ABSTRACT

Title of Document:	LEARNING TO ELICIT, INTERPRET, AND RESPOND TO STUDENTS' HISTORICAL THINKING: A CASE STUDY OF FOUR TEACHER CANDIDATES.
	Michael A. Neel, Doctor of Philosophy, 2015
Directed By:	Dr. David Imig, Professor of Practice, College of Education, Department of Teaching and Learning, Policy and Leadership; Dr. Linda Valli, Professor, College of Education, Department of Teaching and Learning, Policy and Leadership

Teacher education researchers have argued that teacher candidates must learn to attend to students' disciplinary thinking if they are to improve student learning. In history education, such attention must focus on student thinking about evidence because interpretation of evidence is at the heart of historical discourse. This study explores how four teacher candidates who had learned to attend to students' historical thinking in a social studies methods course engaged in the practice of eliciting, interpreting, and responding to that thinking during their internships.

Data collected over a nine-month period included observations of candidates in their methods courses, a pretest administered before the methods

course, observation of at least four lessons per candidate in the internship, interviews with teachers after each observed lesson, and analysis of methods coursework. Case study analyses indicated that two of the candidates elicited, interpreted and responded to students' historical thinking while another did not, and a fourth did so only under certain conditions. The cross-case analysis showed that although all of the candidates used methods course tools in the internship, some were unable to use these tools to elicit students' historical thinking.

While three of the four candidates noticed historical thinking and considered that thinking in determining an instructional response, what candidates noticed was limited to the scope of their instructional objectives. Only one candidate consistently responded to student thinking in evaluative ways, and all four struggled to deliver responses that maintained a focus on student reasoning. Instead, candidates preferred to demonstrate their own reasoning, either by building on a student idea or simply as a means to make a point not directly related to a student idea.

This study highlights the interconnected nature of eliciting, interpreting, and responding to student thinking and offers insight into how teacher educators can facilitate attention to student historical thinking. It also points to factors that are important for the development of this ability including candidate disciplinary knowledge and the social contexts of learning.

Furthermore, this study provides a framework and analytical tools that can enable future researchers to examine this phenomenon more deeply.

LEARNING TO ELICIT, INTERPRET, AND RESPOND TO STUDENTS'
HISTORICAL THINKING: A CASE STUDY OF FOUR TEACHER
CANDIDATES

By

Michael A. Neel

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
David Imig, Professor of the Practice (Co-chair)
Linda Valli, Professor (Co-chair)
Daniel Chazan, Professor
Susan De La Paz, Associate Professor
Daniel Levin, Clinical Assistant Professor
John O'Flahaven, Associate Professor

For Greg Morrow, John Wells, and Ross Brummett,
who believed in me before I ever believed in myself.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

For decades, the preparation of teachers has been a matter fraught with political

controversies, pedagogical uncertainties, and research challenges (Conant, 1963; Dewey, 1904;

Fraser, 2007; Goodlad, 1990; Labaree, 2004; Sarason, 1993; Smith, 1980). Despite ever

increasing attention to teacher preparation, answers to a host of questions, particularly those

about how teacher candidates (TC) learn to teach, how that process can be facilitated, and even

whether preparation results in a product worthy of investment, remain shrouded in uncertainty.

The uncertainty, in part, accounts for the proliferation of models and approaches to teacher

preparation, some of which represent competing hypotheses about the development of teacher

learning.

What little consensus exists about teacher preparation has formed around the idea that it

should be 'practice-based,' a term with little operational continuity across programs (Lampert,

2009; Zeichner, 2012). One such practice-based approach to teacher preparation aims to

improve teacher candidates' readiness to teach by reframing TC learning around the enactment

of *core practices* of teaching (McDonald, Kazemi, & Kavanaugh, 2013). This reframing (largely

of methods courses and not programs at this point) is marked by teacher education pedagogies

built around a set of replicable teaching practices in which teacher knowledge, skill, and identity

are developed through a process of learning teaching while learning to enact specific teaching

practices (Grossman & McDonald, 2008).

Despite enthusiasm for *core practice* approaches in teacher education and theoretical

evidence that the idea is promising, few empirical studies examine TC learning in the context of

a core practice approach (c.f., Thompson, Windschitl, & Braaten, 2013). This dissertation study

is broadly intended to address aspects of this gap by examining teacher candidates' developing capabilities to engage in a core practice targeted by a methods course.

