Title: Supporting Preservice Science Teachers' Ability to Attend and Respond to Student Thinking by Design

This manuscript was accepted by Science Education on May 11, 2015

Hosun Kang[1] and Charles W. Anderson[2]

[1]Corresponding author, University of California Irvine (hosunk@uci.edu)

[2]Michigan State University (andya@msu.edu)

Acknowledgement

Supporting Preservice Science Teachers' Ability to Attend and Respond to Student Thinking by Design

Abstract

A teacher's ability to attend and respond to student thinking is a key instructional capacity for promoting complex and deeper learning in science classrooms. This qualitative multiple case study examines 14 preservice science teachers' (PSTs) responses to learning opportunities created to develop this capacity, as provided by a teacher preparation program. The PSTs engaged in multiple cycles of designing assessments and analyzing student work in coordination with clinical experiences in the field. Drawing upon the notions of responsiveness and noticing, we analyze teaching episodes for whether and how the PSTs in this study attended and responded to student thinking in instructional contexts. Several teaching episodes provide evidence of PSTs' productive responsiveness—suggesting modification in specific elements of instructional design to create better conditions for advancing students' scientific thinking. In general, however, the episodes suggest uneven success in PSTs' responses to student thinking. The findings point to two considerations in designing learning opportunities to enhance PSTs' responsiveness: (a) the use of high-quality assessment tasks that make student thinking visible, (b) helping PSTs to re-frame the problems by de-privatizing PSTs' interpretations of student responses.

Keywords: preservice teacher learning, responsiveness, responsive teaching, student work analysis, and pedagogy of teacher preparation

Introduction

The science education community aims to prepare future science teachers who are capable of promoting all students' deeper learning. Current scholarship in science education advocates complex and deeper learning, characterized by students' robust reasoning, participation in meaningful scientific practices, and sense-making conversations (NGSS Lead States, 2013; NRC, 2007, 2012). To support this kind of learning, teachers must be able to recognize and build on students' ideas and experiences, use students' ways of reasoning as valuable resources, and continuously adapt instruction in response to both the process and progress of student learning (Bransford, Brown, & Cocking, 1999; NRC, 2012; Sawyer, 2006). Attending and responding to student thinking is one core instructional capacity that makes that form of learning possible in classrooms.

Despite the general consensus that developing this instructional capacity is a worthy goal of preservice teacher education (Interstate New Teacher Assessment and Support Consortium Science Standards Drafting Committee, 2002; Kloser, 2014; NRC, 2010), the teacher preparation community has been struggling to figure out how to best support PSTs in cultivating this capacity. The fundamental challenge resides in the fact that we do not yet have well-developed ideas about how and under which conditions PSTs develop this ability, and how their learning progresses throughout their careers. Currently there is little empirical evidence to inform program design and professional learning opportunities that support PSTs in developing this important instructional capacity during the teacher preparation period (NRC, 2010; Windschitl, 2005).

This study intends to fill some gaps in the literature about preservice teacher learning coordinated with a designed learning opportunity. The PSTs of this program engaged in multiple

cycles of designing assessment tasks as a part of planning, implementing plans in their field placement, and collecting, analyzing, and reflecting on student work during their two years in the program. The assessment activities (designing assessment tasks, analyzing student work, and reflecting on their practices) were purposefully designed to draw preservice secondary science teachers' (PSTs) attention to students' scientific thinking and to guide them to respond to it. We examine *how* secondary PSTs responded to this learning opportunity. Specifically, the following research questions are addressed:

1. What was the nature of assessment tasks (items) that were designed or selected by the PSTs?

2. How did PSTs interpret student responses? What did PSTs attend to and how did they make sense of it?

3. What form of instructional change did PSTs suggest (or not)?

4. How were PSTs' attention and responsiveness mediated by assessment activities shaped by their interactions with people in contexts?

We begin by discussing how we conceptualize teachers' responsiveness. The review of previous study provides theories of action behind the pedagogical approach (i.e., engaging PSTs in assessment activities) in relation to responsive teaching. Following the details of our research activities, we present the different responses to this learning activity by the participating PSTs. The variations across the PSTs' responsiveness are accounted for and discussed in relation to the affordances and constraints of the pedagogical approach. The article concludes with some implications and recommendations for designing professional learning programs that cultivate the ability to attend and respond to student thinking.

Conceptual Framework

Teachers' responsiveness and noticing

Teachers' responsiveness is a complex construct to define because responsive teaching involves attending to and addressing multiple problems emerging within a classroom that is populated with diverse students who have different learning needs. Researchers in the fields of science and mathematics education have conceptualized "teacher responsiveness" differently (Elby et al., 2014), and there exists disagreements about what counts as responsiveness in the literature (see Gay, 2000; Hammer, Goldberg, & Fargason, 2012; Rosebery & Puttick, 1998; Sherin, Jacobs, & Philipp, 2011). In this study, we define teachers' responsiveness as the practices of deliberate and ongoing attention and actions that move student learning forward.

This formulation of responsive teaching rests on two premises. First, students have rich nascent resources for reasoning about and making sense of the world around them, and therefore students, even young children, are capable of engaging in complex reasoning and scientific sense-making when appropriate support is provided (Maskiewicz & Winters, 2012; Metz, 1995, 2004; NRC, 2007, 2012). Second, students experience and learn science meaningfully when they are positioned as competent science leaners and their ideas and experiences are recognized, brought forth, and built upon in a supportive learning community (Bransford et al., 1999; Calabrese Barton et al., 2013; Nasir, Rosebery, Warren, & Lee, 2006; NRC, 2005, 2012). In that sense, teachers' responsiveness to students' ideas and ways of thinking, their lived experiences, and their identities as members of multiple cultural and discourse communities is essential to support all student learning (Gay, 2000; Hammer et al., 2012; Thompson et al., In press). In this study, we focus on one dimension of a teacher's responsiveness—attending and responding to student thinking.

Research on teachers' attention and responsiveness to student thinking describes responsive teaching as involving three aspects (see Figure 1). In the first step of responsive teaching, teachers purposefully elicit student ideas (see Hammer et al., 2012). From an instructional standpoint, this step refers to teachers' deliberate efforts to create opportunities for students to show what they know with use of certain forms of assessment (Kang, Thompson, & Windschitl, 2014). Next, teachers interpret and make sense of student responses—how and why students respond in this particular way. Researchers note that a teacher's ability to recognize and interpret the connections between students' ideas and the discipline is critical to being responsive to student thinking (Coffey, Hammer, Levin, & Grant, 2012; Levin, Hammer, & Coffey, 2009). In the final step of responsive teaching, teachers take action based upon these interpretations.

Teachers make their pedagogical decisions based on what they attend to within students' ideas and how they interpret the students' understanding. For instance, in looking over student work, a teacher observes that several of her students are confused about nuclei from the lesson where students made observations of two cells under a microscope (i.e., multiple students understand the nucleus to be a cell in and of itself within the larger cell). The teacher interprets the students' confusion as being related to the limited observations of different kinds of cells within the prior lesson. In this way, this teacher problematizes her own initial instructional design (i.e., not providing sufficient opportunities for students to find patterns across the different types of cells); therefore she modifies the subsequent activities to provide a complementary experience in the following lesson—taking actions to address the perceived problem with regard to student understanding. A responsive teacher continuously engages in this cycle of eliciting, attending, interpreting, and responding to student thinking on the course of instruction at the scale of in-the-moment interaction as well as on a daily basis.

--Insert Figure 1 about here--

Researchers who study the practices of attending and responding to student thinking point out *teachers' ability to notice* as a key mediator that shapes teachers' responsiveness. The assumption is that teachers' pedagogical decisions about actions they take depend on what they notice while interpreting students' responses or situations. In mathematics teacher education literature, van Es and Sherin (2005; 2008) identified three key components of teachers' noticing practices. The first component is identifying "what is important" in a teaching situation. When PSTs interpret students' responses to assessments, for example, they "call out" or "highlight" certain information. The highlighted information shows a PST's act of deciding what is noteworthy and deserves further attention. Secondly, noticing involves using knowledge of the subject matter, of students as science learners as well as of their local context, to reason about events as they unfold. The final aspect of noticing is making connections between specific events and the broader principles of teaching and learning. It requires teachers to categorize and extrapolate from the specific to the general as they respond to the question of, "What is this a case of?" During this process, PSTs (or teachers) connect what they observe in their classrooms to broader principles of teaching and learning, which affects their courses of action.

What one notices is inevitably affected by the structure of expectations about the situation—in other words, the ways in which the situation is framed (Russ & Luna, 2013). A teacher who frames science teaching as working on and with students' ideas is likely to attend to and notice various forms of students' ideas and ways of reasoning, and then act upon them to revise students' thinking. In contrast, a teacher who frames science teaching as delivering canonical scientific knowledge is likely to attend to the correctness of students' responses.

In the fields of sociology (Goffman, 1974), sociolinguistics (Tannen, 1993), and cognitive science (Minsky, 1985; Schank, 1990), framing is conceived as a kind of *schema*, or the "structure of expectations" (Ross, 1975) grounded in one's experience of the world in a given culture (or combination of cultures) (Tannen, 1993). In the field of preservice teacher education, the powerful role of teachers' initial "frame of reference" (Kennedy, 1999) as shaped by their "apprenticeship of observation" (Lortie, 1975) has been well recognized. Kennedy (1999) argues that one important role of preservice teacher education is to change PSTs' initial frames of reference, thereby allowing the PSTs to see situations differently and thus generate different ideas about how they might respond to these situations.

