International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

# Over reported and misunderstood? A study of teachers' reported enactment and knowledge of inquiry-based science teaching

Daniel K. Capps, Jonathan T. Shemwell & Ashley M. Young

**To cite this article:** Daniel K. Capps, Jonathan T. Shemwell & Ashley M. Young (2016) Over reported and misunderstood? A study of teachers' reported enactment and knowledge of inquiry-based science teaching, International Journal of Science Education, 38:6, 934-959, DOI: 10.1080/09500693.2016.1173261

**To link to this article:** http://dx.doi.org/10.1080/09500693.2016.1173261

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
VOL. 38, NO. 6, 934–959
http://dx.doi.org/10.1080/09500693.2016.1173261

Routledge
Taylor & Francis Group

# Over reported and misunderstood? A study of teachers' reported enactment and knowledge of inquiry-based science teaching

Daniel K. Capps[a], Jonathan T. Shemwell[b] and Ashley M. Young[c]

[a]Mathematics and Science Education, University of Georgia, Athens, GA, USA; [b]College of Education and Human Development, University of Maine, Orono, ME, USA; [c]Science Department, Innovation Academy Charter School, Tyngsboro, MA, USA



**ABSTRACT**

Science education reforms worldwide call on teachers to engage students in investigative approaches to instruction, like inquiry. Studies of teacher self-reported enactment indicate that inquiry is used frequently in the classroom, suggesting a high level of proficiency with inquiry that would be amenable to inquiry reform. However, it is unclear whether the high frequency of self-report is based on sound knowledge inquiry. In the absence of sound knowledge, high rates of self-reported enactment would be suspect. We conducted a study to measure teachers' knowledge of inquiry as it related to the known, high frequency of reported enactment. We developed a multidimensional survey instrument using US reform documents and administered it to 149 K–12 teachers at a national science teachers' conference. The majority of the teachers surveyed did not report inquiry enactment based on well-structured knowledge of inquiry. Interviews with participants showed how teachers could readily map non-inquiry activities onto inquiry statements taken directly from reform documents. From these results we argue that teachers often believed they were enacting inquiry, when likely they were not. We further reason that teachers may struggle to interpret and enact inquiry-related requirements of science education reform and will need support distinguishing inquiry from non-inquiry practices.

**ARTICLE HISTORY**
Received 2 April 2015
Accepted 29 March 2016

**KEYWORDS**
Inquiry-based teaching; science education reform

Over the past several decades, science education reform worldwide has given priority to investigative approaches to science teaching wherein learners pursue evidence-based answers to scientific questions (e.g. Department for Education and Skills/Qualification and Curriculum Authority, 2004; Ministry of Education of Singapore, 2007; National Research Council [NRC], 1996, 2000; *Tomorrow 98*, 1992/1994; Tytler, 2007). Various investigative approaches have evolved, including inquiry-based instruction, practical work, project-based teaching, learning through investigation, and most recently, learning

**CONTACT** Daniel K. Capps dacapps@uga.edu Mathematics and Science Education, University of Georgia, 110 Carlton Street, Athens, Georgia 30602, USA

Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/09500693.2016.1173261

through and of the practice of science (Abd-El-Khalick et al., 2004; Dillon, 2008; NRC, 2012). A commonality across all of these approaches is that they are challenging for teachers to enact (Crawford, 2000). Indeed, descriptive studies have shown how teachers can struggle to understand what investigative teaching is and how to enact it in the classroom (Abrahams & Millar, 2008; Abrams, Southerland, & Silva, 2008; Bybee, 2000; Capps & Crawford, 2013a; Ireland, Watters, Brownlee, & Lupton, 2012; Ozel & Luft, 2013). However, studies measuring the extent to which teachers tend to become reasonably proficient at investigative teaching are non-existent. The best available information of this sort comes from surveys in which teachers report how frequently they enact various aspects of investigative teaching in their classroom. Interestingly, these self-report studies tend to show a high rate of enactment (Beatty & Woolnough, 1982; Marshall, Horton, Igo, & Switzer, 2009), hinting at a degree of facility with investigative teaching that is inconsistent with the struggles evident in small-scale descriptive studies. Thus, there is much uncertainty about the extent to which teachers at large may be 'truly' enacting investigative teaching. This uncertainty poses a threat to rolling out new initiatives supporting investigative teaching, for example the shift to the *Next Generation Science Standards* (NGSS Lead States, 2013) in the U.S, as it is unclear how prepared teachers will be to enact investigative instructional approaches that build on ideas from previous reforms.

It would be useful if reported enactment measures could be improved to gain a better sense of teachers' proficiency with the investigative teaching they report on. One avenue of improvement would be to check the degree to which participants interpret survey items based on well-structured knowledge of investigative teaching. High frequencies of investigative teaching would be suspect, for instance, if teachers reporting those frequencies did not have good knowledge of investigative teaching. Thus, reported enactment would be interpreted with greater understanding of teachers' proficiencies. We took precisely this approach in the present study, focusing on a form of investigative teaching known as inquiry-based instruction (AAAS, 1989; NRC, 1996, 2000) which has linkages to a more recent conception of learning of and through the practice of science (NRC, 2012). Specifically, we designed and administered a survey to find out how teachers conceptualized inquiry-based instruction (hereafter referred to simply as inquiry), and how frequently they reported enacting various aspects of inquiry in their classrooms. Participants were a select population of K-12 teachers from across the U.S. whom we expected to be well informed about inquiry, relative to the general U.S. population of teachers. Our findings were that many teachers reported doing inquiry activities quite frequently, but surprisingly few had well-structured knowledge of inquiry. Based on these results, and on existing literature that shows the degree of difficulty that inquiry can pose for teachers, we argue that teacher educators and professional developers should take a conservative view of teachers' general level of proficiency with inquiry. By extension, they should guard against overly sanguine expectations for teachers' preparedness to enact new reforms that incorporate inquiry.

## *Inquiry-based instruction*

Within inquiry, teachers have students use data as evidence to answer scientifically oriented questions (Anderson, 2007; NRC, 1996, 2000). Developing inquiry as an approach to teaching has been a major initiative in science education over the last half century (Crawford, 2014; Deboer, 1991; Jiang & McComas, 2015). Project 2061, a long-term initiative of the American Association for the Advancement of Science (AAAS), advocated integrating scientific investigation with content learning, and placed an emphasis on inquiry as a teaching strategy (AAAS, 1989). The U.S. *National Science Education Standards* (*NSES*; NRC, 1996), along with other reform efforts worldwide (e.g. Ministry of Education of Singapore, 2007; *Tomorrow 98*, 1992/1994), adapted Project 2061's basic premise that students could learn about science by engaging scientifically with phenomena (i.e. through 'inquiry'). One problem, however, was that inquiry was not well defined within these movements. For example, two of the six science teaching standards of the *NSES* discussed inquiry, but nowhere in the standards was inquiry explicitly defined or operationalized. A follow-up publication, *Inquiry and the National Science Education Standards* (*INSES*) (NRC, 2000), attempted to solve this problem. *INSES* was styled as a practical guide for teachers, professional developers, and administrators. It operationalized inquiry through five essential features which included having learners: (1) engage in scientifically oriented questions, (2) give priority to evidence, (3) formulate explanations from evidence, (4) evaluate explanations in light of alternative explanations, and (5) communicate and justify their explanations (NRC, 2000). In addition, the authors provided information on the role of the teacher in inquiry by discussing variations in the amount of structure a teacher might provide, and the extent to which a teacher might involve students in each of the features. Unfortunately, despite *INSES*, confusion lingered over what it meant to teach science as inquiry (Crawford, 2014). This confusion was, in part, the impetus for a new wave of reform resulting in the most recent documents describing investigative teaching in the U.S., the *Framework for K-12 Science Education* (the *Framework*) and *Next Generation Science Standards* (*NGSS*). These documents reconceptualized inquiry as eight scientific practices. Among the various reform documents, the *Framework* and *NGSS* provide the greatest degree of specification of what students should know and be able to do within inquiry.

## *Knowledge and enactment of inquiry*

Our understanding of what teachers know about inquiry is underdeveloped (Keys & Bryan, 2001). One area of teacher knowledge that appears important for the enactment of inquiry is their knowledge of the principles of inquiry teaching (Blanchard, Southerland, & Granger, 2009; Kennedy, 1998). Having a better sense of what teachers know about these principles may enable the development of more accurate ways to measure inquiry enactment on a broad scale, and potentially help to clear up some of the inconsistencies between what teachers think they are doing and what they are actually doing in their classrooms related to inquiry. In what follows we review the literature on teacher knowledge and enactment of inquiry.