In particular, this study is focused on history TCs' ability to elicit, interpret, and respond to student thinking, a practice found to improve student learning across multiple subject areas (Black & Wiliam, 1998; Shepard, 2000). Of course, some student thinking is more important for instructional purposes than other student thinking. In this study, I privilege teacher candidate attention to student thinking about evidence in history because evidence forms the substance of historical arguments and is fundamental to understanding anything about the past (Sexias, 1996; VanSledright, 2004). Furthermore, student work with historical evidence can improve student reading, writing, and reasoning skills (De La Paz et al., 2014; Monte-Sano, 2011b; Reisman, 2012a) in line with requirements of the *Common Core State Standards* and *C3 Framework for Social Studies State Standards*.

Considering the elements above, the purpose of this study is to examine teacher candidates' abilities to elicit, interpret, and respond to secondary students' thinking about historical evidence in light of a methods course designed to develop these abilities. This initial chapter will introduce the fundamental aspects of the study. The chapter is organized topically as follows: 1) Background of the study; 2) Problem and purpose of the study; 3) Significance of the study; 4) Theoretical framework; 5) Research Questions; and 6) Organization of the Study.

Background of the Study

When I started teaching school over a decade ago, I was entirely unprepared for the job before me. As I gained experience as a history teacher, I recognized that my students' interest was due more to my enthusiasm than to any ambitious intellectual engagement. Only after

several more years of teaching did I begin to understand specific problems in my practice that limited students' opportunities to learn.

At the same time I was recognizing the disturbing reality that my practice was insufficient in ways I felt unprepared to address, I became interested in my spouse's preparation as a physician. My initial interest was driven by my perception of medical preparation as a rigorous, practice-oriented, and formative model as compared to my professional preparation. But, when I made teacher education the topic of my doctoral studies, I discovered complexities that challenged my initial naiveté.

Some of the problems I identified in my preparation were actually century-old puzzles with no obvious resolutions. For example, Dewey (1904) asked whether learning to teach required learning about theoretical concepts prior to engaging in teaching or whether concepts were learned as one engaged in teaching. Through further inquiry into such puzzlements, I became focused on questions involving how a person learns to teach and what, if anything, can be done to facilitate that learning.

As I explored the literature on learning to teach and I assumed responsibility for courses on teaching, I found a line of research that specifically addressed many of the questions and problems I faced as a teacher and teacher educator (e.g., Ball & Forzani, 2009; Grossman & McDonald, 2009; Grossman et al., 2009; Lampert, 2009). These authors framed *practice* as the centerpiece of professional learning and *enactment* as a central pedagogy of teacher preparation. I became increasingly interested in using core practices as a lever for teacher learning in my courses. My early experimentation with core practices in methods courses drove me back to the literature in an attempt to better understand the various approaches and eventually led me to the curiosities that form the foundations of this study.

Problem and Purpose

A flurry of articles and book chapters published in recent years outline specific *core practices* of teaching and describe potential approaches to reframing methods courses according to various teaching practices (Ball & Forzani, 2009; Forzani, 2014; Grossman, Hammerness, & McDonald, 2009; McDonald et al., 2014). The emerging literature and available presentations on core practices fall primarily into one of two categories.

First, a growing body of literature focuses on deconstructing teaching into specific practices (e.g., Ball & Forzani, 2009; Grossman & McDonald, 2009; Kazemi, Lampert, & Franke, 2009). These hypothesized core practices or "high-leverage practices" (Grossman, Hammerness, & McDonald, 2009) are specific, routine teacher activities that demand professional judgment, are devoted to planning, enactment, or reflection, and are intended to support student learning (McDonald et al., 2013; Windschitl, Thompson, Braaten, & Stroupe, 2012). Examples of core practices include leading whole-class discussions of content, choosing and using examples of content, and setting up and managing small group work (teachingworks.org).

The second, and directly related, category of literature explores how core practices can be used in teacher education courses (e.g., Ghousseini & Sleep, 2011; Lampert, Beasley, Ghousseini, Kazemi & Franke, 2010; Lampert et al., 2013). Authors of this body of literature explain specific pedagogical strategies underway in courses that are intended to cultivate TC learning around particular core practices.

These pedagogical approaches build on a model of professional preparation initially described by Grossman et al. (2009) as an iterative process of representing a practice, decomposing a practice, and approximating a practice for enactment. Descriptions of these

"pedagogies of enactment" (Grossman et al., 2009) are an important extension of the first body of literature (i.e., theoretical definitions of core practices) because these writings provide concrete examples of core practices as used in actual methods courses.