Taken together, the previous studies suggest that a teacher's responsiveness to student thinking is shaped by what s/he notices during interpretation, and what s/he notices is driven by the structure of the teacher's expectations—the ways in which the situation is framed.

Cultivating Preservice Science Teachers' Capacity for Attending and Responding to Student Thinking via Assessment

PSTs in this program engaged in multiple cycles of structured formative assessment tasks during their field experiences. Researchers point out that formative assessment at its core consists of attending and responding to student ideas and reasoning, with roots in disciplinary activity and goals (Coffey et al., 2012; Levin et al., 2009; Sadler, 1998). In the literature, formative assessment is typically referred to as the process by which teachers use evidence of students' learning to modify their teaching to make it more effective (Black, Harrison, Lee, Marshall, & Wiliam, 2004). Effective formative assessment is distinguished by two key features. One is teachers' genuine attention to and engagement with ideas, continuous with the disciplinary practices science teachers should be working to cultivate (Coffey et al., 2012). The

other is teachers' instructional responsiveness as manifested by modification or adaptation of

teaching (Black & Wiliam, 1998; Wiliam & Thompson, 2007). Furtak (2012) reminds us that

"formative assessment hinges on a criterion of *use*, and when information is not used to improve

performance, it is not formative" (p. 1186). Given the nature of the work, engaging PSTs in a

formative assessment process can, in theory, provide scaffolded opportunities for them to attend

and respond to students' scientific thinking.

In the context of preservice teacher education, teacher educators generally perceive

assessment activities as a promising approach to support PSTs' systemic learning for the

following reasons. First, students' written responses produced from pre-planned assessment tasks

can make a variety of student ideas visible (Furtak & Ruiz-Primo, 2008); and, therefore, provide

easy access to student thinking. In theory, PSTs can learn about their students as "sense-makers,"

paying attention to students' ideas and reasoning, and make the needed modification of their

practices based on the evidence of students' understanding. Second, collecting and analyzing

student work outside the classroom may provide opportunities for PSTs to develop new insights

into situations and student learning. Third, PSTs can take time to analyze student responses in

order to plan their actions based on the information that they gain (Atkin, Coffey, Morthy, Sato,

& Thibeault, 2005).

Studies that empirically examine PSTs' engagement in assessment activities, however,

reveal the depth of challenges in helping PSTs to attend and respond to students' scientific

thinking via assessment activities. In a study of 61 PSTs' formative assessment practices in the

context of a semester-long practicum-based assignment, Otero and Nathan (2008) found that the

PSTs tended to attend to either everyday experience-based ideas (what they call "experience-

based conceptions") or to science-based ideas taught in school (i.e., "academic conceptions").

However, PSTs responded only to science-based ideas even when students' everyday

experience-based ideas were elicited. Otero (2006) also found that a "get it or don't" conception

was commonly used by PSTs when they engaged in formative assessment, with serious impacts

on their instructional practices. Much of novice teachers' knowledge about assessment is

"underdeveloped" (Maclellan, 2004), which makes it difficult for them to make sense of student

responses that require complex reasoning with evidence (Lyon, 2013).

Teaching is a performance, and it is even more difficult to help PSTs develop the ability

to respond to student thinking *in action*. Teachers make pedagogical decisions based on their

interpretations of classroom situations (Kennedy, 1999). A teacher's choice of action at any

moment is inherently responsive to the perceived situations and recognition of problems that

need to be addressed. An important question to teacher educators is whether and how teachers'

choice of actions enhances students' learning by creating better conditions for intellectual and

social interactions. Productively responding to students' scientific thinking involves what Schön

(1983) calls, "reflection-on-action"—teachers spending time exploring why students responded

as they did in in relation to disciplinary learning goals and the circumstances of provided

learning opportunities. It requires PSTs to *reframe* a problem and integrate knowledge about

teaching and learning, which was rarely observed when PSTs engaged in assessment activities

(Lyon, 2013). Furthermore, the complex nature of student teaching, a context typically populated

with multiple perspectives and expectations from both the school and the program, is not always

conducive to PSTs learning how to respond productively to students' scientific thinking. In fact,

it is often impossible for PSTs to make actual changes in instructional design on the following

day or try out new strategies in someone else's classroom.

Research Design

Study Context

The context for this study was a reform-oriented five-year undergraduate teacher preparation program. The PSTs in this program took four field-based, disciplinary-specific methods courses during the last two years of the program, the senior (4th) and internship (5th) years. The historical relationship between the program and local school communities provided relatively strong leeway for the program to structure PSTs' experiences in the field. The program engaged PSTs in mandatory teaching responsibilities and assessment activities at designated times. This institutional context made it possible for the methods course instructors to design deliberate learning opportunities through assignments coordinated with experiences in the field.

PSTs engaged in about eight teaching cycles—planning, enactment, assessment, and reflection—throughout the two years with coaching from instructors and field supervisors. In order to develop the ability to respond to students' scientific thinking, the teaching cycles included an assessment component with three sub-activities: (a) designing assessment tasks; (b) interpreting student responses; and (d) suggesting changes in instruction (see Figure 2). Each phase was guided with tools (e.g., template, rubric) and various scaffolds to draw PSTs' attention to students' scientific thinking and help them to learn how to respond to it. The following section discusses the connection between each activity and aspects of responsiveness illustrated in Figure 1 as well as the underlying assumptions behind this pedagogical approach.

--Insert Figure 2 about here—

Designing assessment tasks—eliciting. Assessment tasks provide information about what students know and are able to do (Kang et al., 2014). The requirement to design an assessment task forces PSTs to think about the outcomes they would expect to see if they

accomplish their instructional goals. The design of the assessment tasks generated from this process inevitably reflects how PSTs frame knowledge and learning of science when they design or select the assessment.

The PSTs of this program either designed or selected two to three assessment tasks at the planning stage of each teaching cycle. Typically, PSTs interacted with methods course instructors, peers, and their mentor teachers during this time to search for resources and ideas for their assessment tasks. The program provided three kinds of support to assist PSTs' planning of high-quality assessment tasks that make students' scientific thinking visible. First, the program provided a template that included rubrics about the quality of assessment tasks and detailed prompts to assess students' scientific thinking. In addition, the instructors posted examples of quality assessment tasks on the course website. Lastly, PSTs received individual feedback from the methods course instructors regarding their assessment tasks (items), including suggestions for modifications before enactment.

Interpreting student responses—attending and noticing. After implementing their lesson or unit plan in the field, PSTs collected and analyzed student work. During this phase, PSTs interpreted student responses while accounting for how and why students produced particular responses. PSTs were expected to attend to and notice various forms of student ideas and thinking such as partial understandings or alternative ideas during this process.

The course instructors required PSTs to select and examine responses from three focus students' at different academic achievement levels, except within the last two teaching cycles. The intention was to engage PSTs in an in-depth inquiry about student thinking rather than in superficial levels of analysis, such as "get it or don't get it." In addition, PSTs discussed their analyses of students' responses with their peers and experienced teachers (field supervisors)

before they produced a written report. A group of five to six PSTs who were under similar student teaching conditions (subject taught, grade level, school, etc.) collectively looked at the produced student work for about 90 minutes with the facilitation of a course instructor or a field supervisor. Finally, the detailed prompts and rubrics in the template guided PSTs' systematic analysis of students' scientific thinking.

Suggesting changes in instruction—responding. The PSTs suggested changes to their instruction based on their interpretations of student responses while producing a written report. The prompts provided in the report template explicitly instructed PSTs to suggest specific changes to different components of their instructional design that might address the framed problem. The caveat to this approach is that PSTs' suggested changes to instruction did not necessarily reflect the actual instructional adaptation or PSTs' ability to enact them. Further, PSTs might have simply tried to "please" the methods instructors to meet their expectations when interpreting student responses and suggesting changes in instruction. However, it was unlikely that PSTs could respond to students' scientific thinking productively in action without first being able to thoroughly interpret student responses and make deliberate decisions based upon those interpretations.

Participants

This investigation uses a multiple case study approach (Stake, 2004; Yin, 1989). We selected participants in a way that maximized the variation among cases to enhance transferability (Merriam, 2009). In this program, each cohort consisted of about 30 secondary science PSTs. The majority of the PSTs were college students with no formal teaching experience. Typically, about one third of PSTs had some experience working with scientists in a laboratory setting, either as undergraduate research assistants or in a graduate program. Each

cohort included one or two PSTs who chose teaching as their second careers. During the two

academic years (2008-10) of this study, we interacted with three cohorts of PSTs who were at

either the intern (5th year) or senior (4th year) stages. Between four and six PSTs were selected

from the volunteers in each cohort based on three criteria: (a) a spectrum of personal

backgrounds (e.g., major, gender, content area, teaching experiences, research experiences in

science, and career choice); (b) school contexts (e.g., suburban academic-oriented schools vs.

urban high-need schools); (c) mentor teachers' relationship with the program. PSTs' personal

backgrounds, and in particular their prior teaching experiences, were considered because of their

potential influence on the ways in which PSTs frame the work of science teaching and learning.

