Fostering Data Literacy through Preservice Teacher Inquiry in English Language Arts

Steven Z. Athanases
Lisa H. Bennett
Juliet Michelsen Wahleithner

University of California, Davis

This is the pre-publication version of the manuscript.
The paper is available in publication form at the website for *The Teacher Educator*

Way to cite this paper:

Athanases, S. Z., Bennett, L. H., & Wahleithner, J. M. (2013). Fostering data literacy through preservice teacher inquiry in English language arts. *The Teacher Educator*, *48*(1), 8-28.

Author Note:

The research reported in this article was funded, in part, by a Catalyst Grant from the School of Education Board of Advisors, University of California, Davis. An earlier version of this paper was presented at the American Educational Research Association Annual Meeting, Vancouver, 2012. The authors thank Meghan Jones for her contribution to data entry and early analyses.

Abstract:

Data literate educators can generate questions about learning in their classes, collect and analyze classroom data to answer them, and develop inferences and explanations. Two elements may promote data literacy: (a) breadth of classroom-based data collection, and (b) depth of analysis. We examined these in inquiries conducted by 80 preservice teachers of secondary English in diverse classrooms over a six-year period. Analyzing products, processes, and self-reports, we found inquiries evidenced a range in breadth of data collection and depth of analysis. Qualitative themes and two cases illustrate data analysis processes and challenges of pattern-finding, initially crude analyses in striving for depth, and the need for data literacy mentoring.

Corresponding Author:
Steven Z. Athanases
szathanases@ucdavis.edu

Fostering Data Literacy through Preservice Teacher Inquiry in English Language Arts

Gina, a preservice English language arts (ELA) teacher, is reflecting on her first experience with teacher inquiry. She identifies components of argumentation she tracked in the writing of her 31 diverse 11th graders in a rural agricultural community, four of them English learners (ELs):

> When you are assessing whether a student has learned some fact or another, a quiz is sufficient. However, when you are assessing whether a student is capable of constructing an argumentative essay with developed argument paragraphs, you have to collect a set of essays….I was able to see how much they understood about making a claim, supporting that claim with evidence, and, most importantly, being able to analyze that evidence to show precisely how it supports their claim.

Gina goes on to report how the process requires ongoing focused attention: "Collecting multiple sets of data helps a teacher check for student growth or persistent problems students are having."

Gina highlights several inquiry processes: delineating elements of an achievement target, aligning assessment with learning goal, sampling repeatedly to track growth, and searching for and analyzing "persistent problems" in student work. Data analysis was not simple: "It only provided more questions and served to complicate my original research question." Still, she uses inquiry to explore how argumentation evolved in her class and for whom, speculates on causes, invokes (and critiques) two literature sources for explanatory power, and ponders future instruction and inquiry.

Gina's reflections on her first inquiry are drawn from a study we conducted examining inquiry products and processes of 80 secondary ELA student teachers (STs) over a six-year period. Gina's reflections reveal her developing *data literacy*. For preservice inquiry, we define data literacy as capacity to conduct focused and purposeful collection and analysis of student work, reflections, and process data, in order to promote reflection on student learning and to guide and inform new understandings of practice (Earl, 2005; Jacobs, Gregory, Hoppey, & Yendol-Hoppey, 2009; Popham, 2008; Schield, 2004). In calls for evidence-based practice, evidence is often considered from outside sources, decontextualized from classrooms. Teachers need support to learn also to generate *practice-based evidence*—derived from their own classroom practice (McNamara, 2002). Such efforts can situate and concretize data and evidence in the learning of students a teacher knows. Teacher inquiry holds potential to foster new teachers' situated processes of asking student learning questions and developing data-based answers.

In recent decades, educators worldwide have used teacher inquiry as part of professional practice. Teacher inquiry positions teachers as agentive professionals who can complement mandated standardized tests with analyses of student work and other classroom data derived from and informative about student learning and instruction. Efforts in teacher education (TE) have worked to initiate even preservice teachers into classroom-based inquiry practices; inquiry now occurs in many TE programs (Borko, Whitcomb, & Byrnes, 2008; Cochran-Smith & Lytle, 2009).

Currently, empirical evidence of values and outcomes of preservice inquiry remains slim (Darling-Hammond & Bransford, 2005; Grossman, 2005). Among areas warranting study is the nature of data collected and evidence generated in inquiry. Studies of these practices can yield information about potential for inquiry to help teachers develop capacity to study learning in their

classrooms and to examine impacts of instruction on students' learning. The present study sought such information. We asked: In what ways, if any, does teacher inquiry in one teacher education program foster preservice teachers' data literacy regarding student learning in their classrooms?

Conceptual Framework

Model of Teacher Inquiry as Systematic

Teacher inquiry typically positions practitioner as researcher rather than object of study, and includes gathering and analyzing data, collaborating in inquiry communities, developing an inquiry stance on learning, and striving toward social justice (Cochran-Smith & Lytle, 2009; Fecho & Allen, 2003; Gore & Zeichner, 1992; Goswami, Lewis, Rutherford, & Waff, 2009; Valli et al., 2006). Inquiry is characterized as intentional and systematic (Cochran-Smith and Lytle, 2009). The model framing the present study builds on this construct of inquiry systematicity.

Experienced teacher researchers have periods of systematic inquiry, but teaching demands at times trump systematicity, and an inquiry focus may recede or get recast (Athanases, 2011). For such teachers, inquiry is not a tidy formula; it occurs not as bounded studies but as ongoing attention to related, developing foci (Fecho, 2004). The inquiry model framing the present study, however, explores systematicity in context of preservice teachers' required work. The model relates to veteran teachers' inquiry but more specifically frames structured processes and bounded studies for STs that may provide heuristics for inquiry conducted later in teachers' careers.

Inquiry may arise from a puzzling moment, student, or learning pattern raising questions (Ballenger, 2009; Gallas, 2003). The model of systematicity framing the present study (Figure 1) features flexibility within structure, and includes levels of activity and processes of abstraction. At ground level are data collection events. Data include demographics, student writing, fieldnotes, quizzes, student interviews, taped discourse, and progress records. Data do not yield much information for teaching without analysis, a level up (Figure 1). Arrows pointing up show an analysis event can yield information for teaching, including trends in student performance, goals most students are and are not achieving, and groups of students struggling with curricular elements.

Information gleaned from analysis can be used in a distal or proximal manner. Distal use includes planning for future ways to teach a text differently or sequence activities better to build understanding. Proximal use alters practice or plans more immediately: redirecting instruction the next day, adding scaffolding for a learning challenge, creating a mini-lecture to address misconceptions or holes in students' knowledge. It also may include constructing activities to differentiate support in ways warranted by information the teacher learned from analysis.

This proximal use is indicated in Figure 1 by a dotted line arrow from each information box to the next collection event. The dotted arrow indicates how a teacher may repeat using information gleaned from analysis not only to redirect teaching but to prompt additional collection and analysis cycles in an ongoing pursuit of information that makes instruction responsive to student needs. This process parallels using assessment results to drive next-steps instruction in a learning culture (Shepard, 2000). The arrow's dotted line denotes how information gleaned from analysis may or may not be used in a proximal way each time. The teacher reflects on her/his own

actions and those of students (Schön, 1983), as well as information learned from data analysis, weighing merits of redirecting activity against time constraints and need for curriculum coverage.

Multiple rounds of data collection, analysis, and reflection may yield synthesis (Figure 1). For an ST with a hunch of what is occurring, a data-based, systematic process with analytic tools can promote conscious understanding to guide future actions and may lead to theory-level logical ordering and coherent framing to deeply understand varied but similar situations (Korthagen, 2010). This highest level of abstraction (Figure 1) is teacher knowledge that can develop from reflecting on rounds of collection and analysis. The knowledge and theory developed also enable *praxis*, practice or action informed by new knowledge and notions of change (Athanases, 2011). The process yields a dynamic, dialectical tension between doing and reflecting on doing (Freire, 1998).