At present, therefore, the literature on core practices in teacher education is either primarily theoretical or descriptive. I identify exceptions to this trend in Chapter 2 (e.g., Janssen, Westbroek, & Doyle, 2014; Lampert et al., 2013; Thompson, et al., 2013). As teacher educators move ahead with core practices of teaching in teacher preparation courses, it is vital that research extend theoretical analysis and description into empirical investigations of TC learning in the context of core practice approaches.

The present study provides just such an analysis of TC ability in a target core practice of history teaching. In this study, I follow four teacher candidates who participated in a methods course designed in accordance with the literature on core practices (McDonald et al., 2013) and intended to facilitate the practice of eliciting, interpreting, and responding to student thinking. The selection of this practice, rather than another relevant teaching practice, is explained in detail in Chapter 3.

Conceptual Framework

The conceptual framework for this study helped to direct the focus of the study on candidate ability in a target practice (i.e., eliciting, interpreting, and responding to student thinking about evidence in history). Although researchers have begun to conceptualize mathematics and science teachers' ability to engage in this practice (e.g., Levin, Hammer, & Coffey, 2009; Singer-Gabella et al., 2015; Thompson et al., 2013), scholars know comparatively little about history teacher candidates' capabilities.

These scholars refer to notions such as "attending to student thinking" (Levin et al., 2009, p. 1) and "leveraging student thinking" (Singer-Gabella et al., 2015) to describe what I call eliciting, interpreting, and responding. Like other scholars, I conceive of this attention to student thinking as a single teaching practice made up of distinct but entangled sub-practices. For the sake of analysis, I find it helpful to distinguish the parts from the whole: the conceptual framework for this study distinguishes between eliciting, interpreting, and responding, despite the interdependent nature of these activities.

Eliciting student thinking in history begins with the establishment of a historical task or problem space in which student reasoning can develop. Productive eliciting is marked by generativity, or the capacity to promote shared investigation of students' understanding of the content (Lampert, 2001; Singer-Gabella et al., 2015; Sleep & Boerst, 2012; Stein, Engle, Smith, & Hughes, 2008). For example, consider a classroom in which students are engaged in the analysis of five primary source documents for the purpose of answering the question, "Was Lincoln a racist?" Compare that scenario to a classroom in which students are answering comprehension questions about a textbook reading on the Lincoln-Douglas debates. Eliciting, therefore, begins with a generative problem space and continues with questions and prompts that initiate new student thinking in accordance with that problem space (Singer-Gabella et al., 2015).

Interpreting consists of a teacher candidate's understanding of student statements elicited during the instructional task. Such understanding requires that a candidate notice student thinking and assess student understanding in the target objectives, a process influenced by a host of invisible factors (Kennedy, 2005; Lampert, 1985). For the purposes of this study, interpretation is limited to those understandings articulated by the teacher candidate during interviews (Sleep & Boerst, 2012).
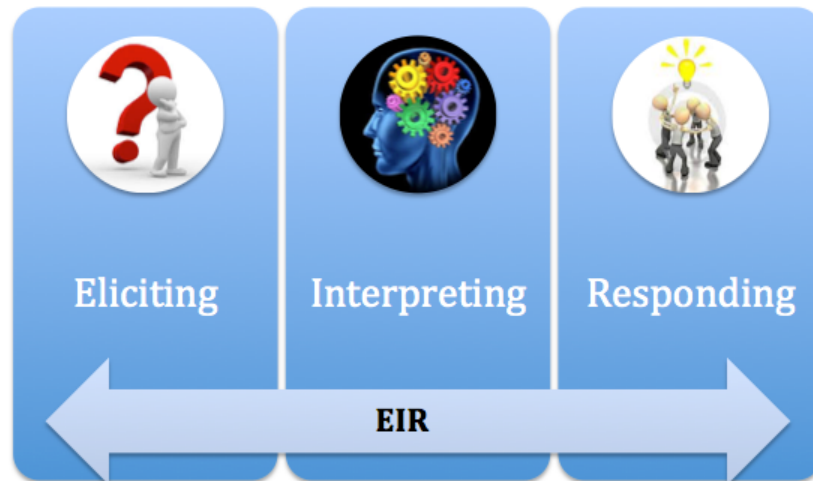
Figure 1.1: Conceptual framework

Responding goes hand in hand with eliciting and interpreting. Responding is the instructional move a teacher makes in response to student thinking about historical evidence, when that thinking is voiced in the classroom (Levin et al., 2009; Pierson, 2008). Responsiveness, then, is the extent to which teachers 'take up' student thinking in their instructional interactions. This eliciting, interpreting, and responding (EIR) framework supported a delineation of the phenomenon of interest in this study and directed the design, data collection procedures, and data analysis for the study.