### *Teacher knowledge of inquiry*

Many studies have shown how teachers' knowledge of inquiry can be incomplete and have variable alignment with inquiry as conceived in reform documents (e.g. Brown, Abell, Demir, & Schmidt, 2006; Demir & Abell, 2010; Lotter, Harwood, & Bonner, 2006). These same studies show that teachers' descriptions of inquiry can include ideas such as hands-on work or discovery learning that can easily be done outside of an inquiry orientation. A recurring pattern in these studies is that teachers tend to describe inquiry in terms of certain salient features (e.g. questioning and investigating), and neglect other important, but somehow less-salient features, such as working with data or reasoning to construct arguments. As an example, Demir and Abell (2010) investigated the knowledge of inquiry held by four beginning teachers from alternative certification programs by asking what teaching science as inquiry meant to them. The teachers described inquiry as featuring student-generated questions and data collection. They neglected evidence, explanation, justification, and communication in their descriptions. Moreover, the authors noted that the teachers included activities that were only loosely related to inquiry such as problem-solving, amount of teacher guidance, and discovery learning. In a study that investigated 19 college science professors' knowledge of inquiry, Brown et al. (2006) found that college faculty also tended to think of inquiry in terms of certain features, mostly as questioning and data collection, neglecting other less-salient features.

Ozel and Luft (2013) used interviews to investigate the understandings of inquiry held by 44 beginning secondary science teachers. They found that teachers in their first few years of service described inquiry in terms of questioning and giving priority to evidence, and left out less-salient inquiry practices such as explanation with evidence, connection to scientific knowledge, and communicating scientific ideas. Capps and Crawford (2013a) looked at the understandings of inquiry held by 26 fifth- through ninth-grade teachers. The researchers asked the teachers to describe inquiry, both in writing and in interviews. The teachers generally equated inquiry with hands-on work and discovery learning. A minority of the teachers described inquiry in terms of using data as evidence to investigate scientific questions. Finally, Kang, Orgill, and Crippen (2008) assessed the inquiry knowledge of 34 teachers from a large urban school district. The authors had teachers classify short teaching scenarios as representative of inquiry or not, and analyzed teacher-defined characteristics of inquiry as either consistent or not consistent with the essential features of inquiry (NRC, 2000). Similar to studies previously discussed, Kang and colleagues found that out of five essential features, in general teachers identified engaging in scientific questions, giving priority to evidence, and formulating explanations in their classifications of the scenarios, but typically left out somewhat less-salient features, such as evaluating explanations and connecting these explanations to scientific evidence and communicating explanations. In their conclusions, the authors speculated that teachers' understanding of inquiry might develop over time, beginning with viewing inquiry in terms of its most surface-level aspects, collecting and explaining evidence. As their understanding of inquiry progressed they might see the need to engage students in scientifically oriented questions. Having students develop explanations and share them to a broader audience would come later.

### *Teachers' enactment of inquiry*

Estimates of the frequency of teacher enactment of inquiry range from very frequent, according to teacher self-report surveys, to very rare when observed directly. Marshall et al. (2009) measured, among many other variables, reported use of inquiry by teachers in a large school district ($N = 1222$ teachers). They included an item which asked survey participants to choose the percentage of instructional time their students were 'engaged in inquiry during a typical lesson' (p. 581). Teachers reported using inquiry an average of 38.7% of the time. This result is difficult to interpret because activities that comprised inquiry were left to the participants' interpretation. A similar limitation can be found in a large-scale study by Banilower, Heck, and Weiss (2007), who analyzed self-report surveys of 18,657 K-8 teachers participating in a curriculum reform project that included inquiry. The authors had many measurement goals, only one of which was to assess the effect of professional development on teachers reported use of investigative teaching practices. Survey items related to investigative teaching asked participants about their use of 'hands-on activities' and 'working on models or simulations'. Though the authors did not attempt to use their data to report on the overall frequency of investigative teaching, had they done so, much like Marshall et al., it would have been unclear whether responses were based on enactment of investigative teaching, or based on activities that had surface features in common with it, like hands-on teaching.

If teachers could provide valid self-reports of their use of inquiry, this information would be immensely useful in assessing the state of inquiry teaching on a broad scale. However (and setting aside untruthful reporting), there are obvious problems with the validity of self-reported enactment. One, which we briefly pointed out in the introduction, is that teachers who do not have well-structured knowledge of inquiry may over- or under-report enactment as they interpret survey items in ways that depart from established norms. Over-reporting is a particular concern if teachers map non-inquiry activities onto survey items asking about inquiry. Another possibility is that teachers may unconsciously bias their response toward more frequent enactment because they know that inquiry activities are generally associated with good teaching, a form of acquiescence bias (Messick & Jackson, 1961). Further problems arise when studies with broad research goals use measurement frameworks that are not specifically geared toward inquiry enactment, so that measures of enactment lack precision (e.g. Banilower, et al., 2007).

Enactment studies based on classroom observations using trained observers have fewer validity concerns than those using self-reported enactment. Yet, problems of interpretation still arise when instruments are not specifically focused on inquiry. For instance, in a large-scale study, Weiss, Pasley, Smith, Banilower, and Heck (2003) reported that 15% of science lessons in elementary schools (K-5) focused on scientific inquiry, while only 2% of lessons in grades 9–12 did. These findings suggest a much lower frequency of enactment of inquiry than self-report surveys. However, the observation protocol upon which these findings were based did not address any specific aspects of inquiry. It instead asked the observer to rate more general teaching practices such as the extent to which 'the design of the lesson incorporated tasks, roles, and interactions consistent with investigative mathematics/science' (Weiss et al., 2003, p. 132). Here again, the problem of interpretation was significant. Apparently, the observers had a consistent enough understanding of what it meant for activities to be 'investigative' to pass reliability checks, but that understanding was not made explicit in the study. Consequently, it is difficult to judge the degree of correspondence between the investigative activities the researchers observed and inquiry as most researchers would define it.

Although there are some potential issues in the interpretation of large-scale studies of inquiry enactment just discussed, Weiss et al. (2003) and Marshall et al. (2009) have provided some of the only information to date about enactment of inquiry more broadly. Whereas Marshall et al. estimate very frequent enactment of inquiry using teacher self-report data, Weiss et al. offer a more conservative estimate of inquiry enactment through direct observation. The high frequency of enactment based on teacher self-report is also at odds with the findings from smaller-scale studies, which suggest that use of inquiry may be quite rare. For instance, Capps and Crawford (2013a) analyzed lessons that 26 teachers identified as being most representative of their inquiry teaching. They compared the features of the lessons to the essential features (NRC, 2000). Only four teachers enacted inquiry with a high degree of alignment with *INSES*, and many teachers enacted no inquiry at all. Ozel and Luft (2013) observed lessons of 44 first-year teachers, again comparing lesson features to the essential features of inquiry. They found that 'there were very few instances of inquiry in the classrooms' (p. 313). Teachers who did enact inquiry spent most of their time on questioning and conducting investigations.

One important reason for measuring reported enactment of inquiry is that it represents what teachers think they are doing. This is valuable information even if it does not represent what teachers are actually doing. For instance, the high frequency of reported enactment found in Marshall et al. (2009), when contrasted with the much lower frequency of enactment found in observation studies, hints at the possibility that teachers may think they are doing inquiry, when in reality they may not be. Of course, a serious investigation of this possibility would depend on increasing the reliability and precision of reported enactment measures using more focused instruments. Furthermore, the value of the self-reported enactment measure would be increased based on the extent to which it could be combined with other variables bearing on the quality of the self-reported enactment. This was the reason for combining measures of knowledge with reported enactment in the present study. We reasoned that these measures together could begin to reveal the degree to which teachers who saw themselves as enacting inquiry might actually be doing so.

## Conceptual framework

As teacher knowledge is generally related to practice (Abell, 2007), a fundamental assumption of our study was that teachers would need sound inquiry knowledge to enact inquiry in the classroom (though knowledge alone would be insufficient). Thus, self-reported enactment would be questionable if it were based on limited knowledge of inquiry. This assumption is based on research showing how inquiry requires sophisticated pedagogical knowledge (Crawford, 2000; Keys & Bryan, 2001), making it unlikely that teachers with vague knowledge of inquiry could teach science in this way. Further, there is at least some evidence that well-structured knowledge of inquiry predicts successful enactment (Capps & Crawford, 2013a), making it unlikely that teachers with low inquiry knowledge could enact it well. This point is reinforced by the consistent finding that knowledge, in general, is essential for the enactment of reform-based teaching practices (Garet, Porter, Desimone, Birman, & Yoon, 2001; Penuel, Fishman, Yamaguchi, & Gallagher, 2007). Thus, having more detailed information regarding what teachers know about inquiry and how their knowledge relates to reported enactment would be useful for untangling the inconsistencies about frequencies of enactment revealed in the literature.

In order to check reported enactment against knowledge, the present study was designed to compare two primary variables. The first was how frequently teachers reported enacting inquiry practices. The second was the degree of structure in teachers' knowledge of the features of inquiry. Underlying both variables was a standard definition of inquiry which we adapted from U.S. national documents.

