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher, 23*(7), 5–12.

Duschl, R. (1990). *Restructuring science education. The importance of theories and their development.* New York: Teachers College Press.

Edelson, D. (2001). Learning-for-use: A framework for the design of technology-supported inquiry activities. *Journal of Research in Science Teaching, 38*(3), 355–385.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the use of Toulmin's Argument Pattern in studying science discourse. *Science Education, 88*(6), 915–933.

Fishman, B., Marx, R., Best, S., & Tal, R. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education, 19*(6), 643–658.

Gee, J.P. (1996). *Social linguistics and literacies: Ideology in discourses* (2nd ed.). London: Falmer.

Germann, P.J., & Aram, R.J. (1996). Student performances on the science processes of recording data, analyzing data, drawing conclusions, and providing evidence. *Journal of Research in Science Teaching, 33*(7), 573–798.

Haberman, M. (1991). The pedagogy of poverty versus good teaching. *Phi Delta Kappan, 73*(4), 290–294.

Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinnings of students and scientists' reasoning about conclusions. *Journal of Research in Science Teaching, 38*(6), 663–687.

Hug, B., & Krajcik J.S. (2002). Students' scientific practices using a scaffolded inquiry sequence. In P. Bell, R. Stevens, & T. Satwicz (Eds.), *Proceedings of the International Conference of the Learning Sciences,* Seattle, WA: Lawrence Erlbaum Associates.

Jiménez-Aleixandre, M.P., Bugallo Rodríguez, A., & Duschl, R.A. (2000). 'Doing the lesson' or 'doing science': Argument in high school genetics. *Science Education, 84*(6), 757–792.

Kanari, Z., & Millar, R. (2004) Reasoning from data: How students collect and interpret data in science investigations. *Journal of Research in Science Teaching, 41*(7), 748–769.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning.* Cambridge, MA: MIT Press.

Krajcik, J., Blumenfeld, P., Marx, R., & Soloway, E. (2000). Instructional, curricular, and technological supports for inquiry in science classrooms. In J. Minstrell & E. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 283–315). Washington, DC: AAAS.

Krajcik, J., McNeill, K.L., & Reiser, B. (in press). Learning-goals-driven design model: Curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education.*

Kuhn, T.S. (1970). *The structure of scientific revolutions.* Chicago, IL: University of Chicago Press.

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

Lynch, S.J. (2000). *Equity and science education reform.* Mahwah, NJ: Lawrence Erlbaum Associates.

Magnusson, S.J., Palincsar, A.S., Hapgood, S., & Lomangino, A. (2004). How should learning be structures in inquiry-based science instruction?: Investigating the interplay of 1^{st}- and 2^{nd}-hand investigations. In Y. Kafai, W. Sandoval, N. Enyedy, A. Nixon, & F. Herrera (Eds.), *Proceedings of the Sixth International Conference of the Learning Sciences* (pp. 310–317). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Masnick, A.M., & Klahr, D. (2003). Error matters: An initial exploration of elementary school children's understanding of experimental error. *Journal of Cognition and Development, 4*(1), 67–98.

McDermott, L.C., Rosenquist M.L., & Van Zee E.H. (1987). Student difficulties in connecting graphs and physics: examples from kinematics. *American Journal of Physics, 55*(6), 503–513.

McGinn, M.K., & Roth, W-M. (1999). Preparing students for competent scientific practice: Implications of recent research in science and technology studies. *Educational Researcher, 28*(3), 14–24.

McNeill, K.L., Harris, C.J., Heitzman, M., Lizotte, D.J., Sutherland, L.M., & Krajcik, J. (2003). How can I make new stuff from old stuff? In J. Krajcik & B. J. Reiser (Eds.), *IQWST: Investigating and questioning our world through science and technology.* Ann Arbor, MI: University of Michigan.

McNeill, K.L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In M. Lovett & P. Shah (Eds.), *Thinking with data: The Proceedings of the 33rd Carnegie Symposium on Cognition* (pp. 233–265). Mahwah, NJ: Taylor & Francis.

McNeill, K.L., Lizotte, D.J, Krajcik, J., & Marx, R.W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences, 15*(2), 153–191.

Metz, K. (2000). Young children's inquiry in biology: Building the knowledge bases to empower Independent Inquiry. In J. Minstrell & E. van Zee (Eds.), *Inquiring into inquiry learning and teaching on science* (pp. 371–404). Washington DC: American Association for the Advancement of Science.

Miles, M., & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Moje, E.B., Collazo, T., Carrillo, R., & Marx, R.W. (2001). Maestro, what is 'quality': Language, literacy, and discourse in project-based science. *Journal of Research in Science Teaching, 38*(4), 469–498.

Moll, L.C., Amanti, C., Neff, D., & González, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice, 31*(2), 132–141.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

Osborne, J. (2002). Science without literacy: A ship without a sail? *Cambridge Journal of Education, 32*(2), 203–218.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What 'Ideas-bout-science' should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching, 40*(7), 692–720.

Palincsar, A.S., & Magnusson, S.J. (2001). The interplay of first-hand and text based investigations to model and support the development of scientific knowledge and reasoning. In S. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 151–193). Mahwah, NJ: Lawrence Erlbaum Associates.

Petrosino, A.J., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning, 5*(2&3), 131–156.

Reiser, B.J., Tabak, I., Sandoval, W.A., Smith, B.K., Steinmuller, F., & Leone, A.J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263–305). Mahwah, NJ: Lawrence Erlbaum Associates.

Reiser, B., Tzou, C., Dodick J., Hug, B., Finn, L.E., Sy, M., & Bruozas, M. (2003). Who will survive? In J. Krajcik & B.J. Reiser (Eds.), *IQWST: Investigating and questioning our world through science and technology*. Ann Arbor, MI: University of Michigan.

Remillard, J.T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246.

Sadler, T.D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*(5), 513–536.

Sandoval, W.A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *The Journal of the Learning Sciences, 12*(1), 5–51.

Schauble, L., Glaser, R., Duschl, R.A., Schulz, S., & John, J. (1995). Students' understanding of objectives and procedures of experimentation in the science classroom. *The Journal of the Learning Science, 4*(2), 131–166.

Schecker, H. (1998). Integration of experimenting and modeling by advanced educational technology: Examples from nuclear physics. In K. Tobin & B.J. Fraser (Eds.), *The international handbook of science education* (Part I, pp. 383–398). Dordrecht, The Netherlands: Kluwer.

Schneider, R.M., Krajcik, J., & Blumenfeld, P. (2005). Enacting reform-based science materials: The range of teacher enactments in reform classrooms. *Journal of Research in Science Teaching, 42*(3), 283–312.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 14*(1), 47–69.

Songer, N.B. (1996). Exploring learning opportunities in coordinated network-enhanced classrooms: A case of kids as global scientists. *The Journal of the Learning Sciences, 5*(4), 297–327.

Warren, B., Ballenger, C., Ogonowki, M., Rosebery, A.S., & Hudicourt-Barnes, J. (2001). Rethinking diversity in learning science: The logic of everyday sense-making. *Journal of Research in Science Teaching, 38*(5), 529–552.

White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3–118.

Wu, H.-K., & Krajcik, J. (2006). Inscriptional practices in two inquiry-based classrooms: A case study of seventh graders' use of data tables and graphs. *Journal of Research in Science Teaching, 43*(1), 63–95.