Research Questions

I took seriously Merriam's (1998) suggestion to start the design of research questions with one's greatest curiosities. So I began simply with the open-ended question, "What do I want to know?" (Merriam, 1998, p. 57). I returned to my UMD application statement of purpose and reviewed the major papers I had written during my coursework. From the beginning, I have had an interest in how a person develops expertise in teaching and what can be done to facilitate that process.

Armed with a basic sense of what I wanted to know, I borrowed from Yin's (2006) encouragement to test all ideas against literature and colleagues. Over time, I worked with colleagues to narrow general interests and broad questions into "a focused list of essential and necessary questions" (Patton, 1990, p. 163). Using the conceptual framework above to guide the inquiry, I designed the following research question and two sub-questions:

How do the teacher candidates elicit, interpret, and respond to student thinking about historical evidence during their teacher preparation program?

1. How do they engage in these practices at the outset of their program?

2. In what ways do their capabilities to engage in these practices change during their program of study?

I designed this inquiry to explore the research questions with a qualitative embedded case study methodology (Yin, 2003). Extensive details about case selection, data collection, and data analysis are provided in Chapter 3.

Significance

The findings of this study address gaps at the intersection of two lines of research. First, the study serves as one of the first empirical inquiries to examine TC classroom ability in a core practice targeted by a methods course. In addition to the relevance of research on core practices in teacher education, it also provides insight into history teacher candidates' ability to engage in eliciting, interpreting, and responding to student thinking, an area of inquiry heretofore limited primarily to mathematics and science education.

Definition of Terms

In order to avoid confusion, I use the following definitions and distinguish between a number of important terms.

Core Practice

Specific, routine aspects of teaching that demand the exercise of professional judgment and the creation of meaningful intellectual and social community for teachers, teacher educators, and students. Core practices are those that occur with high frequency in teaching; that novices can begin to master and enact in classrooms across different curricula or instructional approaches; that allow novices to learn more about students and about teaching; and that are research-based, preserve the integrity of teaching, and have potential to improve student achievement (McDonald et al., 2013).

Teaching Practices

Routine activities teachers engage in devoted to planning, enactment, or reflection that are intended to support student learning (Windschitl et al., 2012, p. 882)

Teacher Candidate

A person enrolled in a formal course of study that leads to licensure and/or certification in teaching; while this person may participate in internship experiences for educational purposes, he/she not a teacher of record.

Organization of the Study

I present this dissertation study in nine chapters. In this first chapter, I have provided a brief overview of the background of the study, statement of the problem, purpose of the study, theoretical framework, research questions, and significance of the study. The next chapter will

demonstrate that established literature points to the need for research that explores teacher candidates' abilities to elicit, interpret, and respond to student thinking about historical evidence in light of a methods course designed to develop these abilities.

With relevant gaps in the literature made plain in Chapter 2, the third chapter indicates how the methodology, research design, selection of participants, data collection, and data analysis procedures align with the study's research questions. Chapter 4 consists of a brief description of the core practice approach taken in the teacher education course. Chapters 5-8 present the findings for each of the four cases, and Chapter 9 is a cross-case analysis of these cases. Chapter 9 concludes with a discussion of the findings' relevance to the literature and a consideration of the study's implications.

Chapter 2: Literature Review

Introduction

In this chapter, I argue that the field of teacher preparation needs research that examines teacher candidates' abilities to elicit, interpret, and respond to students' thinking about historical evidence in light of a core practice approach. First, I examine core practices of teaching, how core practices are used in teacher education courses, and what we know about how teacher candidates (TCs) learn in these contexts. I argue that although core practice approaches appear promising, the field does not know enough about how candidates develop target practices in the context of such approaches.

With a core practice approach made clear, I then examine research on the core practice selected for this study: eliciting, interpreting, and responding to student thinking (EIR). In this second section of the literature review, I argue that while some research has examined mathematics and science TCs' ability to EIR, we know little about similar capabilities of history candidates.