School contexts and mentor teachers' relationships with the program were also considered for

the same reason.

In Year 1 (2008-09), a total of eight PSTs were selected from two cohorts (see Table 1).

The first group of four PSTs from cohort I were interns (5th year), and the second group of four

PSTs, from the cohort II, were seniors (4th year) during Year 1. Two of them continued to

participate in the Year 2 (2009-10) as they proceeded to their internship year. One opted out of

this study, expressing his difficulties in managing time during his internship year. One was

placed at an affluent suburban school for her internship year that was located a long distance

from the research site, with new mentor teachers who had not worked with the program. We

decided to replace these PSTs with two other interns, Adam and Alice. Adam's mentor was a

graduate of the program, and Alice's mentor had been working with the program for several

years. In the second year of this study, in addition to Adam and Alice from Cohort II, another

group of four PSTs were selected from a new cohort (Cohort III). These PSTs only participated

in this study during their senior year.

In addition to PSTs, 12 mentor teachers and two course instructors participated in this study. Among the 12 mentor teachers (three PSTs worked with the same mentor in different years), five graduated from the program and the other four had previously been working with the program as mentor teachers. Three mentor teachers were newly recruited teachers who did not have any relationship with the program. At the end of the year, one of the new mentor teachers, Ms. S., was not recommended to return. Her field supervisor cited management/behavior-oriented teaching and inflexibility as the reasons for the recommendation.

--Insert Table 1 about here--

Sources of Data

The primary source of data was 32 *sets of teaching episodes* generated from 32 teaching cycles taught by 14 participating PSTs. A teaching episode included lesson or unit plans, the descriptions of assessment tasks, worksheets or slides that showed the actual assessment tasks (items), samples of student responses, and PSTs' written analysis and ideas for improvement. Each year we asked the participating PSTs to give us at least one teaching video recorded during their teaching episodes. Half of the teaching episodes included teaching videos (one to two teaching videos per one PST). These data provided information about PSTs' assessment design, interpretations about student responses, and their ideas for instructional modification.

A second major source of data was *interview transcripts*. A semi-structured interview was conducted with the participating PSTs individually at the end of each year. During the interview, the interviewer showed a pre-selected 3-5 minute long segment of the PST's teaching video. Selected clips illustrated students (either individually or in groups) engaging in the tasks and interacting with the PST. During the interview, the interviewer prompted the PST to assess students' general responses to their instruction, and then to assess one or two particular students'

responses. Some examples of questions were, "What was your goal for this lesson," "How did the students of this class do with the content," "Did you notice any difficulty that students were having? Where do you think those difficulties came from," and "What would you do if you taught this lesson again?" The analyses of interview transcripts helped us to triangulate the patterns that emerged from the written teaching report (Denzin, 1978; Denzin & Lincoln, 2005; Merriam, 2009). In addition, the interviewer asked about PSTs' working relationships with their mentor teachers, specifically in designing their assessment tasks. This interview provided insights into PSTs' experiences in a local school context, and in particular their interactions with mentor teachers.

Lastly, we conducted individual *interviews with the two course instructors and 12 mentor teachers*. These were semi-structured, hour-long interviews conducted near the end of the academic year. Similar to the interviews with PSTs, the researcher showed segments of PST teaching videos (the same ones used in the PST interviews) and asked similar questions. These comparison interviews were analyzed to understand the interviewees' relationships with their PSTs as well as the ways in which these significant professionals, who had regularly interacted with the PSTs, framed the goals of science learning. Some examples of questions were, "How do you usually teach this topic and why," "What do you think the PST is trying to accomplish and how do you think about it," "What do you like or dislike about PSTs' approach and why," and "How did you work with your PST when they planned their instruction and after their instruction?"

Data Analysis

At the first stage, we analyzed PSTs' engagement in each of the assessment activities in response to research questions (RQ) 1 to 3. Guided by the responsiveness framework and given

the nature of assessment activities (see Figure 1 and 2), we coded PSTs' assessment tasks, interpretation, and reflection as documented in the 32 teaching episodes, focusing on their attention and responsiveness. Next, cross-analyses were conducted across the 32 teaching episodes and 14 PSTs to examine (a) under which conditions PSTs successfully responded to student thinking, and (b) what made PSTs more or less successful in responding to students' thinking. The following describes the details of the processes.

Phase I: Coding PSTs' assessments, interpretation, and responses

Assessment tasks: epistemic frame built in the design (RQ 1). We analyzed the nature of the assessment tasks in each teaching episode holistically. We examined written descriptions of the goals, objectives, and designed/selected assessments, as well as the teaching artifacts (worksheets, slides), and teaching videos (if available) in each of the 32 teaching episodes. Our analysis was focused on explicating the nature of the opportunities PSTs provided for students to show both what they know and how they know through assessment tasks. The nature of these opportunities revealed how PSTs framed knowledge and learning of science as they designed or selected assessments in that instance.

Based on this analysis, we coded the assessment tasks as either productive or unproductive. Following recent influential documents (NGSS Lead States, 2013; NRC, 2007, 2012), we posited a *productive* epistemic frame as providing opportunities for students to engage in meaningful scientific practices for sense-making. We identified a total of four types of opportunities from our analysis of the assessments. Two types of opportunities were coded as a reflection of productive epistemic frames. In those opportunities the completion of the assessment tasks (items) required students to make connections between observable and unobservable (theoretical) elements of natural phenomena through intensive reasoning.

Specifically, the assessment tasks prompted students to (a) construct scientific explanations or argumentation by reasoning through data, observation, or experiences and (b) use science ideas to account for observable phenomena. The assessment tasks coded as unproductive asked students to either (c) reproduce factual information or canonical scientific knowledge or (d) display skills or procedural knowledge (see the coding scheme and examples of assessment tasks in Table 2).

--Insert Table 2 about here--

Interpreting student responses: attending and noticing (RQ 2). In this analysis, our goal was to investigate PSTs' interpretations of student responses, focusing on their attention and noticing. First, we identified documented student responses and the accompanying PSTs' accounts. Sometimes PSTs drew inferences about individual student responses; at other times, PSTs drew inferences about a pattern of collective student responses. We defined a PST's account, generated around either one individual response or a grouping of responses, as the unit of analysis. We identified a total of 154 units from 32 teaching episodes (4-5 units per one teaching episode). We analyzed each unit to identify what aspect(s) of student responses PSTs attended to. Building on the previous study (Kang & Anderson, 2012), these accounts were then categorized into three groups, attending to cognitive behavior, social behavior, and on-task behavior (see Table 3). When a PST attended to multiple student behaviors, all of that information was coded into a unit.

Next, we analyzed PSTs' noticing—how PSTs reasoned with the attended information to account for how and why students produced certain responses. During this process, PSTs *generated relationships* by connecting bits of attended information to one another. Eight initial codes emerged from the analysis of 154 units representing the relationships generated in this

process. Specifically, PSTs related the attended student responses to (a) aspects of instructional design; (b) students' missing experiences that they didn't take into account, (c) failure to build upon students' prior experiences and knowledge; (d) home or family backgrounds, languages, and cultural resources (e.g., foster care, immigrant family); (e) social interactions in classrooms; (f) students' personalities, characteristics or working pattern; (g) students' attentiveness, attitude, or behaviors; and (h) problems of missing school knowledge prior to the year or misconceptions. We categorized these eight initial codes into three groups reflecting the ways in which PSTs *framed problems*. The three groups were (a) problems of instruction, (b) problems of learning environments, and (c) problems of students (see the final coding scheme in Table 3).

     Suggesting changes in instruction: Productive vs. unproductive responses (RQ 3). In many teaching episodes, PSTs suggested changes collectively at the episode level. Therefore, the final analyses of PSTs' responses were coded at the teaching episode level. We coded the quality of PSTs' responsiveness to student thinking, as reflected in the suggested changes, as either productive or unproductive. Based on the literature on responsive teaching, instructional design (Coffey et al., 2012; Levin et al., 2009; Roseberry & Warren, 2008; Sohmer, Michaels, O'Connor, & Resnick, 2009) and the patterns that emerged from our data, we defined *productive responses* as specific changes in components of instruction (e.g., task design, talk, tools, or scaffolds) that likely create better conditions to promote students' scientific thinking.

     When a PST's responses were coded as productive, the changes suggested by the PST were specific and relevant—meaning they directly addressed the students' difficulties. For example, when a teacher sees that some students fail to notice important patterns across observations, a productive response might involve revising or adding questions to draw students' attention to the important patterns (changes in talk moves). In contrast, *unproductive responses*

do not generate better conditions for student learning. For example, suggestions like "re-teach"

or "do more labs" in a generic sense are unlikely to address students' particular difficulties.

Typically, an *unproductive response* was generic or irrelevant to the manifested student ideas.

Five codes emerged through the initial analysis, of which four were categorized as unproductive

responses (see Table 3). Each of the 154 units were coded using the five sub-codes.