### *Standard definition of inquiry*

At the time of the study, two documents describing inquiry were in broad circulation in the U.S., *INSES* (NRC, 2000) and the just-released *Framework* (NRC, 2012). Both defined inquiry through essential features or practices; however, the newer document was more detailed in that it described several inquiry practices that were not emphasized in the previous reform. This situation challenged us to conceptualize knowledge and enactment of inquiry in a way that would be consistent with *INSES*, but not preclude conceptions of inquiry from the more detailed formulation of the *Framework*. To do this, we created the seven-dimensional definition of inquiry described in Table 1 by combining conceptions of inquiry from *INSES* and the *Framework*. Four of the dimensions, questioning, interpreting data, explaining evidence, and communicating, were essential features of inquiry described in *INSES*. These four also corresponded to four of the scientific practices described in the *Framework*. We created a fifth dimension, investigating, by drawing on one of the abilities to do inquiry from *INSES* and one of the scientific practices from the *Framework*. A sixth dimension, argumentation, was created by differentiating this activity from explanation, as the *Framework* does, in contrast to *INSES,* which tends to refer to explanation and argumentation together. Similarly, we added modeling as a separate dimension, as this activity is prominent in the *Framework*, which points out that it is an important investigative practice in science. It is worth noting that the three dimensions we added from the *Framework* are represented in *INSES* under abilities to do inquiry, though they are not as prominent as they are in the *Framework* (this is especially true for argumentation and modeling). Due to instrumentation space limitations, we did not include using mathematics as a dimension of inquiry, although this dimension is present in both *INSES* and the *Framework*.

**Table 1.** Measurement framework for the seven dimensions of inquiry included in our standard definition.

| Standard definition of inquiry | Label | Essential feature (NRC, 2000) | Abilities to do inquiry (NRC, 2000) | Scientific practices (NRC, 2012) |
|---|---|---|---|---|
| 1. Pursuit of answers to scientific **questions** | Questioning | X | | X |
| 2. Planning, designing, or conducting an **investigation** | Investigating | | X | X |
| 3. Analyzing, evaluating, or **interpreting data** | Interpreting data | X | | X |
| 4. Explaining phenomena from observation or **evidence** | Explaining evidence | X | | X |
| 5. **Using evidence** to advance a claim or conclusion | Argumentation | | X | X |
| 6. **Communicating** scientific information in written or spoken form | Communicating | X | | X |
| 7. Constructing and using **Models** | Modeling | | X | X |

Note: Our standard definition of inquiry was derived from aspects of inquiry described in *INSES* (NRC, 2000) and the *Framework* and *NGSS* (NRC, 2012; NGSS Lead States, 2013). Numbers in the left-hand column are used to highlight each of the seven dimensions in our standard definition. An 'X' in a column indicates the use of the document to derive the dimension of inquiry for our measurement framework. The relationships between our measurement framework, *INSES*, and *Framework* and *NGSS* do not imply that the documents define the different aspects of inquiry identically. Boldface text emphasizes the core of the dimension statement.

### *Teachers' knowledge of inquiry*

Figure 1 shows how we used our seven-dimensional standard definition of inquiry to create a benchmark structure against which to measure teachers' knowledge. It shows three levels of knowledge based on our review of the literature which showed how teachers tended to describe inquiry in terms of certain salient features and not in terms of other less-salient features. The bottom level contains no dimensions of inquiry. The middle level has two dimensions, questioning and investigating; these were the salient dimensions of inquiry which the literature shows to be prominent in teachers' thinking. The highest level contains the remaining five dimensions of inquiry, or the activities that, according to the literature, tend to be less salient in teachers' thinking. In our measurement instrument, we asked participants to describe what inquiry-based science teaching was. We reasoned that teachers might describe no dimensions of inquiry, only the two salient dimensions, or they could go beyond this to also describe some of the less-salient dimensions. Describing inquiry using no dimensions would represent vague knowledge. Describing it in terms of the salient dimensions would represent less-structured knowledge. Describing inquiry in terms of the less-salient dimensions would represent more-structured knowledge.

Our framework for measuring inquiry knowledge was a limited one. It was restricted to declarative knowledge, neglecting procedural and strategic knowledge needed to enact inquiry. Therefore, it should not be interpreted as a comprehensive measure of what teachers may know or need to know about inquiry. On the other hand, as knowledge is thought to be related to practice, it stands to reason that well-differentiated knowledge of inquiry practices is important—and probably necessary—for enacting inquiry. Teachers who lack a well-differentiated knowledge of inquiry will likely struggle to faithfully enact inquiry in the classroom. Thus, it is our argument that we measured an important aspect of inquiry knowledge, adequate for our purpose of providing a check on teachers' reported enactment.

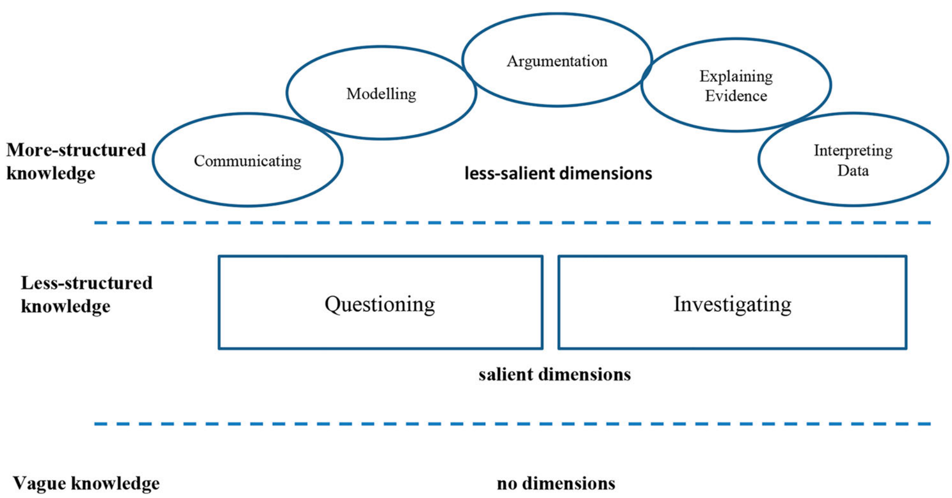


**Figure 1.** Conceptual framing for teachers' knowledge of inquiry used in this study showing the dimensions of inquiry that relate to vague knowledge, less-structured knowledge, and more-structured knowledge.

### *Reported enactment of inquiry*

We defined reported enactment of inquiry as the frequency with which teachers indicated engaging their students in each of the seven dimensions of inquiry listed in the first column of Table 1. Again, our goal was to be consistent with definitions in *INSES* (NRC, 2000), but also to encompass conceptions of inquiry from the more detailed structure of the *Framework* (NRC, 2012). The actual measure listed particular inquiry-related activities (from the *Framework*), and teachers reported how often they enacted these in their teaching. We used the *Framework's* conception of inquiry practices because its goal statements were well formed and accessible. This approach improved upon measures found in existing reported enactment studies (e.g. Banilower et al., 2007; Marshall et al., 2009) in that it focused specifically on inquiry practices and included many more items, providing a greater degree of detail and structure within reported enactment.

## *Purpose*

Given the existence of larger scale studies showing a high frequency of self-reported enactment of investigative teaching (e.g. Beatty & Woolnough, 1982; Marshall et al., 2009), we expected the frequency of reported enactment of inquiry to be high among the teachers in our study. Anticipating this result, our purpose was to see if teachers reporting enactment at high rates also tended to have well-structured knowledge of inquiry. We had no clear expectation of what we would find. As the preceding literature review demonstrated, qualitative studies have shown how teachers can struggle to understand the key aspects of inquiry. If this struggle were also prevalent in the general population of teachers, then faithful enactment of inquiry would be rare, as knowledge of inquiry is likely necessary for faithful enactment to occur. However, there are no studies of representative samples showing what teachers know about inquiry. Assuming a high rate of reported enactment of inquiry, there were two possibilities for knowledge of inquiry that would be valuable to know about: (1) more-structured knowledge and (2) less-structured knowledge. More-structured knowledge, paired with high frequency of reported enactment, would reinforce the validity of reported enactment as a measure of teachers' proficiency with inquiry. It would also be a positive sign[1] with respect to teacher preparedness to enact further inquiry-related reforms, at least among select teachers like those in our study. Less-structured knowledge, paired with high frequency of reported enactment, would indicate a lower proficiency with inquiry, and it would limit the interpretation of reported enactment measures of inquiry to what teachers say they are doing (or perhaps think they are doing), but not what they are truly doing. Correspondingly, this result would indicate a lower level of preparedness to enact further inquiry-related reform than results of reported enactment studies would suggest.

## Method

### *Design*

The study was organized around a survey which measured teachers' reported enactment and knowledge of inquiry. We supplemented the survey with face-to-face interviews which provided information on the ways reported enactment could be either aligned or misaligned with reform-based conceptions of inquiry. We administered the survey and conducted the interviews at a national meeting of the National Science Teachers Association (NSTA). We chose this venue as it attracts thousands of teachers from all grades and times in their careers from across the U.S., providing a diverse population from which to sample. We also regarded attendance at the conference as a sign that teachers were committed to learning about reformed teaching practices and reasoned that this self-selected group of teachers would have high levels of enthusiasm for and exposure to inquiry. Therefore, the sampling would be conservative with respect to whether findings of limited inquiry knowledge should generalize to the broader population of U.S. teachers.