In the final section of this literature review, I consider student thinking about evidence in history, the student thinking prioritized in this study. I argue that although we know that student attention to evidence in history can lead to improved student learning**,** we do not know how teacher candidates learn to EIR to such thinking. Together, the intersection of these gaps suggests the need for research that examines TCs' abilities to EIR to student thinking about historical evidence, in light of a core practice approach.

Teacher Education and Core Practices

In this section, I provide an overview of core practices in teacher education, explain how they are used in some teacher education courses, and clarify how core practice approaches are

distinct from similar approaches in teacher education. I then discuss what research suggests about how candidates develop core practices and I argue that the field needs more research to explore how candidates understand target practices in the context of a core practice approach.

What is a Core Practice Approach? In recent years, scholars of teacher education are increasingly calling for approaches that are 'practice-based' (Ball & Cohen, 1999; Matsko & Hammerness, 2014; Solomon, 2009). One strand of practice-based approaches intentionally designs teacher education around particular practices of teaching. Articulation, exploration, and enactment of these practices then form the foundation of the program of study (Forzani, 2014). Such approaches intended to make the enactment of teaching practices the focal point of learning to teach, rather than bifurcating learning and later application of that learning, as professional preparation has been prone to do (McDonald et al., 2013).

Advocates of the approach aim to develop teaching practices that are "essential for novices" (Forzani, 2014, p. 357) or practices that novices need to develop before being allowed to teach. Of course the work of teaching requires dozens, if not hundreds of practices on a daily basis (Forzani, 2014). These essential, or core, practices are "specific, routine aspects of teaching that demand the exercise of professional judgment and the creation of meaningful intellectual and social community for teachers, teacher educators, and students" (McDonald et al., 2013, p. 1). More specifically, scholars suggest that core practices are teaching practices that:

- Occur with high frequency in teaching,
- TCs can begin to master,
- TCs can enact in classrooms across different curricula or instructional approaches,
- Are research-based and have the potential to improve student achievement,
- Allow TCs to learn more about students and about teaching,

- And preserve the integrity and complexity of teaching (Grossman et al., 2009). While there is yet no formal process for determining what counts as a core practice, the University of Michigan has developed a list of examples (they call 'high-leverage practices') that include such teaching practices as posing questions about disciplinary content, leading whole class discussions about disciplinary content, setting up small group work, and recognizing and identifying typical patterns of student thinking in specific content areas (TeachingWorks, 2013).

Core practice approaches are significantly influenced by research that examined the preparation of several professions that require educative relationships including counselors, teachers, and clergy (Grossman et al., 2009). The findings have pointed to three foundational concepts helpful in describing the mechanics of professional preparation: representations of practice, decompositions of practice, and approximations of practice.

Representations of practice include the ways instructors depict professional practice to candidates and the facets of practice highlighted through such representations. Instructors use components such as video records of practice, case narratives, and field observation to target particular representations of practice in order to create a "professional vision," (Grossman et al., 2009, p. 2069) for the practice. Representation, however, is only a jumping-off point for learning the practice.

The instructor then focuses the trainees' attention on decomposed components of that representation. Decompositions of practice are intended to divide representations into discrete parts in order to make particular aspects of practice both teachable and learnable, a step potentially helpful for novices who struggle to notice complex aspects key to successful practice (Ericsson, 2006). Identification of particular component parts of successful practice in professional preparation, however, is not an end in itself. Upon identification of representations

of practice and decomposition of that representation into component parts, instructors then design "approximations of practice" (Grossman et al., 2009, p. 2058) in order for novices to attempt aspects of practice in settings of reduced complexity.

The authors liken these approximations to kayak lessons on calm waters, in the sense that early attempts, "fraught with awkwardness and uncertainty," (p. 2091) can occur in a safe setting rather than at the risk of public failure. While approximations of practice were never "the real thing," such approximations may best be judged somewhere on a continuum from least complete and authentic to most complete and authentic (Grossman et al., 2009). These research findings establish a language for "pedagogies of enactment" that form the basis for how CP can be used in teacher education programs.

Core practice approaches in the literature are most developed in mathematics and science education (Kazemi, Lampert, & Ghousseini, 2007; Kazemi et al., 2009). Teacher educators and researchers (initially at The University of Michigan) worked together to build on Ball and Cohen's (1999) notion of teacher learning rooted in practice and Cochran-Smith and Lytle's (1999) construct of learning to teach in, from and for teaching practice (Ball, Sleep, Boerst, & Bass, 2009). Over the course of several years, the group re-designed preservice methods courses in order to better equip TCs for the work of ambitious teaching, or "teaching that deliberately aims to get all kinds of students— across ethnic, racial, class, and gender categories—not only to acquire, but also to understand and use knowledge, and to use it to solve authentic problems" (Lampert & Graziani, 2009, p. 491).