--Insert Table 3 about here—

Phase II: Cross-episode analyses—examining PSTs' engagement in assessment

activities in contexts (RQ 4). We examined the relationship between assessment task design,

interpretation, and PSTs' responsiveness within and across the 32 teaching episodes to identify

the conditions under which PSTs productively attended and responded to students' scientific

thinking. We identified twelve teaching episodes that provided evidence of PSTs' productive

responsiveness. With these 12 episodes, we traced back how PSTs' responsiveness was related to

the patterns of assessment design and PSTs' interpretations. This analysis suggested some strong

relationships between PSTs' productive responsiveness and the epistemic frame of the

assessment tasks. Accordingly, we selected 18 teaching episodes that began with a productive

epistemic frame, compared and contrasted PSTs' patterns of interpretations, and analyzed how

those patterns were related to PSTs' productive or unproductive responsiveness. Finally, we

examined the patterns within and across the remaining 14 teaching episodes that began with an

unproductive epistemic frame, focusing on the PSTs' responsiveness.

Next, we shifted the grain size of the analysis from teaching episodes to the individual

cases of the 14 PSTs. We categorized the 14 cases into three groups based on their

responsiveness to student thinking. Employing the constant comparison method (Glaser &

Strauss, 1967), we identified and compared the patterns within individual cases in each group, as

well as across groups. We paid attention to other contextual information provided through

interviews to understand how PSTs' assessment practices were shaped through their interactions

with mentor teachers and teacher educators within contexts. Interview transcripts were analyzed

focusing on the following questions: (a) how did PSTs interact with people at both the school

and the program when they designed their assessments, and what resources were used by PSTs?

(b) how were PSTs' interpretations of their assessments and student responses similar or

different from mentor teacher's and course instructors'? (c) what were the contextual affordances

or constraints for PSTs' productive responsiveness to student thinking? From these analyses, we

intended to theorize PSTs' learning of responsive teaching in relation to the pedagogical

approach of the program.

Subjectivity

There is a potential for bias in the process of collecting and interpreting data due to the

relationships of the authors with the participants. The authors were either course instructors or

field supervisors of five PSTs. On the other hand, the relationships provided some deeper

insights into the nature of local school contexts, professional relationships and interactions with

their peers, mentors, and instructors, and their daily instructional practices in classrooms. We

addressed this issue of potential bias with multiple layers of triangulation (Denzin, 1978) and the

careful design of data collection. None of the authors were involved in the interviews with the

participants who were under their supervision. Multiple sources and types of data collected

through multiple methods were used to increase credibility. The generated coding scheme and

interpretation were discussed and debated at weekly meetings for a total of two years. The

codings, interpretations, and pictorial models were also presented and discussed with three other

science education faculty members through a bi-weekly instructor meeting for about six months

at the early stage of data analysis. In addition, a formal inter-rater reliability check was

conducted with a doctoral student in science education program. We coded a sample of data

together and revised the coding scheme iteratively until we reached up to 80% consistency.

Findings

We present our findings in four sections, reflecting what we learned about the assessment

tasks themselves and the ways in which PSTs interpreted and responded to students. The last

section presents three illustrative cases to show how PSTs' attention and responsiveness to

student thinking via these assessment activities were shaped in contexts.

Assessment Tasks

About a half of the teaching episodes used the productively framed design of assessment

tasks that provided opportunities for students to reveal their sense-making (n=18 out of 32,

56.3%). Among these 18 teaching episodes, two thirds of the episodes provided evidence of

PSTs' productive responsiveness to student thinking (12 out of 18, 66.7%; see Figure 3). There

was no incidence of PSTs showing productive responsiveness to student thinking when

unproductively framed assessments were used. About half of the productively framed episodes

were observed from interns and the other half from seniors (8 from interns, 10 from seniors).

10 out of 18 productively framed assessments were either designed or significantly

modified by the PSTs with help from the program (e.g., using the provided lesson design

framework and templates, using assessment tasks from the course websites with a slight

modification, or co-designing the tasks with course instructors). The assessment tasks of the

other eight teaching episodes came directly from mentor teachers or some commercial curricula

with little modification. Notably, nine out of 12 teaching episodes that provided evidence of

PSTs' productive responsiveness to student thinking were either designed or significantly modified by PSTs. Two of the remaining three were from curricula developed by university-based research groups.

--Insert Figure 3 about here—

Interpretation of Student Responses

Recall that we defined a "unit" of analysis as a set of accounts generated by a PST around a single or collective student response. Using this definition, we identified 154 units across the 32 teaching episodes. However, the total number of information units highlighted by PSTs was 169, because many PSTs attended to multiple pieces of information while simultaneously interpreting student responses (see Table 4). Overall, PSTs attended to on-task behavior most frequently (n=85, 50.3%), followed by cognitive behavior (n=79 out of 169, 46.7%) and social behavior (n=5, 3.0%).

It appeared that PSTs' attention was strongly associated with the nature of assessment tasks. As shown in Figure 4, when productively framed tasks were used, PSTs attended to students' cognitive behavior about three times more than when unproductively framed assessment tasks were used (73.2% vs. 26.5%; see Figure 4).

--Insert Table 4 and Figure 4 about here--

PSTs linked the attended information to various aspects relevant to students' learning. In this process, about 70% of the attended student responses were connected to a problem with students, such as missing school knowledge, students' misconceptions, attentiveness, attitudes, personalities and task behaviors, etc. (69.8%). About one fourth of the attended information was attributed to a problem with instructional design (25.8%). In a few units, PSTs connected student

responses to aspects of the context, such as personal, home, family backgrounds or social relationships and interactions in the classroom (4.4%; see the detail in Table 5).

--Insert Table 5 about here--

The ways in which PSTs made sense of student responses was strongly related to both the frame of the assessment task design and things that PSTs attended to. When productively framed assessments were used, PSTs connected students' responses to problems of instruction eight times more frequently (40.2% vs. 5.2%). In contrast, when unproductively framed assessment tasks were used, 92.2% of the attended information was connected to problems with students. Under the condition of attending to cognitive behavior, PSTs connected the attended student responses to problems of instruction about four times more than the condition of attending to on-task behavior (39.8% vs. 10.1%). In contrast, under the condition of attending to students' on-task behaviors, 89.8% of the attended information was linked to some problems with students.

Suggesting Changes in Instruction: Productive vs. unproductive responses

We analyzed the 12 teaching episodes that provided evidence of PSTs' productive responsiveness to student thinking. This analysis showed that both what PSTs attended to and the ways in which they framed problems while interpreting student responses played important roles in shaping their responsiveness. When PSTs attended to students' cognitive behavior, the likelihood of suggesting productive responses to student thinking was 67.4%. In contrast, it was only 14.1% if PSTs attended to on-task behaviors. When PSTs linked the attended information to problems relevant to instruction, they were three times more likely to make suggestions for productive instructional modification than when they linked the information to problems with students (66.7% vs. 23.8%).

PSTs Attention and Responsiveness to Student Thinking via Assessment Activities in Contexts

The cross-analyses of 14 PSTs' cases revealed three distinctive patterns in the ways in which PSTs attended and responded to student thinking upon their engagement in assessment activities: (a) consistently attending and responding to student thinking (3 PSTs); (b) conditionally attending and responding to student thinking (4 PSTs); and (c) responding to other concerns (7 PSTs) (see Table 4). In the following we present an illustrative case from each of the three groups.

The Case of Leslie: Consistently attending and responding to student thinking with use of productively framed assessment tasks.

Leslie consistently paid great attention to students' ideas and thinking and either suggested or adapted her teaching practices to address students' difficulties instructionally across all four teaching episodes in both her senior and internship years. In her senior year, Leslie was placed in a 9th grade biology class at an urban, high-needs school that had a high percentage of ethnically, racially, and linguistically diverse students from low-income families. Leslie's mentor teacher, Mrs. F, had worked with the program for seven years. Mrs. F's instruction was fairly traditional—some combination of worksheets and hands-on activities, but she was supportive and open to new ideas. During the interview, Mrs. F described Leslie as an "active", "strong", and "confident" candidate who actively searched for useful resources and asked a lot of questions. Mrs. F. said, "[Leslie] would come to the discussions that she would have with some suggestions or ideas already prepared." The course instructor, Dr. G, described Leslie as "one of our top candidates" who was "really attentive to [the assignment] templates." Dr. G commented, "If we made comments, [she was] coming and asking questions about comment, even revising

her unit or lesson plans, even if technically she didn't need to revise it. She followed the templates that we laid out very carefully."

Designing assessment tasks: framing science as engagement in meaningful scientific practices for sense-making. Leslie was assigned by her mentor teacher to teach the topic of chromosomes and the structure and function of DNA (Leslie's teaching episode #2 in Table 4). This lesson was anchored in one puzzling question, "If the DNA from a human cell was stretched out, it would be over 6ft tall! How do you think that a cell fits it all in?" Leslie began the lesson by asking students to think about the best way to pack a lot of clothes into a small suitcase. After eliciting students' ideas like "rolling" and "folding" through discussion, Leslie provided information about the process of folding DNA into chromosome with various pictures. Students labeled the pictures individually or in a group as walking through the worksheet. Next students built models of two nucleosomes with candy. Using this model, they explained the process of folding from the DNA string to chromosomes.