In several studies, we analyzed STs' inquiries at this high level of abstraction, knowledge generated from inquiry. In one study, we found STs used inquiry to attend to culturally and linguistically diverse students and their learning needs by researching contexts; examining performance at group and individual levels; and asking and listening beneath the surface to attitudes, beliefs, and concerns about school learning (Athanases, Wahleithner, & Bennett, 2012a). In other studies (Athanases, Wahleithner, & Bennett, 2012b, 2012c), we found inquiry prompted new ELA teachers to develop pedagogical content knowledge (PCK) (Ball, Thames, & Phelps, 2008; Grossman, 1990; Shulman, 1987) for work with diverse learners. For the present study, we focused on the systematic process of inquiry itself. We examined the lower two levels in the model (Figure 1): data collection events and how STs analyzed raw data to glean information for teaching. We consider these two elements qualities of data literacy for inquiry.

Data Literacy and Teacher Inquiry in English Language Arts

Data literacy refers to the capacity to manage, understand, and critique proliferation of data in a high-tech, networked information age. School curricula include learning from and thinking critically about data, including analysis of data sources and learning to create common measures across raw datasets (Vahey et al., 2010). Data literacy also relates to information literacy and statistical literacy, including capacity to critique data (Steen, 1999). However, data literacy also includes capacity to conduct inquiry: formulating questions; collecting and organizing data systematically; using appropriate tools to analyze and display findings; and drawing conclusions.

In teaching, data literacy refers to knowledge teachers develop to interpret, generate, and use data on teaching and learning (Earl, 2005; Popham, 2008; Schield, 2004). It is more expansive than using standardized test scores and includes design (Jacobs, Gregory, Hoppey, & Yendol-Hoppey, 2009). Data literate educators generate data-based questions, select and evaluate data for answers, and develop inferences and explanations based on interpretations. These elements of an iterative process foster an "inquiry habit of mind"--reflection on "where you are, where you are going, how you will get there, and then turn around and rethink the whole process to see how well it is working and make adjustments" (Earl, 2005, pp. 8-9). This echoes Schön's (1983) reflection in and on action: Unexpected outcomes prompt reflection and future attention to multiple experiences. Data literate educators: (a) conduct collaborative inquiry to promote equitable achievement, (b) work with available data and data educators collect, and (c) learn to understand and name how they derive meaning from data (Love, Stiles, Mundry, & DiRanna, 2008).

As noted, for preservice teacher inquiry, we define data literacy as knowledge of how to conduct focused, purposeful collection and analysis of student work, reflections, and process data, to promote reflection on student learning and to guide and inform understandings of practice. Inquiry includes using available data but also learning to generate original data in the context of one's classroom. Interpreting data and generating evidence from inquiry can be difficult and can yield problematic claims, particularly when data are limited or methods of analysis are unclear (Foster, 1999). For this reason, two elements that may promote data literacy warrant examination: (a) breadth of data collection, and (b) depth of data analysis.

Breadth of data collection. The accountability climate has led teachers, more than ever, to equate assessment with testing and data with test scores. Inquiry into student learning requires a broader definition of data including the range of sources named earlier. Varied purposes require diverse methods, tools, and measures, with multiple, meaningful data sources enabling a balanced understanding of achievement, ways to represent data, and what data can and can't yield. Teachers have privileged access to all forms of student work; easy gathering of formal and informal student reflections; longer-term data collection that *being there* affords; and access to emic perspectives of student life from daily contact. Such opportunities yield data for those with methods, tools, and motivation to gather them and who play ongoing roles of both teacher and researcher. Breadth in data collection can lead to sharpened focus on individual students (Jacobs et al., 2009). This is key in learning the individual needs of diverse students and ELs, to see beyond one-size-fits-all instruction, beyond deficit perspectives, to capture learning that is there. Without multiple measures and levels of data, teachers (as all researchers) can reach faulty conclusions (Love, 2004).

For ELA, diverse data collection tools can capture unfolding learning. Tracking writing development calls for data including students' notetaking, maps, outlines, drafts, revisions, and reflections. Assessing literary analysis includes comprehension quizzes and essays, and discussion data with information on students' comprehension and interpretation of text, language proficiency, and development of argumentation and evidence. Breadth of data collection may include multiple events to track growth, alignment of data to research questions, and varied tools for triangulation. Also key are methods to ask and listen beneath the surface of activity, to understand student interests, reasoning, and challenges, using methods such as questionnaires, surveys, and interviews.

Depth of data analysis. Careful analysis requires repeated review of data supported by tools and processes to glean information. Data such as essays, discussions, and presentation notes are not amenable to easy scoring. Key to analysis of such data is pattern-finding, a process not altogether natural to teachers (Korthagen, Kessels, Koster, Wubbels, & Lagerwerf, 2001)—less so for STs. Such work includes a priori categories to code data, rubrics to quantify success in hitting targets, or repeated readings and analyses to track emerging themes. Also key is recognizing relevance and credibility of data and critiquing tools for fit and capacity to capture nuances. A tool or procedure builds understandings; another leads to other findings--as true for ethnographic work (Erickson, 2004) as for statistical analyses (Utts & Heckard, 2007). Finally, teachers may need to be versed in varied representation forms and interpretive processes, and develop ability to "access, assess, manipulate, summarize, and present data" (Schield, 2004, p. 4). These two key areas, then--breadth of data collection and depth of analysis--served as focal areas for our study.

Methods

Study Overview

This study is part of a program of research on the potential of inquiry to develop teacher knowledge, skills, and dispositions in TE. Aligned with our framework, we used systematicity in data collection and review across six years and multiple data sources for cumulative knowledge development about inquiry in one strand of a program. We focused on inquiry product and process data of secondary ELA STs. The present study features a credential-year course that was site for a first 10-week inquiry. We asked: In what ways, if any, does teacher inquiry in one TE program foster STs' data literacy regarding student learning in their classrooms? A sub-question included: To what degree do STs demonstrate breadth in data collection and depth of data analysis?

Context for the Study

The 10-month post-Baccalaureate program yearly credentials 150+ teachers in California. A prior set of studies found the program fostered advocacy for equity for diverse students (Athanases & de Oliveira, 2008), highlighting attention to ELs (de Oliveira & Athanases, 2007; Athanases & de Oliveira, 2011). Graduates reported how supervisors reinforced equity in diverse, mostly high poverty placements (Athanases & Martin, 2006). The program sought to learn how inquiry might support such teaching and advocacy (Merino & Ambrose, 2009). As of 2003, STs enrolled in two inquiry courses, with a second series in an optional second year M.A. track.

The program provides a fitting context to explore inquiry data and evidence. Archived memos note how earlier inquiry reports provided too little evidence of data and how claims were derived. This led to developing a systematic, data-based model. Supervisors also reported that in schools, a dominance of testing, scripted curricula, and pacing guides constrained STs' practices for diverse students, especially ELs. The TE Director noted that inquiry might help STs with "pushback" by reporting innovation results with clearly displayed data. Time constraints for STs and faculty led to creating strong scaffolds, with a template to make inquiry doable and transparent. The focus was inquiry, not report writing, with an outcome of a PowerPoint presentation with detailed notes. The process repeated in the M.A., with a longer inquiry and more detailed reporting.

Participants

The 80 secondary English STs in this study included 46 White females (57.5%), 15 White males (18.8%), and 19 women of color (23.8%). These 19 were 7 Latinas, 6 East Asian Americans (5 Chinese, 1 Filipina), and 6 Middle East/South Asian Americans. Many STs were bilingual. Inquiries were conducted in grades 7-12, with mostly majority students of color and large numbers of students living in poverty, just under half in classes of 33-100% ELs.