## Instrumentation

The survey was administered on paper and consisted of forced-choice items for reported enactment and a single free-response item about teachers' knowledge of inquiry. In constructing the survey, we drew on recommendations for survey design by Bradburn, Sudman, and Wansink (2004). For instance, all forced-choice items used fully labeled scales, five points for unipolar constructs and seven for bi-polar constructs. In addition to reported enactment and knowledge items, the survey contained items that were not used in the present study. These included items on background information of participants (e.g. years of teaching) and a group of items intended to measure teachers' beliefs about inquiry (e.g. the degree to which teachers felt lack of student motivation was a challenge to implementing inquiry). The results of these measures are not reported in the present study but are fully addressed in Young (2013). The complete survey is available in supplementary online materials (Table S1).

### Reported enactment of inquiry

The reported enactment items on the survey stated activities representing each of the seven dimensions of inquiry in Table 1 and asked teachers how often they did these activities. There were 7 statements for each of the 7 dimensions, for a total of 21 enactment statements. The statements are shown in the first column of Table 2. Having multiple enactment statements for each dimension enabled us to use principal components analysis to check whether teachers interpreted the statements as intended. For each statement, participants were asked to rate how often they had students engage in the activity using a seven-point Likert scale from 1, never, to 7, during every class. All 21 enactment statements were taken nearly verbatim from the goal statements for scientific practices from the *Framework*. In selecting enactment statements from those available in the *Framework*, we avoided statements that exceeded the boundaries of our standard definition of inquiry (e.g. we omitted statements related to engineering practices).

### Knowledge of inquiry

We measured the degree of structure in teachers' knowledge of inquiry with a single free-response item which asked them to describe the most important aspects of inquiry to a non-professional audience. The wording was, 'If you had to tell a group of parents, at an open-house night, what are the most important aspects of inquiry-based science

**Table 2.** Principal components analysis of the 21 inquiry enactment statements.

| | Description | Factor 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Questioning | Ask questions about the natural and human-built worlds | .229 | −.037 | .039 | .865 | .052 |
| | Formulate and/or refine questions that can be asked empirically in a science classroom | .117 | .385 | .247 | .700 | .058 |
| | Ask questions about features, patterns, or contradictions noted in data sets | .257 | .494 | .184 | .560 | .149 |
| Investigating | Decide what data are to be gathered, what tools are needed to do the gathering, and how measurements will be recorded | .155 | .798 | .215 | .101 | .144 |
| | Decide how much data are needed to produce reliable measurements and consider any limitations on the precision of the data | .223 | .850 | .156 | .084 | .110 |
| | Plan experimental or field-research procedures, identifying relevant independent and dependent variables and, when appropriate, the need for controls | .308 | .719 | .220 | .002 | .151 |
| Interp. data | Analyze data systematically, either to look for patterns or to test whether the data are consistent with an original hypothesis | .411 | .723 | .130 | .185 | −.028 |
| | Use spreadsheets, data bases, tables, charts, graphs, statistics, and mathematics to collate, summarize, and display data and to explore relationships between variables | .282 | .740 | .014 | .171 | .141 |
| | Evaluate the strength of a conclusion that can be inferred from any data set, using appropriate grade-level mathematics and statistical techniques | .409 | .749 | .056 | .102 | .136 |
| Exp. evidence | Construct their own explanations of phenomena using their knowledge of accepted scientific theory and linking it to models and evidence | .685 | .364 | .089 | .260 | .099 |
| | Use scientific models and evidence to support or refute an explanatory account of a phenomenon | .739 | .316 | .187 | .194 | .087 |
| | Identify gaps or weaknesses in explanatory accounts | .786 | .194 | .262 | .116 | .152 |
| Argumentation | Construct a scientific argument showing how the data support a claim | .786 | .332 | .157 | .118 | .125 |
| | Identify possible weaknesses in scientific arguments, appropriate to the students' level of knowledge, and discuss them using reasoning and evidence | .825 | .273 | .192 | .117 | .200 |
| | Recognize that the major features of scientific arguments are claims, data and reasons and distinguish these elements in examples | .816 | .292 | .205 | .112 | .091 |
| Communicating | Use words, tables, diagrams, and graphs to communicate their understanding or to ask questions about a system under study | .352 | .342 | .292 | −.006 | .454 |
| | Read grade-level appropriate scientific text with tables, diagrams, and graphs and explain the ideas being communicated | .069 | .179 | .093 | −.030 | .821 |
| | Produce written and illustrated text or oral presentations that communicate their own ideas and accomplishments | .202 | .061 | .032 | .211 | .822 |
| Modeling | Construct drawings or diagrams as representations of events or systems (e.g. to represent what happens to the water in a puddle as it is warmed by the sun) | .097 | .117 | .577 | .419 | .387 |
| | Represent and explain phenomena with multiple types of models (e.g. represent molecules with bond diagrams or 3D models) | .311 | .193 | .825 | .066 | .110 |
| | Discuss the limitations and precision of a model | .388 | .235 | .750 | .175 | .001 |
| % Variance | | 47.69 | 8.04 | 7.05 | 6.39 | 4.94 |
| Eigenvalue | | 10.01 | 1.69 | 1.48 | 1.34 | 1.04 |

Note: Light gray bars in the left-hand column distinguish between each of the seven dimensions of inquiry in our measurement framework. Each dimension is represented by three goal statements from the *Framework*. Dark gray bars indicate that Investigating and Interpreting Data, and Explanation of Evidence and Argumentation each loaded on the same factor, whereas the other three dimensions loaded as single factors.

teaching, what would you tell them?' We wrote the item in this way to elicit the ideas that teachers held about inquiry practices, expressed in their own words.

### Interview protocol

The interview asked teachers to describe what it might look like to carry out particular inquiry practices in their classrooms. The interaction between interviewer and participant was meant to be semi-structured (Wengraf, 2001) so that we could prompt for further information depending on teachers' initial responses.

## Context and participants

The study was conducted at an NSTA national conference held at a large city in the central-continent area of the U.S. We reasoned that teachers attending the national conference would be at least as knowledgeable about inquiry as teachers in the broader population, and probably more so. In general, the NSTA national conference provides a venue for science educators from across the U.S. to connect with one another and share their experiences as well as learn new science content and teaching strategies. Conference attendees are part of a network of science teachers who receive information from various list serves, professional journals, and other information from the NSTA, some of which discuss inquiry teaching. Also, NSTA conferences tend to provide many sessions related to inquiry and inquiry teaching practices.

In total, 149 teachers (approximately 2% of conference attendees) completed the survey, and 11 teachers were interviewed (7% of the sampled population). Approximately 72% of the participants were female, and 28% were male. These percentages reflect gender distribution in the population of teachers in the U.S. as a whole (United States Department of Education, 2012). There was a greater percentage of middle and high school teachers, 84%, than elementary teachers, 16%; and teaching experience ranged from 1 to 43 years, with an average of 13.3 years. All but four teachers provided the zip code where they taught, showing that participants represented 29 states and 1 U.S. Territory. Not surprisingly, the largest number of study participants came from the state in which the conference took place (29 participants). Of those from other states, 61 came from the central-continent region, followed by the southeast, 20, northeast, 15, and other regions, 21. We used the U.S. Department of Agriculture's 2013 Rural-Urban Continuum Codes to estimate the population density where the respondents taught. Most teachers (84%) worked in metropolitan areas. This proportion is representative of the U.S. population distribution (United Nations Department of Economic and Social Affairs, 2011).

## Procedures

Before the study began, we piloted an early version of the survey with 21 K-12 teachers. There were roughly equal numbers of teachers from elementary, middle, and high schools. The piloting teachers took the survey and wrote comments on items they found difficult to interpret. Additionally, we conducted interviews with seven pilot teachers to obtain more detailed information about their interpretation of the items.

For the actual study, we administered the survey from two booths in the exhibition hall of the conference. We used two locations to prevent surveying only those teachers who were interested in particular kinds of exhibits. One booth was in the central thoroughfare and the other was in a side aisle. We also solicited participation as we walked through the exhibition hall. To solicit participation, we introduced ourselves as education researchers and asked if the attendee would be willing to fill out a 15-minute survey about their teaching. Teachers who agreed to participate were shown to a nearby table where they filled out the paper survey. Those who completed the survey were entered in a drawing to win one of four $25 gift cards.

The interview procedure began by looking to see if teachers checked a box on the back of their completed survey stating that they would be willing to participate in a five-minute interview about topics related to their responses. If the box was checked affirmatively, we verbally expressed our wish to interview the teacher and asked permission to audio-record the interview. Then, a researcher sat down with the teacher and conducted the interview based on the teacher's response to self-reported enactment items on the survey. As part of this process the interviewer chose one or two statements that the teacher reported enacting frequently. The interviewer then asked the teacher to describe what it might look like to carry out the selected enactment statements in the classroom.