Leslie designed four assessment questions. First, she asked students to (a) "Label the picture. Be sure to use the following words: Histone (protein), DNA double helix, Chromatin, and Chromosome." Then, students had to answer three open-ended questions: (b) "Explain how DNA goes from a very long thin molecule to the thick structure of condensed chromosomes" (c) "What are two reasons for folding DNA into chromosomes?" (d) "Next week we will be learning about when cells spilt in two. Why might chromosomes be important in the splitting process?" The design of this assessment provided opportunities for students to engage in modeling, and reason with both the information and the model to solve the puzzling question. The epistemic frame built in the design of this assessment task was coded as "productive" because science learning was framed as engagement in meaningful scientific practices for sense-making.

Interpreting student responses: attending to student thinking and framing the problem in relation to instructional design. Students produced various interesting responses to this assessment task in both drawing and written form. Leslie attended to the details of student ideas and thinking, and identified their difficulties in sense-making (see the coding results of teaching episode #2 in Table 4). She then framed the problems by connecting them back to her instructional approach (problem of instructional design). For example, Leslie stated, "About half of my students thought that DNA turns into histones. In addition to the confusion around histones, I noticed that about 10% of my students thought chromatin and chromosomes were structures that DNA wrapped around. My students realized that chromosomes result from a series of folding and rolling steps, but they were confusing which step was which." Leslie then connected the student's confusion around histone with her unclear representation during the instruction, stating, "There were a few things that I did during my lesson that might have confused them. For example, I showed a tangled mess of yarn and a ball of yarn to represent DNA being wrapped around a histone. The students never saw the bead the yarn was wrapped around, so they might have thought that the yarn itself represented the histone. If this was the case then it is easy to see why so many of my students thought that DNA turned into a histone."

Suggesting changes in instruction: A specific modification of task design and scaffolding. Leslie's responses were specific and tied to the difficulties that she identified in students. Leslie suggested two specific changes in components of instruction: "have students work together to draw a picture on the chalkboard or build a model that represents DNA at each stage of folding" (adapting task design) and "in the discussion, I would emphasize that a histone is a protein and the only structure that DNA wraps around" (purposeful talk moves). She stated,

If I were teaching the next day, I would do a warm up question about DNA folding and then have a class discussion reviewing each structure. I would say something like "most of you seem to understand that strands of DNA are rolled and folded to make chromosomes, but many of you are confusing what each structure is composed of. Let's review this again." In addition to this I might have my students work together to draw a picture on the chalkboard or build a model that represents DNA at each stage of folding. In the discussion I would emphasize that a histone is a protein and the only structure that DNA wraps around. I would also emphasize that chromatin and chromosomes include the DNA in their structure.

Across all four teaching episodes taught in her senior and internship years, Leslie continuously used assessment tasks that provided opportunities for students to engage in meaningful scientific practices and asked questions that generated complex and long responses. Leslie's reflection included specific ideas for changing components of instruction that were directly related to the manifested student difficulties.

The Case of David: Attending to on-task behaviors and addressing students' wrong or incomplete responses with generic motivational strategies.

David was one of the seven PSTs in our "responding to other concerns" group. He was one of three PSTs who continuously used assessment tasks that were coded as *unproductive*. During his internship year, David worked with an exemplary mentor, Mrs. M, in an academically-oriented high school near the university. Mrs. M was a veteran teacher who graduated from the same program 12 years ago. She was a highly regarded mentor teacher who provided PSTs with a good model of science teaching. During the interview, Mrs. M expressed her frustration in working with David throughout the internship year. Mrs. M said, "The planning

portion, I didn't really feel like it was a strong relationship" and "Just in terms of behavior, he was very closed. In terms of what you would see, his physical stance was with his arms crossed. Kind of staying at a further distance from me, and just saying, 'Okay, okay, okay,' and not really asking the 'why' part.' It was very short, like, 'I've heard you' and that was it." The course instructor, Dr. R described David's engagement in her methods course using the words like, "so flippant about things," "he is a joker," and "confident about his ability." Dr. R commented, "I think [David] thought he could be a fine teacher. I don't think he thought the program would do him a whole lot of good, you know. I think that he thought of himself as somebody who was a pretty good teacher coming in." Ironically, during the interview David commented that his instruction was pretty similar to Mrs. M and there was no big difference except "she has more stories to tell" due to her experiences.

Designing assessment tasks: framing science as reproducing canonical scientific knowledge. This teaching episode was focused on the human digestive system in 9th-grade biology (David's teaching episode #1 in Table 4). This unit was anchored on the phenomenon of a man eating a hamburger—What happens to a hamburger as it goes through the digestive system? While planning this unit, David selected following two questions as his assessments: (a) "Label a human body silhouette picture with the organs on it" and (b) "Name two organs that digested a certain macromolecule." The key objective of this unit as identified by David was to "identify the correct location and relative size of the organs in each body system, and explain each organ's general function within its specific system." This assessment was coded as an "unproductive frame" because it only showed whether students were capable of reproducing facts or known scientific ideas.

Interpreting student responses: attending to task behaviors and connecting it to the problem of students. While analyzing student work, David primarily attended to correctness of student responses and students' on-task behavior (e.g., completing vs. not completing the work). For example, David narrated Pat's (one of his low-achieving students) responses as follows:

> Pat got all but three of them correct and for some reason she still labeled the bladder as part of the system. I am a little confused here because we haven't even talked about the urinary system yet and I made it a point to say that urine was not involved with the digestive system and that it only creates solid waste. This student is definitely more talkative in class, so there is a chance that she may have missed this. The only other organ she missed was the gall bladder which we didn't spend much time on other than just saying it was up under the liver and it stores the bile created by the liver.

David connected Pat's incorrect answer to her attentiveness to the instruction. Through the analysis of student responses, the main problem was framed as students' attentiveness, motivation, and behavioral issues (i.e., problems of students).

Suggesting changes in instruction. To improve the instruction for this topic, David suggested adding "some sort of activity or project."

> I really liked my lesson sequence for this unit and think that it went well…the only thing I would change is that for the final assessment of this lesson sequence I might have them do some sort of activity or project with it where they maybe make a working model of the digestive system or something cool like that.

Despite the potential for increasing students' attentiveness to the task with this hands-on activity, the instructional modification of "doing some sort of activity" was less likely to address

Pat's difficulties in understanding why the bladder is not part of the digestive system. This kind of positive and overall satisfactory comment about the lessons was another salient pattern among the PSTs in this *Responding to other concerns* group. Typically the suggested ideas for adapting instruction were generic or too general, and were not related to the students' manifested intellectual difficulties.

The Case of Monica: Attending to students' social behaviors and responding with general strategies of action.

Monica was one of three PSTs who was successful in providing evidence of responsiveness in one of two teaching episodes—the one she taught at the later stage of the year. There were noticeable differences between the two teaching episodes (see Monica's #1 and #2 in Table 4). One difference was her changes in the design of assessment tasks to provide new forms of opportunities (from an unproductive to a productive frame). The other was Monica's increasing attentiveness to students' scientific thinking with the use of the productively framed assessment tasks. But across the two episodes, Monica consistently framed problems in relation to her instructional design.

Throughout her time in the program, Monica had struggled to teach science "differently," meaning "not lecturing all the time." She said, "Up until college, none of my classes really focused on inquiry or application as far as I can remember. Like, as a student, you're not really thinking about that." During the interview, Monica commented, "I get tired of lecturing and [students] don't like lecturing. My first lesson I taught to seniors, I lectured the entire hour. Their teacher must have threatened them because they were really good."

In her internship year, Monica worked with Ms. S who was fairly traditional in her instructional approach and less supportive of the program's approach. Monica aspired to create

an "emotionally supportive but cognitively challenging" classroom learning community where "everyone can be successful." Monica's mentor, Ms. S said, "[Monica] was quite a researcher, so she was really good about going out and trying to find something new." Monica actively interacted with the methods course instructors, field supervisors, and her peers to receive feedback on her plans. Her course instructor, Dr. R said, "Monica asked a lot of questions both before she would plan, but also while she was teaching. She would sometimes write me an email, or when we saw each other in class, she really wanted to talk about some things."

Designing assessment tasks: from knowing canonical science ideas to engaging in sense-making. The topic of her 1st teaching episode was osmosis and diffusion within a cell in a 9th-grade biology course. The assessment used in this episode consisted of five true/false questions, three short answer questions, and one multiple-choice question. One key assessment question that Monica "really hoped [students would] understand by the end of unit" was a multiple-choice question about the mechanism of osmosis. In the pre-assessment, students were asked to choose one correct prediction about the movement of water inside a U-shaped tube over time, from among four choices. In the end-of-unit assessment, Monica asked students to mark the status of the water inside a U-shape tube at equilibrium. Both assessment tasks were coded as part of an "unproductive frame" because, by design, the assessments revealed whether students knew the canonical science ideas (osmosis) that were covered in the lecture.

In the other teaching episode (Monica #2 in Table 4), Monica taught a complex inheritance unit, following Mendelian genetics in 9th grade. During the sequence of activities, students made observations of the five pictures showing that offspring displayed a mixture of parents' traits as opposed to displaying a single parent's trait—the five examples of complex inheritance from students' everyday lives. Students discussed their observations to formulate

questions or hypotheses, and engaged in a modeling activity that illustrated incomplete or co-dominant inheritance patterns. The two key assessment questions that appeared on the worksheet were as follows: (a) What are some differences that you see between this type of inheritance and the dominant/recessive pattern you saw last week? (b) Can you think of any examples of this type of inheritance that you've seen before? If so, what have you seen? In addition, students had to go back to the questions that they formulated at the beginning of this unit and answer those questions based on what they learned from the modeling activities. This assessment was coded as productive because it provided opportunities for students to engage in scientific practices (pattern finding), and to use this idea to make sense of complex real-world phenomena.