Roles of the Researchers

Researchers included three former K-12 teachers, White educators who have worked on ELA issues and diverse learners. First author (Steven) was ELA inquiry instructor and lead researcher. Other authors (Juliet and Lisa) recorded inquiry class fieldnotes, wrote memos on

themes in inquiry processes and products, and analyzed some of STs' work independent of first author. The study fits a research tradition of insiders studying programs and practices within an institution (Grossman, 2005). This enables us to provide rich insider perspectives, collect and archive large amounts of data over time, and produce knowledge through study of our context (Athanases et al., 2012a). Our detailed description of program contexts, inclusion of participant voices, and details of roles of researchers/authors who were not program or course insiders at time of the study address concerns of a program insider studying practice in an institution. We ensured supplemental data collection, analysis, and critical review of processes and STs' work by those external to the course "to interrogate findings and challenge the possibilities of self-fulfilling findings" (Clift & Brady, 2005, p. 333).

Data Collection

Data included 80 inquiries over six years, 2004-09, as PowerPoint presentations plus detailed Notes on slides (ranging 18-35 slides). Inquiries documented focus; class, school, and community contexts; research questions; and evidence justifying need for study focus and action plan. Inquiries also documented literature sources and their uses; study overview; and methods for data collection and analysis. Also required were analysis results with commentary; synthesis of inquiry learning; and next steps for future inquiry. Other data included data analysis field memos; questionnaires in which STs reflected on processes and products; taped discussions from inquiry class and solo and small group conferences; and annual final-day taped crosstalk. Data on course instruction included fieldnotes and taped classroom events. Researchers' analytic discussions and memos complicated and critiqued the inquiry instructional process and STs' engagement with it.

Data Analysis

We constructed a large database, with demographics, topics, and research questions, as well as coded data on issues and domains of study, teaching strategies used in inquiry, assessment modes used, and data collected and analyzed. Discourse data were transcribed. Aligned with our framework and research questions, we focused on two data literacy domains for inquiry: breadth of data collection and depth in data analysis.

To analyze breadth in STs' data collection, we first counted collection events per inquiry. STs were expected to document three collection events for conducting and reporting analyses. We defined a data collection event as a time when an ST purposely collected work or gathered other student data for investigating a framed student learning issue. Many STs documented more than the required three events, and we wanted the number of events STs explored in their inquiries. Second, we wanted to know *how*, across the corpus of inquiries, STs collected data and of what kinds. Following Stiggins (1994), we coded collection tools, using four assessment modes: selected response (multiple choice, matching); essay (constructed written responses); performance (observed demonstrations of understanding); and personal communication (student self-reports and informal gathering of student perspectives). Within modes we identified tools across events in 80 inquiries. We reviewed and tabulated *number* and *diversity* of collection tools per inquiry.

Finally, to analyze variation in collection per inquiry, we assigned inquiries to one of three data literacy opportunity profiles, characterizing the degree to which inquiries afforded opportunity to explore and potentially develop data literacy. *Entry-level* indicated an ST used one

collection tool throughout, regardless of number of events. *Broadening* indicated two tools and minimum of four events, one above the required three, indicating greater breadth of data collection exploration, as the ST sought to learn about student learning in more than one way and tried for more data sampling. *Broad* indicated three or more tools and five or more collection events, a broadened database providing multiple, varied opportunities for data literacy development.

For depth of analysis, we used archived scores from class assigned to three required analysis events per inquiry. These excluded baseline, exit, or supplementary collection events an ST completed. For each event, 3 indicated ST thoroughly explained how data connected to the research question; included a clear example of tool and response; and analysis showed depth and complexity of interpretation. A 2 indicated connections to research question were unclear; or tool and response were unclear; or analysis needed greater depth. Score of 1 included problems in two or more areas. Scores were derived annually, with interrater reliability scoring conducted by instructor and T.A. The process involved rubric review and practice scoring a subset of inquiries. Assessors reviewed scores for agreement, with discrepant scores reviewed until agreement was reached. Double-scoring followed with several inquiries, until agreement was reached on all rubric items. Remaining inquiries were single-scored, with double-scoring of difficult cases. For three focal events, scores of 1-3 were averaged for a data analysis average score.

We also examined crosscutting issues in inquiry products and processes, guided by the constant comparative method (Merriam, 1998). We coded and analyzed questionnaires, interviews, and discussions for themes in STs' data collection and analyses--and triangulated, examining ways themes emerged in STs' actual inquiry work. Using product and process data, we constructed two cases to illustrate and complicate findings. Process data included STs' field memos that offered preliminary analyses of a dataset, reflected on patterns, and raised questions. Other process data included coded and analyzed ST questionnaires and inquiry class discussions. We selected STs for cases with contrasting data literacy opportunity profiles but both with high average data analysis scores, highlighting different routes to achieving strong analyses. Cases also analyzed data literacy growth, as both STs struggled but grew within the scope of their first inquiries. Developing data literacy while learning to teach is no easy task; our cases highlight this double challenge.

Results

Breadth of Data Collection in Preservice Teachers' Inquiries

Number of data collection events. All 80 STs used a required three events, ranging from 3 to 14 (mean of 4.85). Twenty-four (30%) used just the required three, and a majority (71, 89%) used from three to six events. At the high end, nine STs (11%) used 7 to 14 events. Often these are repeated instances of sampling student work or perspectives using the same tool; in some cases, varied tools were used for multiple events. At an overall mean of 4.85 events, STs demonstrated that even in their first, fairly short-term inquiry, they used multiple opportunities to collect data.

Data collection tools across inquiries. Table 1, Column 1 (Data collection events) displays tools used in collection events across inquiries. Of 388 total events, more than half (206, 53.1%) used essay mode. Table 1 shows events coded as essay were more than the combined total coded as personal communication (77), performance (51), and selected response (54). Table 1,

Column 2 (ST inquiries) displays distribution of collection tools across inquiries, or *how many* STs used each tool. Examining assessment modes (Table 1, both columns) confirms that essay dominated (81.3% of STs), with selected response least common (Column 2, 29%). Though selected response has dominated classroom assessment across subject areas (Stiggins, 1999), it is not surprising that essay dominated collection in our database, given a language-production focus of English and the course press to match tools with inquiry purposes. However, multiple-choice comprehension and vocabulary quizzes are widely used in texts and tests in schools where STs were apprenticed, and state assessments that drive much instruction use almost exclusively multiple-choice items. Therefore, it is notable that STs explored essay responses in which students needed to produce language and for which STs needed to generate and use tools of analysis.

Just under one-fifth (19.8%) of all data collection events and over half of all inquiries (41, 51.3%) used personal communication, indicating many STs used tools such as surveys, questionnaires, and interviews to learn students' perceptions and challenges related to classroom work. One ST remarked, "Datasets based on surveys and observations…offer more qualitative input, and may help us to design better informed solutions to problem areas." This comment represents an understanding of differing values of assessment modes and a growing data literacy.

Multiple measures in inquiries. The range of tools (Table 1) suggests multiple tools and measures were used in inquiries. We also wanted to know how many different tools *each* ST used—or breadth of tool use per inquiry. Figure 2 shows the range in number of collection tools STs used within their inquiries. As the Figure shows, 11 (13.8%) STs used the same tool throughout the inquiry. However, nearly two-thirds (51, 63.8%) used two to three different data collection tools. Nearly one-fourth of STs (18, 22.6%) used four to eight different tools to guide their inquiries. These STs incorporated multiple methods to track and study student progress.

How varied were tools per inquiry? We found 27 STs (34%) used collection tools from just one mode. For example, an ST used short answer and paragraph, both in essay mode. However, nearly two-thirds (52 STs, 65%) used collection tools from two to three different modes. Such cases may have included short answer (essay mode) to study students' comprehension; observation notes (performance mode) to capture students' participation patterns in reading discussions; and survey (personal communication) to learn students' challenges or resistance with comprehension strategies. This variation demonstrates potentially greater breadth in tool use. One ST reflected that she "would have really liked to have more data about the process. I could have video/audio-taped a lesson, or taken notes as students worked. This would have shown me who was struggling more than others." This ST reveals how she learned to view data as a way to learn about her students, information she could not obtain without being purposeful in her collection.