## Data analysis

### Reported enactment of inquiry

A preliminary step in the analysis of teachers' reported enactment of inquiry was to conduct a principal components analysis to check whether the 21 enactment statements grouped as intended into seven triplicates corresponding to the seven dimensions of inquiry in our measurement framework. The analysis yielded five factors with eigenvalues greater than one instead of seven (see Table 2). Three of these factors mapped directly to the dimensions of inquiry we expected. These were questioning, communicating, and modeling. The remaining two factors included two dimensions each. One of the pairs combined investigating with interpreting data. We think it likely that this pair grouped together because data analysis is naturally coupled with investigation, so those teachers who reported engaging their students in investigations also reported having students work with data. The other pair was for explaining evidence and argumentation. In this case, we think it likely that teachers did not differentiate between using evidence for explanation and using evidence to advance an argument. This was not surprising given the fact that theory distinguishing between argumentation and explanation is relatively new in science education (Osborne & Patterson, 2011). Whatever the cause of these groupings, in all further analysis and description of reported enactment, we report on investigating/interpreting data and argumentation/explaining evidence together. We calculated the frequency of enactment for each dimension by taking the mean of the frequencies for each of the items the dimension contained. There were six items for each of the two combined dimensions and three items each for the non-combined dimension. We also calculated a single index for the overall frequency of reported enactment for the purposes of comparing to teachers' knowledge. To create this index, we summed the number of dimensions of inquiry that each teacher reported enacting at or above a 'high-frequency' threshold of 2–3 times a month. We selected 2–3 times a month as the threshold based on the rationale that if three or more different dimensions of inquiry were enacted at this frequency, then the total amount of enactment would approach twice a week, a fairly high frequency.

### *Knowledge of inquiry*

We used two separate analyses to transform teachers' written descriptions of the most important features of inquiry teaching into a quantitative measure of their knowledge of inquiry. One approach was predetermined through its direct linkage to our measurement framework. The other was an inductive analysis that was independent of our framework. The purpose of the inductive analysis was to capture knowledge that was not anticipated by the framework.

The first step in the predetermined analysis was to code whether or not a teacher's response mentioned any of the seven dimensions in the measurement framework (see Table 1). The criteria for mentioning each dimension are provided in Table 3, along with example responses. As the examples show, teachers did not have to exhibit detailed knowledge of a dimension in order to be credited for mentioning it. This liberal interpretation of teachers' responses enabled us to be conservative in reporting what was not in teachers' knowledge structures (i.e. to have especially firm measurements for what teachers did not know).

The second step of the predetermined analysis was to group teachers' responses into three levels consistent with our conceptual framework's tiered structure of inquiry knowledge. The levels were: (1) vague knowledge, where no dimensions of inquiry were described; (2) less-structured knowledge, where only one or two of the salient dimensions (i.e.

**Table 3.** Criteria used to measure whether teachers' responses included dimensions of inquiry in our standard definition along with an example for each dimension.

| Standard definition of inquiry | Needed to be mentioned | Example |
|---|---|---|
| Pursuit of answers to scientific questions | Asking questions | *Students do science.* **Students find solutions to teacher driven problems/questions**. |
| Planning, designing, or conducting an investigation | Conducting an investigation | *Inquiry is where students ask questions and* **develop an experiment** *based on the questions. They will be doing science, not just learning it. Your kids will experience science in a new way* |
| Analyzing, evaluating, or interpreting data | Analyzing data | *Students do more critical thinking, less memorization, collect and* **analyse data**. *Assessment is not like traditional assessment they are familiar with* |
| Explaining phenomena from observation or evidence | Forming explanations | *Observe critically the situation at hand. As yourself what is happening/what I am seeing. Think critically when* **building an explanation for the phenomena/ situation** *using the evidence at hand.* |
| Using evidence to advance a claim or conclusion | Making evidence-based claims | *Science as a way of knowing:* **reasoning**, *asking questions, testing questions, analysis of data,* **conclusions based in evidence**. |
| Communicating scientific information in written or spoken form | Communicating about their inquiry | *Inquiry-based science teaching allows students to learn about science in a safe and structured environment. As we progress through a unit, students will be given the tools to investigate a topic and will be built upon in time. The goal in the end is to have students think like scientists. They will be able to question scientifically and develop experiments that they design. And most importantly, they will be able to* **explain what they learned and why it is important (to me, their classmates, and themselves**). |
| Models or modeling | Using models | *Hands on and teacher is the facilitator. Students learn to think through a problem.* **A lot of modelling**. |

Note: The entire written answer is copied verbatim. Boldface type indicates the part of the response that met the acceptance criterion.

questioning and/or investigating) were described; and (3) more-structured knowledge, where any of the less-salient dimensions of inquiry was described. In using this tiered structure, we assumed that the two dimensions in the middle level would be more prominent in teachers' thinking than those in the higher level. Moreover, we assumed that teachers would not describe higher level dimensions without describing the lower level dimensions. We used the knowledge of inquiry measure to check and report on the validity of this assumption. Notably, teachers only needed to mention one of the dimensions of inquiry to reach a given level (e.g. mentioning one of the five less-salient dimensions of inquiry was sufficient to reach the more-structured level). Once again, the point was to have a conservative measure of what teachers did not include in their knowledge of inquiry.

For the non-predetermined analysis, we developed categories inductively, and independently from the measurement framework, based on terms, phrases, and ideas teachers used to describe inquiry. To conduct the analysis, the third author examined a subset of teacher responses to identify and define 'candidate' themes. The research team then reviewed these themes, combining those that overlapped, and refining their definitions. Through this process, we developed 12 distinct themes. We then used the themes to code all of the survey responses for the presence or absence of each one. The categories and their frequencies are described in the Results section. Two researchers independently coded all 149 responses using both the predetermined and inductive coding schemes. The coders were the third author and another researcher not closely associated with the project. Agreement between coders was 90% or greater for all themes, calculated as the number of responses agreed upon divided by the total number of responses. We took this rate of agreement to be sufficient evidence of the reliability of the coding.

### Interviews about reported enactment

To analyze the interviews, we established criteria for the validity of self-report by judging alignment between the activities teachers described in their interviews with the reported enactment statements from the survey. To prepare for the analysis, we took each statement on the survey that surfaced in an interview and defined what we considered to be the minimum level of enactment to be aligned with that statement. Then, after transcribing the interviews, we coded each described activity for whether or not it met the minimum enactment criterion. Statements who did not meet this criterion were judged as misaligned. Table 4 shows an example of an aligned and misaligned response. Agreement was reached by consensus. Further examples are provided in the Results section.

## Results

As expected, conference attendees reported enacting the various dimensions of inquiry quite frequently. However, their knowledge of the dimensions of inquiry was not well structured. In what follows we present details of these findings and how they relate to one another.

### Reported enactment of inquiry

Table 5 shows the percentage of teachers who reported enacting dimensions of inquiry at or above the high-frequency threshold of 2–3 times a month. The percentages in this table

**Table 4.** Criteria used to judge the alignment between the activities described in interviews compared to the reported enactment section of the survey.

| Statement with minimum criteria underlined | Minimum enactment | Example response | Code | Rationale |
|---|---|---|---|---|
| Students represent and explain phenomena with multiple types of models | Using multiple representations or models. May or may not be student generated | For wave propagation, we do regular spring waves and all that, but then I have them physically model waves, everything from the football kinda football stadium wave to uh to do compressional waves | Aligned | Teacher has students use multiple models to represent phenomena |
| Students construct their own explanations of phenomena using their knowledge of accepted scientific theory and linking it to models and evidence | Students construct or are supported in constructing scientific explanations drawing on their knowledge of science | In the classroom they have a writing prompt that they do every day in their science notebook. Sometimes it's about what we've been studying and sometimes it's more anticipatory. I ask them to answer the question regardless of whether or not they feel qualified or know the right answer | Misaligned | A daily writing prompt is not indicative of constructing explanations and linking them to science knowledge |

Note: The table shows two examples. The top example aligns with the minimum enactment criterion and the bottom example does not. A rationale is provided with each example.

**Table 5.** The numbers of dimensions of inquiry that teachers enacted at least 2–3 times a month.

| | Number of dimensions[a] | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | Total |
| Percent of teachers | 2.0 | 4.0 | 6.7 | 20.8 | 23.5 | 43.0 | 100 |

[a]The maximum number of dimensions was five, not seven, because two of the dimensions were indistinguishable in the factor analysis. See the Methods section for more information.

are notably skewed toward higher rates of enactment. About two-thirds, or 66.5% of teachers, reported enacting all, or all but one, of the inquiry dimensions at or above the threshold frequency, and just under half, 43%, reported enacting all of the dimensions at this frequency. These figures correspond to a majority of the teachers surveyed reporting engaging their students in at least two different dimensions of inquiry each week, or doing so on 40% of school days, even allowing for some overlap (i.e. situations were a single activity hits upon two different dimensions of inquiry). Thus, our findings roughly agree with those of Marshall et al. (2009) who found that teachers reported engaging students in inquiry an average of 38.7% of the time.