Interpreting student responses: attending to learning environments that affect students' engagement and framing the problem in relation to learning environments. The analyses across teaching episodes suggested that Monica was very attentive to students' personal backgrounds, language, and social relationships both in and outside her classroom. She drew upon that information to make sense of student responses. Monica's report, especially for the 1st teaching episode, included a variety of information about the three focus students. For example, Monica noted that Ken, her high achieving student, moved to this school from Japan a year ago, and had been struggling with scientific terminology despite his proficiency in English. Monica stated about Tom, who was a lower academic achiever, "Although my initial opinions were based on his standardized test scores, Tom is a very bright student who expresses his love of science—struggles to stay focused in social situations, but an active participant in class overall."

While making sense of students' responses produced from her assessment questions, Monica attended to students' cognitive and social behavior, and she connected the information to some aspects of the learning environment offered to the students. For example, in the post

assessment about osmosis, all the students except Tom failed to generate a scientifically accurate

prediction of the change of water level. Tom correctly drew the change of the water level, but

showed that all of the particles moved to the same side, which was contradictory to his prediction

about the changes of the water level. Monica noted, "Tom was the only one to make the water

level higher on one side than on the other, but he also showed that all of the particles moved to

the same side as well. I'm not sure if he thinks that when water moves, the particles must travel

with it, or whether this was just a guess." Monica also connected Tom's responses to some

personal factors outside of the classroom ("recently placed in foster care," "the number of class

days he missed") and then problematized the general instructional approach (i.e., lab vs. lecture):

> Tom seemed to struggle a bit on this assessment. He was recently placed in foster care
>
> and I think the number of class days he missed drastically affected his results. In general,
>
> he seemed to struggle most on the application questions that not only used information
>
> from this unit but also required students to draw upon knowledge from prior units…In
>
> general, Tom seemed to excel on questions that he had the most experience with (labs,
>
> concepts we discussed frequently in class). The concepts that he learned only in lecture
>
> format however (active/bulk transport) he seemed unprepared to discuss.

In contrast, Monica's interpretations about student responses in the Genetics teaching

episode was much more detailed and attentive to cognitive behavior relevant to scientific sense-

making. For instance, Monica noticed students' difficulties in differentiating observation from

inference, and its consequential effect on their sense-making: "From what I observed, many of

the students were making inferences rather than observations. For example, instead of stating

that the flower contained both pink and white petals, students would say that both pink and white

were dominant."

Suggesting changes in instruction: modifying the design of tasks. In the osmosis teaching episode, the suggested ideas for adapting instruction were promising in principle, but not specifically tied to the students' manifested difficulties. Monica suggested a modification of her planning approach to "incorporate more student experiences initially." Despite the fact that Monica noticed all students except Tom responded to her "most important question" inaccurately, and even Tom, who drew the water level accurately, did not show his understanding about the mechanism of osmosis, she did not make any suggestions to address this problem.

In the genetics teaching episodes, Monica noted that students "did not do a great job of connecting real world examples to these patterns." She specifically pointed out students' confusion between two distinct patterns (co-dominance vs. incomplete dominance), and described the actual changes that she made on the following day (adding complementary tasks).

Discussion

This study examined 14 PSTs' responsiveness to student thinking while engaged in assessment activities. Based on the findings, we offer two speculations about how teacher education programs might be able to better help PSTs attend and respond to student thinking. The first has to do with ensuring that PSTs have an opportunity to use more productive assessment tasks. The second has to do with when and where PSTs engage in interpreting student responses.

Using High-Quality Assessments Creates Opportunities for PSTs to Attend and Respond to Students' Scientific Thinking

Perhaps the most important finding from this study is that PSTs' productive responsiveness was *only* observed when they used productive assessment tasks—that is, tasks that appeared to provide opportunities for students to engage in scientific sense-making. In addition, the use of high-quality assessments was strongly related to PSTs' attention to students' cognitive behaviors and responsiveness to student thinking. We know that the design qualities of the assessments themselves significantly affected the quality of produced responses (Herman, 1992; Supovitz, 2012). Well-designed assessments make various ideas and student thinking visible (Furtak & Ruiz-Primo, 2008; Kang et al., 2014), which creates a condition for PSTs to notice students' ideas and ways of reasoning. In our data set, when productively framed assessments were used, student responses tended to be long and complex. In contrast, students mostly produced short and simple responses when unproductively framed assessment tasks were used. When unproductively framed assessments were used, PSTs mostly attended and responded to students' social or task behaviors.

While the program involved in this study sought to promote more productive assessments, PSTs used them in only 14 out of 32 teaching episodes (43.8%). We can only speculate on the reasons why high-quality assessments were not selected more frequently. It can be hypothesized that the variance has to do with disparities between program goals and the philosophies of mentor teachers—what is sometimes called the "two-worlds pitfall" (Anagnostopoulos, Smith, & Basmadjian, 2007; Feiman-Nemser & Buchmann, 1985). The nature of student teaching is such that PSTs encounter multiple frames of science knowledge and

learning as well as competing expectations between program and school, which might make it difficult for PSTs to select and use high-quality assessments provided by the program.

In our study, this hypothesis is partially supported by two cases in the *Conditional Group*, including Monica's. Both of these PSTs worked with mentor teachers who had strong ideas about the goals of science learning that contradicted those held by the program. Initially both of the PSTs mostly relied on their mentor teachers' curricula, including assessment tasks. Later they designed their own assessment tasks, while actively using tools supplied by the program and negotiating them locally as they worked with mentor teachers (see the case of Monica). However, the fact that only two out of 14 cases provided support to this hypothesis, and more importantly some cases like David provided counter-evidence, seems to undermine this theory of the two-worlds problem. Instead, we conjecture that the selection of high-quality assessments has something to do with complex social interactions among PSTs, course instructors, and school professionals in the process of planning (see Martin, Snow, & Franklin Torrez, 2011).

It is also possible that the relatively low percentages of the high-quality assessments have to do with PSTs' perspectives on the goals of science learning and knowledge, grounded in their past experiences—PSTs' initial, and apparently un-changed epistemic frames during the preparation period. Three of the cases in the *Responding to other concerns* group support this hypothesis. These three PSTs commonly worked with exemplary mentor teachers who were highly regarded by the program receiving relatively consistent and coherent messages from both program and school. Despite the available assessment tasks from the program and the feedback provided by course instructors, in the end, high-quality assessments were not selected by these PSTs and the suggested ideas for modifying the design of assessments were not taken seriously across all the observed teaching episodes.

What made the PSTs less interested in changing their assessments? One hypothesis is that it might be because these PSTs thought of their assessments as good and productive. Like most teachers, PSTs selected items that they believed would inform them about students' success in achieving their instructional goals. Thus PSTs who came with unproductive epistemic frames selected low-cognitive-level assessment items that were consistent with their frames. For example, David, who consistently used unproductively framed assessments, expressed satisfaction with his instruction and students' responses, which was contradictory to mentor and course instructor's strong concerns. By choosing unproductively framed assessment tasks to begin with, David created qualitatively different conditions to learn about student thinking.

In contrast, some PSTs like Leslie—who came in with a strong interest in students' ideas and thinking and framed teaching as working on students' ideas to begin with—actively interacted with course instructors and chose the high-quality assessment tasks that revealed rich information about students' ideas. Those PSTs created ample opportunities for themselves to attend and respond to student thinking through their active choice of assessment tasks. The assessment activity was designed to help PSTs learn about student thinking and how to respond to it. However, the analysis showed that *PSTs actively created different qualities of opportunities for their own learning with their choices of assessment tasks*. And over time, the PSTs like Leslie learned more about students' ideas and developed their capacity for responsive teaching.

This pattern of "the rich get richer, the poor get poorer" has also been suggested by other researchers. In a study about beginning teachers' collective inquiry about student work, for example, Windschitl and colleagues (2011) found that beginning teachers who had complex and problematized views on teaching and learning, such as Leslie, derived the most benefit from the collegial inquiry into student work. In contrast, the teachers with simplistic and unproblematized

views, such as David, did not show significant learning gains through the analysis of student

work.

Lastly, the reason why high-quality assessment tasks were not used by PSTs might have

to do with the intertwined nature of instruction, assessment, and standards. Teachers cannot

assess something that they do not teach. In general, teachers do not teach the things that are not

listed in the standards. Even though some PSTs aspired to use productively framed assessment

tasks provided by the program, they might not be able to select those assessments if their

instruction (or their mentor's) did not match the assessments. In her teaching report, Monica

described her selection of assessment tasks as "at this point I am still using my mentor teacher's

materials." Monica might not have been able to select assessments during the period when she

primarily followed Ms. S's instructional approaches. David might not have seen a reason to

change his assessment in part because he thought his assessment aligned with the standards.

In short, our analysis clearly suggests that the use of high-quality assessment tasks that

visualize students' thinking are necessary to promote PSTs' responsiveness to student thinking.