Data literacy opportunity profile. Of 80 inquiries, 27 (33.8%) were coded as *entry-level*, indicating the ST used just one collection tool throughout, regardless of number of events. At a level of *broadening*, 22 STs (27.5%) used two tools and a minimum of four collection events, one above the required three. This profile indicated greater breadth of data collection, as the ST sought to explore student learning in multiple ways and tried for more data sampling. Coded with a profile of *broad*, 31 STs (38.8%) used three or more tools and five or more data collection events--a broadened database providing multiple, varied opportunities for data literacy development.

Depth of Data Analysis in Preservice Teachers' Inquiries

Overall scoring performance. Nearly half of STs (Figure 3) received a data analysis average score of 3. Another 22 STs (27%) received between 2 and 3, approaching depth. Still, 19 STs (nearly 25%) received average scores of 2 or lower, indicating data analysis problem areas.

Qualitative themes. Three themes emerged about STs' analysis processes. One theme was *reframing*. STs mapped studies but analysis often led to sharper focus or reframe. One ST selected vocabulary from research terms and model papers to scaffold ELs' research papers on careers. She used "a mix of different types of terms--conjunctions, career-specific vocabulary, transition words and general academic terminology." After initial instruction and analysis of quiz results, she found

> nearly all students had made errors in identifying the transition words. As I thought more and more about why students were having trouble with these words in particular, I realized that I needed to change the focus of my inquiry to address this.

In such cases, STs described ways they used information gleaned from analysis to make next-steps instructional choices and to reframe and redirect inquiries.

A second theme was *pattern-finding*. STs valued rubrics as tools for pattern-finding from raw data. One ST noted value in having "some way of converting the information you collect into numbers so that you can compare and analyze the data for patterns if possible." Another reflected on how inquiry makes teachers "investigators of patterns": "Instead of just realizing that certain students are having difficulties, say, decoding assignments, the process invites us to investigate what is causing the problem. And then taking it a step further, how to be of help to the students." This ST names what we found in STs' work: pattern-finding can yield information on details of student struggles and can guide action. Finally, we found STs using pattern-finding for synthesis:

> When I compared scores of all of my data-gathering, it pushed me to look for trends and explain them. I began to see patterns of things students consistently did well or struggled with and I thought it was very valuable to be able to so closely compare everything.

This highlights how reflecting on multiple data collection and analysis cycles promotes synthesis.

A third theme was the value of *mining student responses as data*. While some short-answer work yielded only abbreviated language production and missed opportunities for deep analysis, in other cases, language from even short answers enabled STs to learn about students' learning. We found that more important than length of student response prompted by a task was quality of analytic treatment undertaken by the ST in mining responses as data. For example, for her inquiry on improving reading comprehension strategies, one ST taught students to interact with text by making textual annotations. She collected and analyzed these, coding each by reading strategy it represented. The ST mined these short responses for key information. In addition to providing a graphic representation and discussion of patterns of strategies her students as a whole employed, the ST included sample student work and her analysis from high, mid, and low achievers. Though data she collected were phrases or sentences, careful analysis allowed the ST to glean a tremendous amount of information about students' strengths and weaknesses as readers.

Cases of Data Literacy Processes and Development in Preservice Teacher Inquiry

The cases illustrate our findings and complicate them. Vann's inquiry was coded as *broadening* in data literacy opportunities, while Cara's was coded as *broad*, using twice as many data collection events as Vann and more tools. However, both received high average data analysis scores of 3. Both struggled early in their inquiries but demonstrated growth in data literacy.

Vann: Developing and testing a writing assessment tool. Vann provides a case of an ST who demonstrates evolving ability to think deeply about student writing.

Inquiry overview. Vann, a White male, conducted inquiry in a school with students from diverse backgrounds (European, Latino, African American, Hmong, Chinese, and Filipino, among others). A majority (70%) received free/reduced lunch. In his class of 28 sophomores, five were ELs. Observing lack of evidence in student essays, Vann focused his inquiry on scaffolding use of evidence: "My goal is to have all my students using quotes and evidence effectively and explaining the quotes they used and connecting their quotes to their ideas and arguments."

Data literacy opportunity profile. Vann's inquiry, coded as broadening, had two collection events beyond the three required and two tools (short answer responses and full essays, both within essay mode). He structured data collection as baseline, three focal events, and exit data. After a rough start, Vann created a rubric to analyze all data he would collect from later events. Though he relied heavily on a single data collection tool (student essays) and a single analytic tool (a rubric), he hoped that repeated use of that rubric might help him detect nuanced changes in students' writing. His use of the tool evolved, demonstrating growth in data literacy.

Data exploration: First rough steps. Vann began his inquiry asking students to explain the meaning of a word in a poem, supported by textual details. Vann analyzed responses on two criteria: "used details from the poem to support their ideas" and "any kind of real explanation or connection of the quote or details to their idea." He reported only presence/absence of these, that 10 of 25 students used details for support, and of those 10, only 5 explained the quote or made connections from quote to overall point. What Vann gleaned from this analysis was problematic:

> I looked at the performance of the students who did use quotes and they performed well, so my best speculation is that those students who did not use quotes were simply lazy or did not pay attention in class. So this means that the instruction was not poor in providing students with the knowledge of how to use quotes, but it seems to be a lack of interest or motivation that caused them to not use the quotes.

Vann did not unpack ways students used quotes, and his speculation about why students did not use quotes came without evidence and without testing against other speculations or hypotheses. He also exonerated himself without reflection on his practice.

Ongoing data collection and analysis process. For their next essays, Vann prompted students to include a supporting quote, explanation, and connection to overall point. He reviewed data with the instructor in an extended conference. The instructor pressed Vann to search for nuances of achievement, teasing apart criteria and articulating performance levels for them. Vann reported feeling empowered by what he developed. He analyzed the data by looking for presence

of a quote but then added a level of analytic specificity by scoring students from 0 (not present) to 3 (strong) on three elements: relevance, explanation, and connection to overall argument.

Vann first found all students used a quote. In his second level of analysis, he learned most selected relevant quotes (mean score 2.8) and provided strong explanations (mean score 2.5). The analysis prompted further reflection. Vann noted scores likely were inflated because he directed students to specific points in the text to find quotes. This claim was borne out by his observation that students received no direct instruction to connect quotes with overall arguments, a criterion for which scores averaged 1.86. Vann considered his instruction further: "Had I provided a model, it would have been clearer as to whether students simply did not understand what was expected or that they struggled to or chose to ignore the context portion of the question." Based on analyses, Vann decided his scaffolds needed to be more explicit regarding what to include in essays.

For his next dataset, Vann used the same rubric, regretting that 9 of 21 students did not include quotes: "This is way too high of a number which tells me something is wrong. I am not sure if it is laziness, lack of understanding, poor instruction, or inability to follow instructions." The rubric still allowed him to analyze patterns in essays that *did* include evidence. He found that, across areas (relevance, explanation, and connection), students scored a bit lower than the previous time but those who did use quotes overall selected relevant evidence that supported arguments. Vann's analysis prompted further reflection, and he used new information in a proximal manner, vowing that for the next task he would "provide more models and guided practice with these students in hope that, with experience as well as better scaffolding and improved instruction on my part, they will come to understand the task better and be able to perform it."

Next time, Vann provided explicit instructions and a model. However, only eight students completed the task. Despite the slim dataset, Vann mined it for information on students' learning and the strength of scaffolds. Of eight essays, six provided explanation for quotes. Again, Vann averaged scores per criterion across essays, then examined individual student scores. Realizing he needed more data to confirm findings, Vann still speculated that scaffolding led to all students using quotes and "an increase in scaffolding has a significant increase in the students' ability to explain the meaning of quotations." However, he did not address who the eight students were who submitted the assignment. He speculated that low submission rate was due to burnout, which may suggest those completing the assignment were highly motivated and perhaps typically high performers. Stronger analysis would have compared just these students' scores across datasets.