Table 6 shows the percentages of teachers who reported enacting various frequencies for each measured dimension of inquiry. The shaded portions of the table represent frequencies above the high-frequency threshold of 2–3 times a month. The cumulative frequencies show that two of the dimensions, investigating/interpreting data and argumentation/explaining, were reportedly enacted less frequently than the

**Table 6.** Percentages of teachers reporting enactment at each frequency.

| Dimension of inquiry | Never | Less than once a month | Once a month | 2–3 times a month | Once a week | Several times a week | During every class | Cum. lower frequency | Cum. higher frequency |
|---|---|---|---|---|---|---|---|---|---|
| Questioning | 0.7 | 7.5 | 16.8 | 26.8 | 28.8 | 18.1 | 1.3 | 25.0 | 75.0 |
| Investigating/ Interpreting data[a] | 3.4 | 17.0 | 31.8 | 27.9 | 13.9 | 5.3 | 0.7 | 52.2 | 47.8 |
| Arguing/ Explaining[a] | 4.1 | 14.3 | 33.6 | 24.4 | 12.8 | 8.8 | 2.0 | 52.0 | 47.8 |
| Communicating | 0 | 4.7 | 18.1 | 25.5 | 35.5 | 15.5 | 0.7 | 22.8 | 77.2 |
| Modeling | 0 | 10.1 | 21.4 | 26.8 | 28.3 | 11.4 | 2.0 | 31.5 | 68.5 |
| Overall mean | 1.6 | 10.7 | 24.3 | 26.2 | 23.9 | 11.8 | 1.3 | 36.7 | 63.3 |

Note: Shading represents frequencies at or above the high-frequency threshold value of 2–3 times a month.
[a]These dimensions are reported together because they could not be distinguished in the factor analysis. See the Methods section for more information.



others. On average, 47.8% of teachers reported engaging their students in these dimensions at least 2–3 times per month compared to 71.0% of teachers who reported engaging their students in questioning, modeling, and communicating evidence at this frequency.

It was interesting that the data and evidence-centred activities within inquiry were reported less frequently than the other activities, given that data and evidence are central to inquiry (Anderson, 2007; NRC, 1996, 2000). One explanation could be that teachers based their responses in part on activities within questioning, modeling, and communicating that were done outside of inquiry. Taken out of context, several of the enactment statements for questioning, modeling, and communicating could fit activities that have little to do with inquiry. For example, a teacher could have her students ask questions about the natural and human-built worlds, construct drawings or diagrams as representations of events or systems, or produce written and illustrated text or oral presentations that communicate ideas and accomplishments outside of an inquiry approach. By contrast, the statements pertaining to investigating/interpreting data and argumentation/explaining, such as deciding what data are to be gathered and or how measurements will be recorded, would be much more difficult to interpret outside of inquiry. Thus, we might have obtained somewhat lower frequencies of reported enactment had we more explicitly linked questioning, modeling, and communicating to evidence-based understanding of phenomena. Nevertheless, even if the lower rates for the two evidence-based dimensions are taken as more representative, the reported enactment frequency would still be very high.

## *Knowledge of inquiry*

Table 7 compares the number of teachers who mentioned either of the two salient dimensions of inquiry (less-structured knowledge) with the number who mentioned any of the five less-salient dimensions (more-structured knowledge). The upper left cell of the table shows that there was a strikingly high number of teachers, 89, who mentioned no dimensions from either level, suggesting that they had vague knowledge of inquiry. Furthermore, and as we anticipated, teachers mentioned the two salient

**Table 7.** Comparison of the number of teachers who mentioned either of the salient dimensions of inquiry with those who mentioned any of the less-salient dimensions of inquiry.

| Salient | Less salient | | |
|---|---|---|---|
| | No | Yes | Total |
| No | 89 | 4 | 93 |
| Yes | 38 | 18 | 56 |
| Total | 127 | 22 | 149 |

inquiry dimensions, questioning and investigating (indicating less-structured knowledge), more than twice as often as they mentioned the less-salient dimensions (indicating more-structured knowledge). Only four teachers violated the tiered framework by describing less-salient dimensions without mentioning either of the more-salient dimensions. We concluded from this pattern that the tiered framework (see Figure 1) was sufficient to define the level of structure in teachers' knowledge of inquiry. The four teachers who violated the pattern were assumed to have more-structured knowledge and are included in this level for all further analysis.

Table 8 shows the percentage of teachers whose knowledge fell into each level. In line with the frequencies just reported, most teachers, 59.7%, had vague knowledge of inquiry. Roughly a quarter of teachers, 25.5%, had less-structured knowledge, and 14.8% had more-structured knowledge. Among teachers' responses showing more-structured knowledge, the dimensions interpreting data and argumentation were the most commonly mentioned (eight and seven teachers, respectively), while the remaining dimensions were spread among seven teachers, with one teacher mentioning modeling, three mentioning explaining evidence, and three mentioning communicating.

Table 9 shows the themes teachers used to describe inquiry that did not correspond to any of the measured dimensions. In general, the themes were similar across the three groups. For example, teachers in all three levels regularly described inquiry in terms of *student-centered learning*, *learning through exploring or discovering*, and *critical thinking or problem-solving*. Among teachers with some degree of knowledge, we see these themes as potentially representing supplemental knowledge, such as additional ideas about inquiry, or preferences within inquiry such as a preference for engaging students in hands-on science investigations or generating scientific questions through exploration. However, for teachers with vague knowledge of inquiry, the themes seem more likely to represent alternative ideas. Among teachers with vague knowledge, the four most common themes were *student-centered learning*, *learning through exploring or discovering*, *critical thinking or problem-solving*, and *hands-on learning*. Overall, many of these modes of learning are consistent with inquiry, but they could also be done outside of an inquiry orientation, with hands-on learning being a prime example. Indeed, two of the themes, *hands-on learning* and learning through *exploring and discovering*, are commonly viewed as alternative conceptions of inquiry (NRC, 2000).

**Table 8.** Percentage of teachers in each level of knowledge.

| Distribution of teachers' degree of structure | % of teachers |
|---|---|
| Vague knowledge (no dimensions of inquiry) | 59.7 |
| Less-structured knowledge (one or two salient dimensions) | 25.5 |
| More-structured knowledge (included any of the less-salient dimensions) | 14.8 |

**Table 9.** Themes teachers used to describe inquiry arranged by the number and percentage of teachers who used them from each level.

| | Vague | | Less structured | | More structured | |
|---|---|---|---|---|---|---|
| | # of teachers ($N = 89$) | % | # of teachers ($N = 38$) | % | # of teachers ($N = 22$) | % |
| Student-centered learning | 34 | 36.6 | 23 | 60.5 | 9 | 41.0 |
| Exploring-discovering | 26 | 29.2 | 11 | 28.9 | 2 | 10.0 |
| Critical thinking—problem-solving | 24 | 27.0 | 7 | 18.4 | 5 | 22.7 |
| Hands-on learning | 19 | 21.3 | 4 | 10.5 | 1 | 4.5 |
| Models what real scientists do | 13 | 14.6 | 7 | 18.4 | 3 | 13.6 |
| Engagement in science | 12 | 13.5 | 5 | 13.2 | 3 | 13.6 |
| Deeper understanding of science content | 14 | 15.7 | 3 | 7.9 | 1 | 4.5 |
| Relevancy | 11 | 12.4 | 0 | 0 | 1 | 4.5 |
| Preparation for future | 6 | 6.7 | 1 | 2.6 | 1 | 4.5 |
| Okay to get the wrong answer | 4 | 4.5 | 2 | 5.3 | 2 | 10.0 |
| Teamwork | 7 | 7.9 | 0 | 0 | 0 | 0 |
| Knowledge construction | 5 | 5.6 | 1 | 2.6 | 0 | 0 |



### *Interviews about reported enactment*

Post-survey interviews with 11 teachers yielded 17 descriptions of reported enactment. In line with the discrepancy between knowledge and reported enactment in the survey, just over half of the teachers interviewed, and two-thirds of the descriptions, illustrated activities that did not meet our acceptance criteria (i.e. they were misaligned with the enactment statements of inquiry). The remaining six descriptions, made by five teachers, included activities that aligned with the enactment statements.

The interviews, though small in number, provided an additional check on the validity of teachers' knowledge of inquiry scores on the survey. Although we could not confirm all three levels of knowledge proposed in the conceptual framework, the interviews did provide evidence that teachers with vague knowledge of inquiry misinterpreted inquiry enactment statements when asked about them in the interview, and only those teachers with more-structured knowledge of inquiry (those that included dimension(s) of inquiry in their response to item 7) were able to accurately interpret the enactment statements. Specifically, we found that four of the six teachers whose interview responses were misaligned with inquiry enactment statements had vague inquiry knowledge as measured by item 7. The other two teachers with misaligned responses mentioned only salient dimensions of inquiry in the survey. All five of the teachers who described activities aligning with the enactment statements mentioned either salient or non-salient dimensions of inquiry in their response to item 7.

To demonstrate the character of teachers' responses, we provide three examples. One is aligned, and two are misaligned with the enactment statements. We begin with the aligned case, where a teacher described having her students construct explanations using their knowledge of science.

*Interviewer:* When you say that you have students construct their own explanations of phenomena using their knowledge of accepted scientific theory and link it to evidence, what might this look like in your classroom?

*Teacher 1:* We often go through the process of the scientific method a little bit more than what might be planned out in a typical textbook lesson, and then whenever students have done the experiment and they're finishing the activity, they have to do a concluding part and I have them use the science that they know to explain what happened.