Despite this obvious condition, creating productive situations for PSTs in student teaching

contexts is not a simple task for teacher educators. The analysis reveals the contentious nature of

selecting and using assessment tasks at the intersection of multiple expectations from PSTs,

school professionals, and university programs.

Helping PSTs to Re-frame Problems by De-privatizing Interpretation

The use of high-quality assessments was necessary, but not sufficient, to promote PSTs'

abilities to attend and respond to students' scientific thinking. One third of the teaching episodes

that began with high-quality assessment tasks failed to yield evidence of PSTs' responsiveness to

student thinking (n=6 out of 18, 33.3%). Analysis of these six teaching episodes showed PSTs'

difficulties in noticing and reasoning about student thinking even when they were provided rich information about students' ideas. Other researchers have documented similar findings (see Gearhart et al., 2006; Lyon, 2013; Maclellan, 2004; Morgan & Watson, 2002). They pointed out that novice teachers' underdeveloped knowledge about content, teaching and learning, assessment, facility, or skills about assessment can explain these difficulties. Building on this previous work, in this study we attended to PSTs' framing and ability to notice in accounting for the five PSTs' uneven success in responding to students' thinking under the condition of using high-quality assessment tasks.

We conjecture that some PSTs' failure to attend to students' cognitive behavior, even with use of high-quality assessment tasks, had to do with novice teachers' *frame shifting* abilities. For instance Shannon, who was partnered with Leslie and taught the same microscope lesson using identical teaching materials and assessment tasks to a different period in Mrs. F's classroom, only noticed students' task behaviors. Both Leslie and Shannon described their difficulties in managing the classroom in their reports as well as during the interviews. Leslie noticed and responded to various student ideas along with students' task behaviors. But Shannon only highlighted students' task behavior while analyzing student work (see Table 4). Classrooms are complex spaces, especially to novice teachers, where teachers must be able to simultaneously monitor multiple problems within a crowd of students (Kennedy, Under review). Expert science teachers who carry out effective instruction are capable of attending and responding to various issues moment by moment while *flexibly shifting their framing*—their sense of "what is going on here." We speculate that novice teachers' inflexibility in shifting their framing has significant influence on their responsiveness.

Attending and responding to students' cognitive, social, and task behaviors are all legitimate and essential skills to successfully helping students to learn. In order to support novice teachers as they gradually but systematically develop this expert-like ability, teacher educators can reduce the complexity of the work of teaching by having PSTs focus on one aspect at a time (Lampert et al., 2013). Within the structure of this assessment activity, PSTs were expected to focus on students' ideas and thinking and learn how to respond to it. However, for some PSTs like Shannon and Mary, who seemed to be overwhelmed with management-related concerns in the early stages of their student teaching, engaging PSTs in this assessment activity did not seem to be sufficient to draw their attention to students' cognitive behavior and notice students' ideas and thinking. Alternatively, Mary's and Shannon's difficulties in attending to students' cognitive behaviors through the use of high-quality assessments might have been related to their underdeveloped knowledge and skills regarding content, teaching and learning, and assessment—the essential components for noticing (Sherin & van Es, 2005; van Es & Sherin, 2008). Their weak content knowledge and skills might have led them to attend to students' task behaviors, instead of cognitive behaviors. However, Mary succeeded in attending and responding to student thinking in her next teaching episode a month later. Assuming that Mary's knowledge and skills would not have grown to such a degree within a single month, this knowledge theory appears to be less convincing in accounting for novice teachers' failure of attending and responding to students' scientific thinking than our hypothesis of flexible framing shift when attending and interpreting student responses.

Some PSTs' failure to respond to students' thinking under the condition of using high-quality assessment tasks points to the problem of the privatized nature of the interpretation in the design of this learning opportunity. One important pattern that emerged from the six teaching

episodes was the way in which PSTs framed the problem while interpreting student responses, and how it led to proposed instructional modifications. In most cases, the "breakdown" on the pathway from attending and responding to student thinking was observed when PSTs reasoned with student responses during interpretation. All five PSTs linked students' failure to provide expected responses to some kind of problem *with* the students. The most popular link was to "lack of prior knowledge" without unpacking what that meant in the instructional context. Notably, in those cases, the suggested instructional modifications were generic or un-relevant, reflecting PSTs' limited repertoires and strategies.

In contrast, the PSTs who consistently attended and responded to student thinking, including Leslie, actively used the language from the program as her pedagogical reasoning resource to account for the observed student responses. PSTs' responsiveness hinged upon their interpretation of students' responses, in particular, the ways in which they framed the problem during the process, but there was little evidence that this process was interrupted during their engagement with this assessment activity. Within the structure of the activity and despite the support given by the program, such as one session of collective inquiry into student work, the use of the alternative reasoning resources was largely dependent upon PSTs' personal choices.

Taken together, setting up opportunities for PSTs to learn about students' ideas and thinking and how to respond to it requires more than having PSTs use high-quality assessments. Both PSTs' frame shifting that results in noticing students' intellectual difficulties and their interpretation of student responses need to be carefully scaffolded; therefore, PSTs can re-frame the problem by considering alternative reasoning resources beyond their own ideas.

Conclusion

Over the last several decades, the field of education research on student learning has made significant progress. This leads teacher educators to have strong hypotheses about student learning and essential instructional capacities that are necessary to support students' learning. In contrast, the research on teacher education is still at the adolescent stage (Grossman & McDonald, 2008). The field is only beginning to understand how teachers develop essential instructional capacities and design professional learning opportunities informed by empirical evidence. This study contributes to our knowledge base about preservice teachers' learning about one important instructional capacity—attending and responding to student thinking.

The findings point to two considerations in designing learning opportunities to enhance PSTs' responsiveness to student thinking: (a) the use of high-quality assessment tasks that make student thinking visible, (b) helping PSTs to re-frame the problems by de-privatizing PSTs' interpretation about student responses. This study demonstrates that without using high-quality assessment tasks and without pressing PSTs to re-frame problems in relation to their instructional design, it is unlikely that PSTs will attend and respond to student thinking.

Even though key components of effective learning environments are apparent, designing learning opportunities that create conditions for PSTs' to be responsive to student thinking during the preparation period is a complex task. Teacher educators must strategically work with a *system* in a way to address various kinds of contextual and institutional challenges in designing learning activities for PSTs. The program in this study, for example, later changed the structures of the activities based on the lessons from this research project. In the modified assessment activities, each PST was required to use at least three assessments—one from their mentor teacher, one recommended by the program, and a third one that the PST prefers.

In the field of mathematics education, some researchers suggest using a carefully chosen set of '*instructional activities*' to support novice teachers in advancing their practices toward ambitious goals of mathematics learning (Lampert & Graziani, 2009). Instead of depending on novice teachers' designing and selecting assessment tasks, using pre-selected high-quality activities and assessment tasks as a common context for learning about teaching can be another way of addressing the problems. Recent professional development approaches that relocate the places of teacher learning into classrooms, such as Studio approaches or collective Lesson Study, can be promising models for de-privatizing the work of interpretation (see Lampert et al., 2013; McDonald, Kazemi, & Kavanagh, 2013; Teachers Development Group, 2010). This kind of design increases the space for teacher educators to step in and to provide alternative pedagogical reasoning resources in the moment; therefore increasing the likelihood for PSTs to better respond to students' thinking while re-framing the problem collectively and collaboratively.

We know that what teachers do in the classroom has a huge impact on students' learning. It is more influential for students' academic futures than any other in-school factor, including quality of curricula, type of school, length of school year, or peer influences (see Murnane & Steele, 2007; Rockoff, 2004; Sanders & Rivers, 1996). Preparing future science teachers who are capable of attending and responding to various students' needs is one imperative agenda. We can achieve this goal only when we, the teacher education community, cultivate responsive mechanisms while continuously revising our understanding of how, and under what conditions, PSTs productively learn from practices. This study sheds light on the complex processes of PSTs' learning in the context of a teacher preparation program, which provides a foundation for teacher educators' responsive practices.

References

Anagnostopoulos, D., Smith, E. R., & Basmadjian, K. G. (2007). Bridging the University–School Divide. *Journal of Teacher Education, 58*(2), 138-152. doi: 10.1177/0022487106297841

Atkin, J. M., Coffey, J. E., Morthy, S., Sato, M., & Thibeault, M. (2005). *Designing everyday assessment in the science classroom*. New York: Teacher College Press.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 9-21.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.

Calabrese Barton, A., Kang, H., Tan, E., O'Neill, T. B., Bautista-Guerra, J., & Brecklin, C. (2013). Crafting a Future in Science. *American Educational Research Journal, 50*(1), 37-75. doi: 10.3102/0002831212458142

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2012). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching, 48*(10), 1109-1136. doi: 10.1002/tea.20440

Denzin, N. (1978). *The research act: A theoretical introduction to sociological methods* (2nd ed.). New York: McGraw-Hill.

Denzin, N., & Lincoln, Y. (2005). Introduction: The discipline and practice of qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (Vol. 3, pp. 1-32). Thousand Oaks, CA: Sage.

Elby, A., Richards, J., Walkoe, J., Gupta, A., Russ, R. S., Luna, M. J., . . . Sherin, M. G. (2014). *Differing notions of responsive teaching across mathematics and science: Does the discipline matter?* Paper presented at the 2014 International Society of the Learning Sciences, Boulder, CO.