For exit data, Vann used his rubric to track growth since his first dataset. Submission rate of essays again was low, though essays submitted showed gains in all areas. Vann reflected on the slim dataset but also critiqued his design. He noted that unlike when writing essays as baseline data, students at exit had scaffolds that Vann left in because he ran out of instructional time to remove them gradually. He acknowledged the limitation: "It means the exit data assessment scores will most likely be higher than the baseline scores…it does not show whether implementing scaffolds as a teaching strategy will increase student scores if those scaffolds are then taken away." The ability to critique his data for what the data can tell him again signals Vann's evolving thinking about how to use data to guide instruction and the limitations of those data.

Data literacy development. Vann's evolving data literacy was evident in his development of an analytic tool to capture nuances of learning, and in his capacity to use data to reflect on his

instruction and implicate himself in reflection on outcomes. Also, Vann developed analytic clarity about tracking student averages and using such data in a proximal way for next-steps instruction. Moreover, Vann developed capacity to critique datasets for what they can reveal. Vann speculated about why some students did not submit essays, and why students achieved certain outcomes and not others. A broader data collection toolkit would include questionnaires, interviews, and other personal communication data. Such data might have enabled Vann to move past pure speculation to data-based information about student understanding, perspectives, and possible resistance.

Key to Vann's data literacy development was use of the same tool to analyze writing throughout his inquiry, which proved essential to his developing knowledge for teaching writing:

> What is satisfactory? What is excellent? What isn't? So you describe each level in your own words, so you get a better idea of what you want. So you have a better idea of how you're going to instruct the students to get there or what they need. By doing that you have more of an expert understanding of what you're looking for in everything you've done.

Developing the rubric forced Vann to both articulate characteristics of a response to literature essay and to plan instructional moves to guide students' development. He came to realize students need multiple scaffolds and models to understand rhetorical moves. The tool, then, assisted Vann in looking for ways to help students reach excellence, in thinking in multiple ways about student work, and in developing new knowledge about content, pedagogy, and data literacy.

Cara: Using multiple data collection events and tools to study discussions. The case of Cara provides an example of an ST who initially struggled with the process of data analysis but developed her ability through multiple opportunities to engage authentic, meaningful data.

Inquiry overview. Cara, a White woman, conducted her inquiry in a 9th-grade honors English class in a diverse high school just outside a major city. The ethnic makeup of Cara's class was: 15 Asian, 9 African American, 6 Latino/a, 2 East Indian, and 1 Caucasian. All students were English proficient. Cara sought ways to improve quality and quantity of participation in *Romeo and Juliet* discussions. Cara asked this research question: "If we read the text in class, allow students to do a quick write on the topic before the class discussion and vary the discussion technique, will student participation increase and will students' responses improve?"

Data literacy opportunity profile. Cara's data collection was coded broad, including 10 separate data collection events and three data collection tools. Data collection tools included surveys, transcribed discussions, and written literature responses. Multiple collection events (beyond three required) and multiple measures provided Cara with ample opportunity to conduct analysis with meaningful data and to earn a 3 for data analyses. These opportunities, along with repeated analyses, enabled Cara to make great strides in data literacy for inquiry.

Data exploration: First rough steps. Cara struggled to analyze her early data. Of a possible score of 10, Cara scored 5 on her first field memo and 6 on her second. The scores reflected lack of depth in analyses. The first memo focused on baseline discussion data in a prior unit on *The Odyssey*. Cara found students "did not talk as thoroughly or as in-depth as I would have liked and hence the whole classroom discussion was not as thorough as I would have liked." In feedback on the memo, the instructor noted that students require scaffolding for discussion. The

memo revealed that Cara struggled at this point with elements of data literacy. As she examined baseline discussion data, she assigned scores of advanced, basic, and below basic without a rubric that articulated criteria. Cara neglected to provide samples to illustrate scores, attached unanalyzed data to her memo, and made general statements, such as "They wrote a lot" and "As seen by my notes, more students talked." The instructor explained to Cara that her data were not adequately analyzed and that she needed to be more detailed and thorough in written analyses.

In response to input on students needing discussion scaffolds, Cara conducted a survey to determine students' prior knowledge about discussions. She reported in her second field memo, "Six of the responses were negative towards classroom discussion. 18 of the students had positive responses to class discussions." The instructor wrote, "Way too general! You are not revealing how you analyzed these surveys. Where is the rubric? Where is an explanation of data sources?"

Despite this rocky start, Cara gleaned information from her analyses: that students had not been explicitly taught structures and expectations of discussions, that they did not use the text or personal experiences to support responses, and responses did not build on those of classmates. Cara's baseline findings prompted her to design scaffolds for participation in discussions on *Romeo and Juliet*. These lessons became the basis of Cara's focal data collection events.

Ongoing data collection and analysis process. For her first *Romeo and Juliet*-themed dataset, Cara asked students to generate open-ended questions, then facilitated discussion using the questions. Cara collected and analyzed the questions and taped the discussion. She prepared a sample discussion transcript for a student-generated question: Does Nurse want what is best for Juliet? Cara found that while students could recite basic information about the play, they still did not cite examples from the text and did not support claims with evidence. She also found through transcript analysis that students did not connect responses to what had previously been said:

> The first student's response [to the question about Nurse's relationship to Juliet] is about the sleeping potion. The next student's response is the incorrect fact that Nurse did not think that Juliet was serious about Romeo. The next student sort of corrects the student but then jumps to how Nurse tries to convince her to marry Paris. Each student could have been the first to respond to the question; the responses did not build on each other.

While Cara noted students were not building on prior remarks, she found that having students generate questions before discussion promoted student-initiated questioning, a departure from the teacher-directed questioning Cara documented in her baseline data. Cara counted total student responses, offered percentages of time students and teacher spoke, noted who initiated discussion, and depicted data graphically. She was pleased with the ratio of student to teacher response but still sought to improve both quality of response and number of student participants.

In Cara's next dataset, students considered who was at fault for the deaths of Romeo and Juliet. Students wrote quick-write arguments for why six different characters might be blamed. Cara designated six places in the room, asking students to move to the area for the character they felt was to blame. Once in groups, Cara asked students to identify reasons their selected character might be blamed. She facilitated a full-class discussion in which groups shared reasons with other groups challenging. Cara provided transcribed discussion and detailed analysis. She found more students linked responses to prior comments, more referenced text, and students did not require as

much prompting to carry discussion without teacher facilitation. Again, Cara reported number of responses, percentages of student and teacher turns, and who initiated conversation, depicting data graphically. She found a greater percentage of students (64.7%) participated in discussion.

Cara's next *Romeo and Juliet*-themed dataset was a think/pair/share discussion with students arguing whether Romeo and Juliet were in love. Students did a quick-write, discussed in pairs, then engaged in full-class discussion. Cara used transcribed discourse to demonstrate how discussion ensued without prompts or her intervention. In her analysis, Cara graphed student and teacher participation and student response. Cara also offered her hypothesis about why some students did not participate, found patterns across results, and discussed implications of her study.

At the end of her inquiry, Cara surveyed students to uncover which discussions they liked best, how prepared they felt to participate in each discussion type, and how they rated their own participation. She then compared her baseline data, which featured students' participation in discussions about *The Odyssey,* with her *Romeo and Juliet*-themed inquiry data and offered analysis. She found that numbers of students participating and teacher involvement did not change but the quality of involvement increased. She attributed the change to several features of her inquiry activity, including reading the play out loud as a class, requiring quick-writes before discussions as a method of organizing thoughts, and variety in discussion strategies.

Growth in data literacy. Through inquiry, Cara learned to facilitate literature discussion and conduct data analysis. Despite a rough start, in her third memo, Cara demonstrated a level of specificity required in conducting and reporting data analysis. The memo included a discussion transcript, a rubric to analyze it, and reports of patterns. These elements enabled Cara to focus on areas such as number of responses per student, ratio of student to teacher talk, and number of responses built on previous statements. Positive feedback from the instructor let Cara know she was on track to analyze subsequent datasets. Through repeated, scaffolded data analysis with a variety of tools, Cara began to develop data literacy as she conducted inquiry.