According to this description, the teacher regularly has her students draw on their understanding of scientific knowledge to explain what they observe in experiments. Thus, this description met our criterion for enactment of constructing explanations from science knowledge within inquiry.

The two misaligned activities, presented below, illustrate how teachers could indicate engaging their students in inquiry activities, when, objectively, the activity they described was not inquiry. The first excerpt shows an example of a teacher who described an activity that was misaligned with an enactment statement about modeling. The second excerpt shows an example of a teacher who described an activity that was misaligned with an enactment statement about questioning.

*Interviewer:* Let's see, you said you had your students construct drawings or diagrams as representations of events or systems often. I'm wondering, what might this look like in your classroom? Can you give me examples of what you might do?

*Teacher 2:* I do foldables a lot. I use graphic organizers a lot. I have my kids sketch everything because I sketch. I sketch on my board, I sketch on my white-board, I sketch on my overheads, you know, I sketch all the time. And it helps me remember and organize my information. So I have my kids do the same thing.

*Interviewer:* There's a couple of things here, like ask questions about the natural and human-built worlds and formulate and refine questions that can be asked empirically. How do you have students do these in your class?

*Teacher 3:* A lot of times we will uh, you know, there's some time at the beginning of class or at the end of class and even if it's not on task, so a lot of times students will come up and ask questions about what they did over the weekend. Why did this happen? Why do you think that is? There's no real formal time, it's more of a, just let them do it.

In both instances, the activities described were misaligned with inquiry. Teacher 2 interpreted the enactment statement about modeling to encompass any type of sketch or drawing, instead of those that students generate to represent their understandings of events or systems. Although the type of activity the teacher described may well have included some inquiry representation, his response seemed to be based on sketching or drawing in a general sense. In the second example, Teacher 3 described a scenario that reflected surface-level similarities to questioning; however, she interpreted the enactment statement to refer to students asking her about the nature of scientific phenomena in their everyday lives, instead of engaging her students in pursuing answers to scientifically oriented questions.

### *The relationship between knowledge and reported enactment*

Table 10 shows the number of dimensions of inquiry teachers reported enacting at or above the threshold frequency of 2–3 times per month, broken down by the knowledge level. The two columns on the right of the table (i.e. 4 or 5 dimensions enacted) represent a very high frequency of enactment, since enacting four or more inquiry practices 2–3 times a month would add up to enacting, at minimum, 8–12 different inquiry practices over a four-week period, or approximately two dimensions per week. Comparing the three levels of knowledge for these high-frequency columns shows that teachers at each knowledge level reported similarly high frequencies of inquiry enactment (i.e. 60–70% of teachers from each level reported enacting four or more inquiry practices 2–3 times a month or more). Thus, high-frequency reported enactment had no relationship to knowledge level. Most concerning is the large proportion of teachers in the top right-hand corner of the table (i.e. teachers with vague knowledge who reported high-frequency enactment), showing that nearly 70% of teachers with vague knowledge reported enacting 4 or 5 dimensions at or above the threshold frequency.

## Discussion

Aware of prior research showing high rates of self-reported enactment of inquiry (e.g. Marshall et al., 2009), we set out to investigate if these rates corresponded with sound knowledge of inquiry. We found that most teachers in our study did not have well-structured inquiry knowledge, despite reporting very frequent enactment. Indeed, nearly two-thirds of the participants described inquiry in a way that had no correspondence with normative definitions of inquiry (i.e. they had vague knowledge of inquiry). It was also not uncommon for teachers with vague inquiry knowledge to describe inquiry using potential alternative conceptions such as hands-on learning. Our sample was drawn from a select group of teachers, and we analyzed the data using liberal criteria for knowing the various dimensions of inquiry. Therefore, we think that the vague knowledge we observed may be widespread among U.S. science teachers. If so, and reasoning that well-structured knowledge is necessary for inquiry enactment (Capps & Crawford, 2013a; Garet et al., 2001; Penuel et al., 2007), then the overall level of proficiency with inquiry among U.S. teachers is likely not very high, and it is certainly much lower than self-reported enactment measures alone would suggest.

**Table 10.** Number of teachers who reported enacting inquiry 2–3 times or more a month with teachers in each of the three levels of inquiry knowledge.

| Knowledge of inquiry | Measured dimensions of inquiry enacted 2–3 times per month or more | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | Total |
| Vague (59.7%) | 2 | 3 | 5 | 18 | 19 | 42 | 89 |
| | 2.3% | 3.4% | 5.6% | 20.2% | 21.3% | 47.2% | 100% |
| Less-structured (25.5%) | 1 | 3 | 3 | 8 | 9 | 14 | 38 |
| | 2.6% | 7.9% | 7.9% | 21.1% | 23.7% | 36.8% | 100% |
| More-structured (14.8%) | 0 | 0 | 2 | 5 | 7 | 8 | 22 |
| | 0% | 0% | 9.1% | 22.7% | 31.8% | 36.4% | 100% |
| Total | 3 | 6 | 10 | 31 | 35 | 64 | 149 |
| | 2.0% | 4.0% | 6.7% | 20.8% | 23.5% | 43.0% | 100% |

Evidently, self-reported enactment of inquiry is not a good indicator of what teachers are doing in their classrooms related to inquiry, as limited knowledge of inquiry would make them prone to misreporting the frequency of their inquiry enactment. Nevertheless, reported enactment can be a very useful metric, as a measure of what teachers think they are doing. Moreover, it can be especially revealing when combined with other metrics, as a way of pointing out potential gaps between what teachers think they are doing and what they are actually doing. The present study serves as an example of this. Taken alone, the most informative interpretation of the high rates of reported enactment found here, and in prior studies, would be that teachers think they are doing inquiry quite often. In this study, over 70% of the teachers reported engaging their students in questioning, modeling, and communicating at least 2–3 times a month or more, an encouraging finding if true. However, when this result is combined with the finding that most teachers had vague knowledge of inquiry, and with interviews that demonstrated how teachers can map non-inquiry practices such as questioning and modeling onto inquiry, it becomes more probable that there is a gap between what teachers think they are doing related to inquiry, and what they are actually doing. To be clear, we did not actually observe this gap, nor can we comment on how large it may be. Though the evidence points to its existence, which would be a major obstacle for implementing reforms aimed at transforming teaching practice towards inquiry, notably the *NGSS* in the U.S. (NGSS Lead States, 2013). After all, if teachers already see themselves as meeting the requirements of a new reform (in our study, reporting that they regularly use the inquiry practices described in the new reform), what impetus do they have to change their practice?

An important limitation of our measure of teachers' inquiry knowledge is that it relied on a single omnibus item. With only one item, the possibility of invalid measurement was much higher than if multiple items had been used. Of particular concern is the possibility that, despite our liberal coding scheme, the item was overly difficult, artificially pushing the knowledge measure to low levels. Interviews with a subset of teachers undermined this possibility. They showed that nearly two-thirds of the teachers described activities that were either not inquiry or easily exceeded the bounds of inquiry, thus corroborating the finding that teachers frequently had vague inquiry knowledge. Nevertheless, a multi-item, multi-dimensional instrument would have provided a more reliable and precise measure of inquiry knowledge. Therefore, we would warn against interpreting the percentages of teachers who fell into each knowledge category as being anything like quantitative estimates of proportions in the population. Rather, these percentages should be taken as coarse indications of the relative magnitudes of the different proportions, with the resulting implication that the proportion of teachers with well-structured inquiry knowledge is apt to be small compared to the proportion with vague knowledge.

Our instrument for measuring reported enactment provided much greater structure and definition than measures previously used, making it a more informative measure. However, this instrument, too, could be improved upon. Specifically, it could be expanded to obtain useful information about what teachers see themselves as doing with respect to inquiry. For example, it might be informative for teachers to self-assess the quality of their enactment of various inquiry activities by indicating the extent to which the activities they reported on were enacted in ways that were central to inquiry, as opposed to being more peripheral to it. Again, when paired with other metrics such as knowledge of inquiry, this information could be used to better understand the extent to which self-reported enactment of inquiry accords with visions of inquiry from current reform movements.

Finally, interviews with a subset of survey participants showed how teachers can include non-inquiry activities when reporting on inquiry, as was the case with Teacher 2 and Teacher 3. Apparently, teachers can easily map non-inquiry activities onto inquiry, as was the case for Teacher 3. Anticipating this mapping, teacher educators could design learning opportunities that would emphasize the differentiation of inquiry practices from similar non-inquiry practices, as other scholars have suggested (e.g. Demir & Abell, 2010). For example, Teacher 2 might benefit from learning how to distinguish when representations (like sketches and drawings) would serve as scientific models, and when they would not. Similarly, Teacher 3 could contrast examples, or perhaps even gradations, of questioning that are more or less inquiry-oriented and be supported in identifying the key features of questioning that relate to inquiry. Teachers with vague knowledge of inquiry who see it as hands-on learning or learning through undefined exploring and discovering activities could also learn about inquiry through differentiation. In this case, structured contrasts could show how approaches like hands-on learning can be done with and without an inquiry orientation.