Feiman-Nemser, S., & Buchmann, M. (1985). Pitfalls of experience in teacher preparation. *Teachers College Record, 87*(1), 53-65.

Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching, 49*(9), 1181-1210. doi: 10.1002/tea.21054

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education, 92*(5), 799-824. doi: 10.1002/sce.20270

Gay, G. (2000). *Culturally responsive teaching: Theory, research, & practice*. New York, NY: Teachers College Press.

Gearhart, M., Nagashima, S., Pfotenhauer, J., Clark, S., Schwab, C., Vendlinski, T., . . . Bernhaum, D. J. (2006). Developing expertise with classroom assessment in K-12 science; Learning to interpret student work. *Educational Assessment, 11*(3 & 4), 237-263.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*: Aldine Transaction.

Goffman, E. (1974). *Framing analysis: An essay on the organization of experience*. New York: Harper & Row.

Grossman, P., & McDonald, M. (2008). Back to the future: directions for research in teaching and teacher education. *American Educational Research Journal, 45*(1), 184-205.

Hammer, D., Goldberg, F., & Fargason, S. (2012). Responsive teaching and the beginnings of energy in a third grade classroom. *Review of science, mathematics and ICT education, 6*(1), 51-72.

Herman, J. L. (1992). What research tells us about good assessment. *Educational Leadership, 49*(8), 74-78.

Interstate New Teacher Assessment and Support Consortium Science Standards Drafting Committee. (2002). *Modeling standards in science for beginning teacher licensing and development: A resource for state diaglogue*. Washington, DC: Council of Chief State School Officers,.

Kang, H., & Anderson, C. W. (2012). *The mechanisms of secondary science teacher candidates' learning to teach*. Paper presented at the NARST, Indianapolis, IN.

Kang, H., Thompson, J., & Windschitl, M. (2014). Creating opportunities for students to show what they know: the role of scaffolding in formative assessments. *Science education, 98*(4), 549-742.

Kennedy, M. M. (1999). The role of pre-service teacher education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 54-86). San Francisco: Jossey Bass.

Kennedy, M. M. (Under review). How does professional development improve teaching? *Review of Educational Research*.

Kloser, M. (2014). Identifying a core set of science teaching practices: A delphi expert panel approach. *Journal of Research in Science Teaching, 51*, 1185-1217. doi: 10.1002/tea.21171

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., . . . Crowe, K. (2013). Keeping It Complex: Using Rehearsals to Support Novice Teacher Learning of Ambitious Teaching. *Journal of Teacher Education*. doi: 10.1177/0022487112473837

Lampert, M., & Graziani, F. (2009). Instructional Activities as a Tool for Teachers' and Teacher Educators' Learning. *The elementary school journal, 109*(5), 491-509.

Levin, D. M., Hammer, D., & Coffey, J. E. (2009). Novice Teachers' Attention to Student Thinking. *Journal of Teacher Education, 60*(2), 142-154. doi: 10.1177/0022487108330245

Lortie, D. C. (1975). *Schoolteachers: A sociological study*. Chicago: University of Chicago Press.

Lyon, E. G. (2013). Learning to Assess Science in Linguistically Diverse Classrooms: Tracking Growth in Secondary Science Preservice Teachers' Assessment Expertise. *Science Education, 97*(3), 442-467. doi: 10.1002/sce.21059

Maclellan, E. (2004). Initial knowledge states about assessment: novice teachers' conceptualisations. *Teaching and Teacher Education, 20*(5), 523-535. doi: http://dx.doi.org/10.1016/j.tate.2004.04.008

Martin, S. D., Snow, J. L., & Franklin Torrez, C. A. (2011). Navigating the Terrain of Third Space: Tensions With/In Relationships in School-University Partnerships. *Journal of Teacher Education, 62*(3), 299-311. doi: 10.1177/0022487110396096

Maskiewicz, A. C., & Winters, V. A. (2012). Understanding the co-construction of inquiry practices: A case study of a responsive teaching environment. *Journal of Research in Science Teaching, 49*(4), 429-464. doi: 10.1002/tea.21007

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core Practices and Pedagogies of Teacher Education: A Call for a Common Language and Collective Activity. *Journal of Teacher Education, 64*(5), 378-386. doi: 10.1177/0022487113493807

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco: Jossey-Bass.

Metz, K. E. (1995). Reassessment of Developmental Constraints on Children's Science Instruction. *Review of Educational Research, 65*(2), 93-127. doi: 10.3102/00346543065002093

Metz, K. E. (2004). Children's Understanding of Scientific Inquiry: Their Conceptualization of Uncertainty in Investigations of Their Own Design. *Cognition and Instruction, 22*(2), 219-290. doi: 10.1207/s1532690xci2202_3

Minsky, M. (1985). *The society of mind*. New York: Simon and Schuster.

Morgan, C., & Watson, A. (2002). The Interpretative Nature of Teachers' Assessment of Students' Mathematics: Issues for Equity. *Journal for Research in Mathematics Education, 33*(2), 78-110.

Murnane, R. J., & Steele, J. L. (2007). What is the problem? The challenge of providing effective teachers for all children. *The Future of Children, 17*(1), 15-43.

Nasir, N. i. S., Rosebery, A. S., Warren, B., & Lee, C. (2006). Learning as a cultural process: Achieving equity through diversity. In K. R. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (pp. 489-504). New York: Cambridge University Press.

NGSS Lead States. (2013). Next Generation Science Standards: For States, By States

NRC. (2005). *How students learn science in the classroom*. Washington, DC: National Academies Press.

NRC. (2007). *Taking science to school: Learning and teaching science in grade K-8*. Washington, DC: National Academies Press.

NRC. (2010). *Preparing Teachers: Building Evidence for Sound Policy*. Washington, D.C.: The Naitonal Academies Press.

NRC. (2012). *A framework for K-12 science education: practices, crosscutting concepts, and core Ideas*. Washington, DC: National Academies Press.

Otero, V. K. (2006). Moving Beyond the "Get it or Don't" Conception of Formative Assessment. *Journal of Teacher Education, 57*(3), 247-255. doi: 10.1177/0022487105285963

Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *The American Economic Review, 94*(2), 247-252. doi: 10.2307/3592891

Roseberry, A., & Warren, B. (2008). *Teaching science as English language learners: Building on students' strengths*. Arlington, VA: NSTA Press.

Rosebery, A. S., & Puttick, G. M. (1998). Teacher professional development as situated sense-making: A case study in science education. *Science Education, 82*(649-677).

Ross, R. (1975). Ellipsis and the structure of expectation. *San Jose State Occational Papaers in Linguistics, 1*, 183-191.

Russ, R. S., & Luna, M. J. (2013). Inferring teacher epistemological framing from local patterns in teacher noticing. *Journal of Research in Science Teaching, 50*(3), 284-314. doi: 10.1002/tea.21063

Sadler, D. R. (1998). Formative assessment: Revising the territory. *Assessment in Education: Principles, Policy & Practice, 5*, 77-84.

Sanders, W. L., & Rivers, J. C. (1996). Cumulative and residual effects of teachers on future student academic achievement: University of Tennessee Value-Added Research and Assessment Center.

Sawyer, K. R. (2006). The new science of learning. In K. R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 1-16). New York, NY: Cambridge University Press.

Schank, R. (1990). *Tell me a story: A new look at real and artificial memory*. New York: Scribner.

Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (2011). *Mathematics teacher noticing: Seeing through teachers' eyes*: Taylor & Francis US.

Sherin, M. G., & van Es, E. A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education, 13*(3), 475-491.

Sohmer, R., Michaels, S., O'Connor, K. M., & Resnick, L. B. (2009). Guided construction of knowledge in the classroom: The troika of talk, tasks and tools.

Stake, R. (2004). Qualitative case studies. In N. Denzin & Y. Lincoln (Eds.), *The Sage Handbook of Qualitative Research, 3rd edition* (pp. 443-466). Thousand Oaks, CA: Sage.

Supovitz, J. (2012). Getting at student understanding--The key to teachers' use of test data. *Teachers College Record, 114*, 1-29.

Tannen, D. (1993). *Framing in discourse*. New York: Oxford University Press.

Teachers Development Group. (2010). Mathematics Studio Program: Transforming a school's culture of mathematics professional learning. In T. D. Group (Ed.): Teachers Development Group.

Thompson, J., Hagenah, S., Kang, H., Colley, C., Windschitl, M., Stroupe, D., & Braaten, M. (In press). Rigor and responsiveness in classroom activity. *Teachers College Record*.

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education, 24*(2), 244-276.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: Will it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). New York: Routledge.

Windschitl, M. (2005). Guest editorial: The future of science teacher preparation in America: Where is the evidence to inform program design and guide responsible policy decisions? *Science education, 89*(4), 525-534.

Windschitl, M., Thompson, J., & Braaten, M. (2011). Ambitious pedagogy by novice teachers: Who benefits from tool-supported   collaborative inquiry into practice and why? *Teachers College Record, 113*(7), 1311-1360.

Yin, R. K. (1989). *Case study research: Design and methods* (2nd ed.). Newbury Park, CA: Sage.