Discussion

Preservice teacher inquiry provides a means to develop capacity to collect, analyze, and glean information from classroom-based data. Even while learning to teach, STs in our study more than met minimum required data collection events and overall used multiple measures. As noted earlier, many equate data with test scores. Casting data far more broadly in classrooms, STs in our study benefited from multiple data collection events, varied collection methods and tools, and practice analyzing data for diverse purposes. In these ways, STs were constructing knowledge about learning through authentic experiences and were using information to guide instruction.

Those who varied modes and/or used many data types may have had more opportunities to learn to represent and triangulate data to build understanding with a union of insufficient measures. Just over half of STs used personal communication methods to gather data that revealed attitudes, challenges, and perceptions related to classroom work. This is a promising mode to learn of the interaction of students, instruction, and subject matter and warrants further exploration for equipping STs with tools they might not otherwise explore as new teachers. Our case of Vann highlights how, without such data, the ST is left to speculate about reasons for patterns--from low submission rate of student work to variation in achievement on different performance criteria.

Roughly half of STs earned high data analysis scores. This is notable, given the dominance of essay as STs' favored data collection tool. Essays are language-rich, not easily quantified; STs needed to review such work repeatedly, exploring ways to parse language, code it, characterize it, evaluate it. These analytic processes enabled STs to deepen their knowledge of learning targets and what it takes to hit them, elements of data literacy. Our case of Vann illustrates this analytic care in reviewing focused achievement targets in student essays and evolving clarity in grasping students' learning about writing. As they conducted analysis, both Vann and Cara deepened their PCK in specific dimensions of ELA curriculum. Though necessary for teaching, such knowledge is difficult to develop early-career but can be scaffolded by inquiry (Athanases et al., 2012b).

Breadth of data collection and depth of data analysis, our two data literacy targets, hold promise in developing teachers' capacity to observe students' learning, frame emerging problems, collect relevant work, and analyze information to understand patterns. Such work can enable teachers to monitor ongoing learning and to respond. Moreover, such data literacy can bring equity issues to the fore, enabling teachers to disaggregate data to understand which individuals and groups are performing at what levels on what things; what issues are shaping students' successes, problems, and resistance; and what kinds of differentiation and further scaffolding are needed. We saw evidence of this as STs reported data trends, for example, related specifically to ELs.

Our study is limited by a focus on one program. Inquiry models include other foci; we therefore specified elements of the model and why it held promise for examining data literacy development through inquiry. Some of these elements include a course series on systematic inquiry, scaffolds including data analysis workshops, and mentoring and cohort groups. Data literacy is multi-faceted. In this study we focused on two elements—breadth of data collection and depth of analysis. Other elements warrant attention, including how STs align data with questions, display data, build evidence, and substantiate claims. This study also is limited by relatively short inquiries. To mitigate this concern, we examined inquiries of 80 STs over a six-year period. A follow-on study also will provide longitudinal perspectives of a group of now early-career teachers.

Finally, our study also demonstrates that while possible, data literacy development in preservice inquiry is challenging for STs. Even with a highly scaffolded process and community supports, nearly 25% of STs earned average data analysis scores of 2 or below on a 3-point scale, indicating problems in analyzing and reporting data. STs needed to engage in many data-based and analytic processes mostly new to them, while facing countless demands of learning to teach. The cases of Vann and Cara, who both achieved strong data analysis average scores, shed light on learning to work with data, challenges of pattern-finding, initially crude analyses as STs strive for depth, and a need for critical feedback and mentoring for data literacy growth. Even more veteran teachers need sufficient time, resources, guidance, and institutional support to conduct inquiry (Cochran-Smith & Lytle, 2009; Freedman et al., 1999; Huffman & Kalnin, 2003). For STs, supports may be more critical, when time constraints can be daunting and prior inquiry experience minimal. Still, when teachers learn to collect and analyze data as STs, they may internalize these processes so they can draw on them throughout their teaching careers (Korthagen, 2010). Our study highlights the promise of systematic and scaffolded inquiry, even in preservice, in helping to develop data literate, evidence-generating professionals closely tracking diverse students' learning in response to a range of content-based pedagogical actions.

References

Athanases, S. Z. (2011). Research as praxis: Documenting the dialectical relationship between theory and practice. In D. Lapp & D. Fisher (Eds.), *Handbook of research on teaching the English language arts, 3rd Ed* (pp. 358 -363). Sponsored by the International Reading Association and the National Council of Teachers of English. Erlbaum/Taylor Francis.

Athanases, S. Z., & de Oliveira, L. C. (2008). Advocacy for equity in classrooms and beyond: New teachers' challenges and responses. *Teachers College Record*, *110*(1), 64-104.

Athanases, S. Z., & de Oliveira, L. C. (2011). Toward program-wide coherence in preparing teachers to teach and advocate for English language learners. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms* (pp. 195-215). NY: Taylor & Francis.

Athanases, S. Z., & Martin, K. J. (2006). Learning to advocate for educational equity in a teacher credential program. *Teaching and Teacher Education, 22*(6), 627-646.

Athanases, S. Z., Wahleithner, J. M., & Bennett, L. H. (2012a). Learning to attend to culturally and linguistically diverse learners through teacher inquiry in teacher education. *Teachers College Record*, *114*(7).

Athanases, S. Z., Wahleithner, J. M., & Bennett, L. H. (2012b). Preservice teacher inquiry and content knowledge for teaching. Manuscript submitted for publication.

Athanases, S. Z., Wahleithner, J. M., & Bennett, L. H. (2012c). Learning about English learners' content understandings through teacher inquiry: Focus on writing. Manuscript submitted for publication.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407.

Ballenger, C. (2009). *Puzzling moments, teachable moments: Practicing teacher research in urban classrooms*. NY: Teachers College Press.

Borko, H., Whitcomb, J. A., & Byrnes, K. (2008). Genres of research in teacher education. In M. Cochran-Smith, S. Feiman-Nemser, & D.J. McIntyre (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (pp. 1017-1049). NY: Routledge.

Clift, R. E., & Brady, P. (2005). Research on methods courses and field experiences. In M. Cochran-Smith & K. M. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 309-424). Mahwah, NJ: Erlbaum.

Cochran-Smith, M., & Lytle, S. L. (2009). *Inquiry as stance: Practitioner research for the next generation*. NY: Teachers College Press.

Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco: Jossey-Bass.

de Oliveira, L. C., & Athanases, S. Z. (2007). Graduates' reports of advocating for English language learners. *Journal of Teacher Education*, *58*(3), 202-215.

Earl, L. (2005). From accounting to accountability: Harnessing data for school improvement. Submission from Australian Council for Educational Research Conference 2005--Using Data to Support Learning. 6-10. http://research.acer.edu.au/research_conference_2005/8

Erickson, F. (2004). Demystifying data construction and analysis. *Anthropology and Education Quarterly*, 35(4), 486-493.

Fecho, B. (2004). *"Is this English?" Race, language, and culture in the classroom*. NY: Teachers College Press.

Fecho, B., & Allen, J. (2003). Teacher inquiry into literacy, social justice, and power. In J. Flood, D. Lapp, J. Squire, & J. Jensen (Eds.), *Handbook of research on teaching the English*

language arts 2nd ed. (pp. 232-246). Mahwah, NJ: Lawrence Erlbaum.

Foster, P. (1999). "Never mind the quality, feel the impact": A methodological assessment of teacher research sponsored by the Teacher Training Agency. *British Journal of Educational Studies*, *47*(4), 380-398.

Freedman, S. W., Simons, E. R., Kalnin, J. S., Casareno, A., & The M-CLASS Teams (1999). *Inside city schools: Investigating literacy in multicultural classrooms*. NY: Teachers College Press.