## Conclusion

The present study provides grounds for reflecting on what has become an enduring question in science education: Why is the concept of inquiry, and more broadly investigative teaching, so elusive for teachers? One answer may be that these approaches are not defined in a user-friendly way. Following other researchers (e.g. Capps & Crawford, 2013a, 2013b; Demir & Abell, 2010; Kang et al., 2008; Ozel & Luft, 2013), our study measured what teachers knew of inquiry using a standard definition advanced in recent reform documents (NRC, 2000, 2012). These definitions are principally focused on specifying the different elements or features of inquiry (i.e. five essential features or eight practices). As such, they are fundamentally lists of information with little organizational structure. The conceptual framework for the present study (see Figure 1) suggests an alternative with greater structure. It proposes that there are two salient dimensions within inquiry, questioning and investigating, with various supporting practices, for instance interpreting data and modeling. Similarly, in addition to defining its eight scientific practices, the most recent guiding document for inquiry in the U.S., the *Framework* (NRC, 2012), conceptualizes the activity of scientists and engineers as being three spheres, investigating, evaluating, and developing explanations and solutions. Setting aside the question of whether these structures represent true or correct conceptions of inquiry or practices of science/engineering, what they both provide is a greater degree of conceptual organization than is found in lists of essential features or practices. As is well understood in educational psychology, information that is organized is easier to learn (Durso & Coggins, 1991; Slavin, 2003). Thus, it may be equally or more important for U.S. teacher educators to emphasize the holistic conception of scientific practice found in the preliminary pages of the *Framework* than the list version that follows. More fundamentally, we argue, it is the job of teacher educators to make students aware of different ways in which inquiry can be conceptually organized, supporting them in building their own rich frameworks for understanding inquiry as an approach to teaching, and not rely on lists of features or practices.

## Note

1. Positive interpretations would be limited by the fact that we only measured declarative knowledge of inquiry, and so did not include the procedural and strategic knowledge that would also be needed for successful enactment (see the conceptual framework). This did not turn out to be an issue.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Notes on contributors

***Daniel K. Capps*** is an Assistant Professor of Science Education in the Department of Mathematics and Science Education at the University of Georgia in the USA. His research focuses on teacher knowledge and practice related to inquiry-based instruction.

***Jonathan T. Shemwell*** is an Assistant Professor of Science Education and Cooperating Assistant Professor of Physics at the University of Maine in the USA. His research interests focus on cognitive representations, scientific argumentation, and alternative modes of science learning.

***Ashley M. Young*** is a high school science teacher at the Innovation Academy Charter School in Tyngsboro, Massachusetts, USA.

## References


Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R. Hofstein, A., … Tuan, H. (2004). Inquiry in science education: International perspectives. *Science Education*, *88*(3), 397–419.

Abell, S. K. (2007). Research on science teacher knowledge. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 1105–1149). Mahwah, NJ: Lawrence Erlbaum.

Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, *30*(14), 1945–1969.

Abrams, E., Southerland, S. A., & Silva, P. (2008). Inquiry in the classroom: Identifying necessary components of a useful definition. In Abrams, E., Southerland, S. A., & Silva, P. (Eds.), *Inquiry in the classroom: Realities and opportunities* (pp. xi–xiii). Charlotte: Information Age Publishing.

American Association for the Advancement of Sciences. (1989). *Project 2061: Science for all Americans*. New York: Oxford University Press.

Anderson, R. D. (2007). Inquiry as an organizing theme for science curricula. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 807–830). Mahwah, NJ: Lawrence Erlbaum.

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the national science foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching*, *44*(3), 375–395.

Beatty, J. W., & Woolnough, B. E. (1982). Practical work in 11–13 science: The context, type and aims of current practice. *British Educational Research Journal*, *8*(1), 23–30.

Blanchard, M. R., Southerland, S. A., & Granger, E. M. (2009). No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers. *Science Education*, *93*(2), 322–360.

Bradburn, N. M., Sudman, S., & Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design–for market research, political polls, and social and health questionnaires*. San Francisco, CA: John Wiley & Sons.

Brown, P. L., Abell, S. K., Demir, A., & Schmidt, F. J. (2006). College science teachers' views of classroom inquiry. *Science Education*, *90*(5), 784–802.
Bybee, R. W. (2000). Teaching science as inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 20–46). Washington, DC: American Association for the Advancement of Science.
Capps, D. K., & Crawford, B. A. (2013a). Inquiry-based instruction and teaching about nature of science: Are they happening? *Journal of Science Teacher Education*, *24*(3), 497–526.
Capps, D. K., & Crawford, B. A. (2013b). Inquiry-based professional development: What does it take to support teachers in learning about inquiry and nature of science? *International Journal of Science Education*, *35*(12), 1947–1978.
Crawford, B. A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching*, *37*(9), 916–937.
Crawford, B. A. (2014). From inquiry to scientific practices in the science classroom. In N. Lederman & S. Abell (Eds.), *Handbook of research on science education* (Vol. 2, pp. 515–544). New York: Routledge.
DeBoer, G. E. (1991). *A history of ideas in science education*. New York: Teachers College Press.
Demir, A., & Abell, S. K. (2010). Views of inquiry: Mismatches between views of science education faculty and students of an alternative certification program. *Journal of Research in Science Teaching*, *47*(6), 716–741.
Department for Education and Skills/Qualification and Curriculum Authority. (2004). *Science – The national curriculum for England*. London: HMSO.
Dillon, J. (2008). *A review of the research on practical work in school science* (pp. 1–9). London: King's College.
Durso, F. T., & Coggins, K. A. (1991). Organized instruction for the improvement of word knowledge skills. *Journal of Educational Psychology*, *83*(1), 108–112.
Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945.
Ireland, J. E., Watters, J. J., Brownlee, J., & Lupton, M. (2012). Elementary teacher's conceptions of inquiry teaching: Messages for teacher development. *Journal of Science Teacher Education*, *23*(2), 159–175.
Jiang, F., & McComas, W. F. (2015). The effects of inquiry teaching on student science achievement and attitudes: Evidence from Propensity Score Analysis of PISA Data. *International Journal of Science Education, 37*(3), 554–576.
Kang, N. H., Orgill, M., & Crippen, K. J. (2008). Understanding teachers' conceptions of classroom inquiry with a teaching scenario survey instrument. *Journal of Science Teacher Education*, *19*(4), 337–354.
Kennedy, M. (1998). *Form and substance in in-service teacher education*. Arlington, VA: National Science Foundation.
Keys, C. W., & Bryan, L. A. (2001). Co-constructing inquiry-based science with teachers: Essential research for lasting reform. *Journal of Research in Science Teaching*, *38*(6), 631–645.
Lotter, C., Harwood, W. S., & Bonner, J. J. (2006). Overcoming a learning bottleneck: Inquiry professional development for secondary science teachers. *Journal of Science Teacher Education*, *17*(3), 185–216.
Marshall, J. C., Horton, R., Igo, B. L., & Switzer, D. M. (2009). K-12 science and mathematics teachers' beliefs about and use of inquiry in the classroom. *International Journal of Science and Mathematics Education*, *7*(3), 575–596.
Messick, S., & Jackson, D. N. (1961). Acquiescence and the factorial interpretation of the MMPI. *Psychological Bulletin*, *58*(4), 299–304.
Ministry of Education of Singapore. (2007). *Primary science syllabus*. Singapore: Author.
National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
National Research Council. (2000). *Inquiry and the National Science Education Standards*. Washington, DC: National Academy Press.

National Research Council. (2012). *A Framework for K-12 science education: Practices crosscutting concepts, and core ideas*. Washington, DC: National Academy Press.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

Osborne, J. F., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education*, *95*(4), 627–638.

Ozel, M., & Luft, J. A. (2013). Beginning secondary science teachers' conceptualization and enactment of inquiry based instruction. *School Science and Mathematics*, *113*(6), 308–316.

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, *44*(4), 921–958.

Slavin, R. (2003). *Educational psychology: Theory and practice* (7th ed.). Boston: Allyn & Bacon.

*Tomorrow 98: Report of the superior committee on science mathematics and technology in Israel*. (1992/1994). Jerusalem: Ministry of Education and Culture.

Tytler, R. (2007). *Re-Imagining science education: Engaging students in science for Australia's future*. Camberwell: Australian Council for Educational Research (ACER).

United Nations Department of Economic and Social Affairs. (2011). *World urbanization prospects*, the 2011 Revision [Data file]. Retrieved from http://esa.un.org/unup/CD-ROM/Urban-Rural-Population.htm

United States Department of Education. (2012). *Teacher trends*. Retrieved from http://nces.ed.gov/fastfacts/display.asp?id=28

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Looking inside the classroom: A study of K–12 mathematics and science education in the United States*. Chapel Hill, NC: Horizon Research.

Wengraf, T. (2001). *Qualitative research interviewing: Biographic narratives and semi-structured methods*. Thousand Oaks, CA: Sage.

Young, A. M. (2013). *Teachers' understanding of inquiry and reported use of scientific practices: A survey of NSTA conference attendees* (Unpublished master's degree thesis). University of Maine, Maine.