Freire, P. (1998). *Pedagogy of freedom: Ethics, democracy, and civic courage*. Lanham, MD: Rowman and Littlefield.

Gallas, K. (2003). *Imagination and literacy: A teacher's search for the heart of learning*. NY: Teachers College Press.

Gore, J. M., & Zeichner, K. M. (1991). Action research and reflective teaching in preservice teacher education: A case study from the United States. *Teaching and Teacher Education 7*(2), 119-136.

Goswami, D., Lewis, C., Rutherford, M., & Waff, D. (2009). *On teacher inquiry: Approaches to language and literacy research*. New York: Teachers College Press.

Grossman, P. (1990). *The making of a teacher*. NY: Teachers College Press.

Grossman, P. (2005). Research on pedagogical approaches in teacher education. In M. Cochran-Smith & K. M. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 425- 476). Mahwah, NJ: AERA.

Huffman, D., & Kalnin, J. (2003). Collaborative inquiry to make data-based decisions in schools. *Teaching and Teacher Education, 19*, 569-580.

Jacobs, J., Gregory, A., Hoppey, D., & Yendol-Hoppey, D. (2009). Data literacy: Understanding teachers' data use in a context of accountability and response to intervention. *Action in Teacher Education*, *31*(3), 41-55.

Korthagen, F. A. J. (2010). Situated learning theory and the pedagogy of teacher education: Towards an integrative view of teacher behavior and teacher learning. *Teaching and Teacher Education*. *96*(1), 98-106.

Korthagen, F.A.J., Kessels, J., Koster, B., Wubbels, T., & Lagerwerf, B. (2001). *Linking practice and theory: The pedagogy of realistic teacher education*. Mahwah, NJ: Lawrence Erlbaum, Rpt. Routledge (2008).

Love, N. (2004). Taking data to new depths. *Journal of Staff Development, 25*(4), 22-26.

Love, N., Stiles, K. E., Mundry, S., DiRanna, K., (2008). *The data coach's guide to improving learning for all students: Unleashing the power of collaborative inquiry*. Thousand Oaks, CA: Corwin.

McNamara, O. (2002). *Becoming an evidence-based practitioner: A framework for teacher-researchers*. NY: Routledge Falmer.

Merino, B. J., & Ambrose, R. C. (2009). Beginning teachers' inquiry in linguistically diverse classrooms: Exploring the power of small learning communities. In C. J. Craig & L. F. Deretchin (Eds.), *Teacher Education Yearbook XVII: Teacher learning in small group settings* (pp. 242-260). Lanham, MD: Rowman & Little.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey Bass.

Popham, W. J. (2008). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, *48*(1), 4-11.

Schield, M. (2004). Information literacy, statistical literacy and data literacy. *IASSIST Quarterly, Summer/Fall*, 6-11.

Schön, D. A. (1983). *The reflective practitioner, how professionals think in action*. NY: Basic.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Shulman, L. S. (1987). Knowledge and teaching foundations of the new reform. *Harvard Education Review*, 57(1), 1-22.

Steen, L. A. (1999). Numeracy: The new literacy for a data-drenched society. *Educational Leadership, 57*(2), 8-13.

Stiggins, R. J. (1994). *Student-centered classroom assessment*. NY: Merrill/Macmillan.

Stiggins, R. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*, November, 191-198.

Utts, J. M., & Heckard, R. F. (2007). *Mind on statistics*, 3rd Ed. Pacific Grove, CA: Duxbury.

Vahey, P., Rafanan, K., Swan, K., van 't Hooft, M., Kratcoski, A., Stanford, T., and Patton, C. (2010, May). *Thinking with data: A cross-disciplinary approach to teaching data literacy and proportionality*. Presented at the American Educational Research Association Annual Meeting, Denver, CO.

Valli, L., van Zee, E. H., Rennert-Ariev, P., Mikeska, J., Catlett-Muhammad, S., & Roy, P. (2006). Initiating and sustaining a culture of inquiry in a teacher leadership program. *Teacher Education Quarterly 33*(3), 97-114.

Table 1
Total Number of Data Collection Events and ST Inquiries in Which Each Tool Was Used

	Data collection events	ST inquiries
Mode of assessment *	Number (and percent) of total events using (n=388)	Number (and percent) of teacher inquiries using (n=80) **
Essay	**206 (53.1)**	**65 (81.3)**
Short answer	110 (28.4)	54 (67.5)
Paragraph	53 (13.7)	28 (35.0)
Essay	38 (18.4)	23 (28.8)
Research paper	2 (0.5)	2 (2.5)
Poem	3 (0.77)	1 (1.3)
Personal Communication	**77 (19.8)**	**41 (51.3)**
Survey	35 (9.0)	25 (31.3)
Self-report of work	15 (3.9)	11 (13.8)
Questionnaire	10 (2.6)	8 (10.0)
Interview	12 (3.1)	5 (6.3)
Self-evaluation	5 (1.3)	3 (3.8)
Performance	**51 (13.1)**	**27 (33.8)**
Notes/observations	30 (7.7)	20 (25.0)
Rubric as record	11 (2.8)	7 (8.8)
Audio	6 (1.5)	4 (5.0)
Video	4 (1.0)	4 (5.0)
Selected Response	**54 (13.9)**	**23 (28.8)**
Quiz	29 (7.5)	15 (18.8)
Test	25 (6.4)	14 (17.5)

* Mode of assessment coded using Stiggins' (1994) analytic framework for student-centered classroom assessment

** Because most inquiries used more than one mode of assessment, the total number for assessment modes reported in Column 2 (ST inquiries) is greater than 80. Also, in this same column, the total number for tools within each mode is greater than each mode total, because many inquiries used more than one tool per assessment mode. For instance, an inquiry may have included both a short answer assessment and a research paper as assessment. This inquiry then would be counted in the total for each assessment tool but only once for the mode as a whole.

Figure 1. Systematicity in Teacher Inquiry

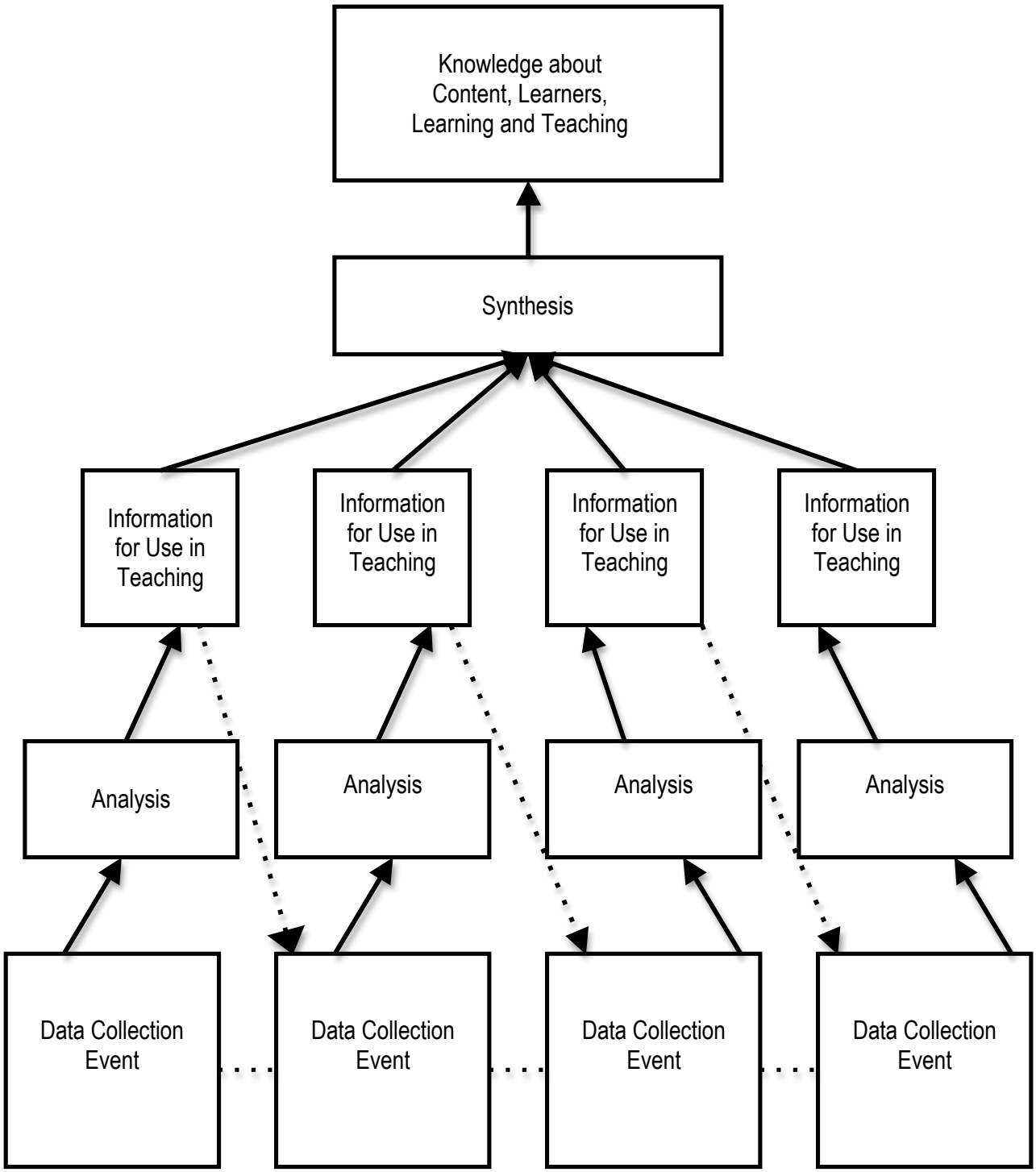


Figure 2

Total Number of Data Collection Tools Used Per Student Teacher Inquiry

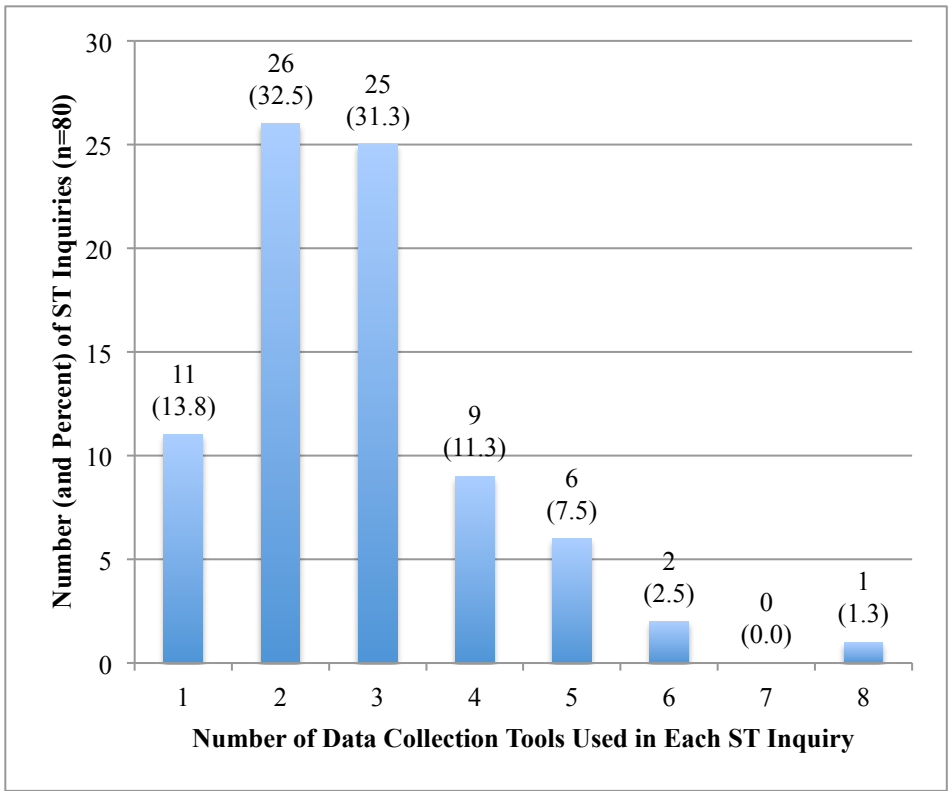
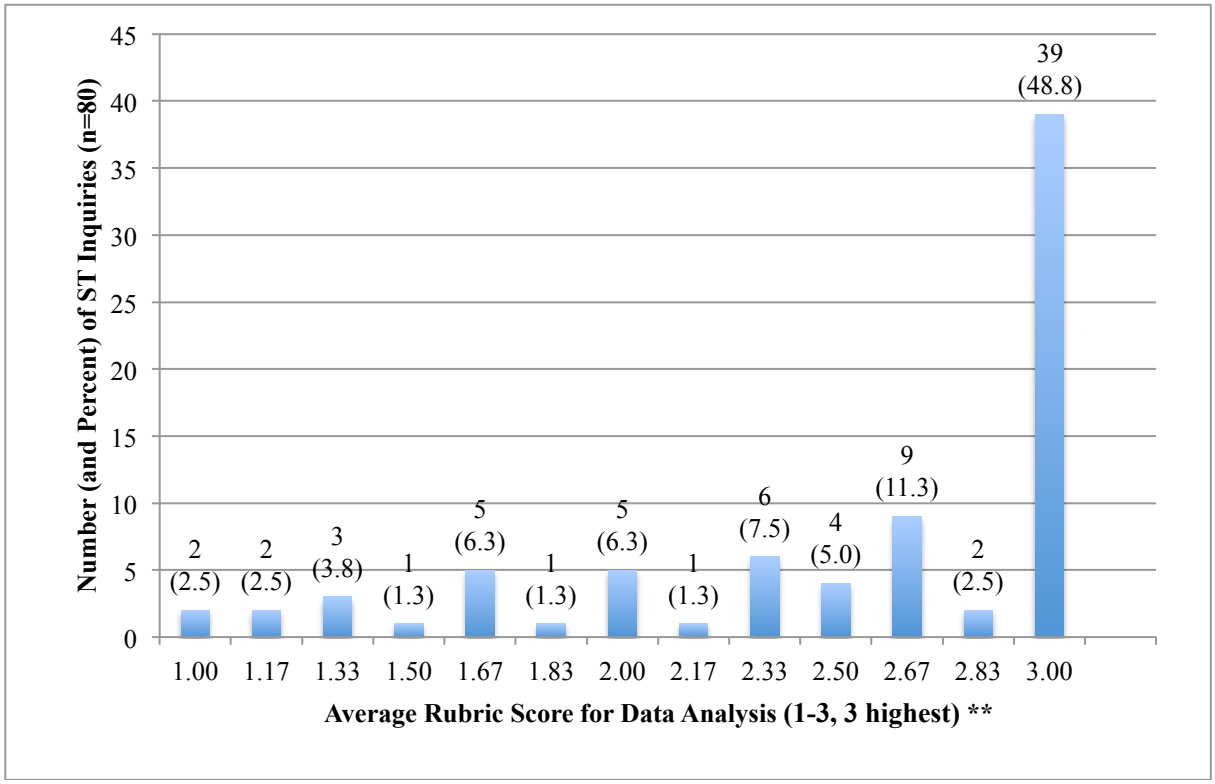


Figure 3

Number of Inquiries Receiving Each Average Rubric Score for Data Analysis *



* Scoring rubric for each data analysis round/event:

3= thoroughly explains how data connect to question; clear example of tool and response; and analysis shows depth and complexity of interpretation
2= connections to research question are not quite clear; or tool and response not quite clear; or analysis could show greater depth
1= problems in 2 or more of these: connections to research question, tool and response, analysis

** Average score calculated from three individual data analysis scores per ST