Citing an Italian language school as a practical model of their developing conceptual frame (Lampert & Graziani, 2009), the consortium articulated a cycle of modeling, planning, rehearsing, enacting and reflecting on an instructional activity. Using this frame, the consortium

redesigned math methods courses around instructional activities that could be used to "practice" certain practices of teaching mathematics. The approach was intended to make teaching practice "studyable" (Ghouseini & Sleep, 2011) in the sense that practices be modeled, analyzed, enacted, and reflected upon (Lampert et al., 2010).

Others have advocated slightly different, though not irreconcilable, notions of core practice approaches. For example, science teacher educators at Washington have been pursuing a similar course for years and arrived at four CPs that the program uses to organize TC learning including: 1) framing a big idea; 2) eliciting student ideas; 3) helping students make sense of material activity; and 4) pressing students for evidence-based explanations (Windschitl et al., 2012).

Admittedly, versions of core practice approaches differ in particular details such as practice grain size and specificity to content. For example, is a core practice a general teaching practice in which every teacher of every subject engages (e.g., Grossman & McDonald, 2009)? Or, is a core practice more narrowly bound to a particular subject area (e.g., Windschitl et al., 2012)? The common denominator across versions I have explored is that core practice approaches are built on the assertion that 1) much of teaching can be decomposed into specific practices without reducing the complexity of teaching practice, and 2) TCs can learn these practices through a cycle of learning that includes study of a model, rehearsal, enactment and reflection on that enactment.

Core practices in teacher education courses? Teacher educators have since advanced Grossman et al.'s (2009) work and now provide an increasingly coherent idea of a core practice approach for methods courses (e.g., Ghousseini & Sleep, 2011; Lampert, Beasley, Ghousseini, Kazemi & Franke, 2010; McDonald, et al., 2013; McDonald, Kazemi, Kelley-Petersen,

Mikolasy, Thompson, Valencia, & Windschitl, 2014). Descriptions of these "pedagogies of enactment" (Grossman et al., 2009) are an important extension of theoretical literature that conceptualizes core practices of teaching because these descriptions provide specific examples of these pedagogies from methods courses. For example, teacher educators are reporting practices of modeling (McDonald et al., 2013), rehearsals (Lampert et al., 2013) and the use of artifacts for representation of practice (Forzani, 2014; Ghousseini & Sleep, 2011). This body of literature is invaluable for teacher educators who design and implement a methods course based on a core practice approach.

Although variance exists, core practice approaches generally resemble a model for teacher learning articulated by McDonald et al. (2013). As depicted in Figure 2.1, this model highlights learning as an iterative process of representing a practice, decomposing a practice, approximating a practice for enactment, and reflecting on the enactment.

Figure 2.1: Core Practice Cycle (McDonald et al., 2013)

McDonald et al. (2013) approach the CP by way of an "instructional activity," visualized as a "container" for targeted core practice (Lampert et al., 2013). For example, if the CP under study is leading whole class discussions, the teacher educator designs an instructional activity that "contains" the practice. The instructional activity then can be modeled, analyzed, rehearsed, enacted and reflected upon as a means to engage the core practice.

Kazemi, Lampert, and Ghousseini (2007) provide an example of an elementary mathematics instructional activity called *choral counting* — a mathematics strategy that helps students build specific skills and patterns. Choral counting is one of several instructional activities intended to support elementary TCs in developing skills and knowledge that go well beyond the boundaries of the choral counting activity (Lampert et al., 2010).

In the choral counting instructional activity, TCs participate in a cycle of learning that includes 1) observation and deconstruction of a video model of choral counting, 2) design and

rehearsal of choral counting, 3) enactment of choral counting in an actual classroom, and 4) reflection on the enactment of choral counting. This cycle, which follows the McDonald et al. (2013) framework in Figure 2.1, engages in each of the phases of learning to practice that Grossman et al. (2009) conceptualized (c.f. McDonald et al., 2014; tedd.org).

As a TC engages in iterative cycles of instructional activities that build capacity in a core practice, candidates are expected to develop a "mental schema for an instructional 'chunk' that can routinely be utilized" (Lampert et al., 2010, p. 137) by adapting it to specific contexts. That is, TCs learn to teach through a reflective process of routine implementation and selective adaption of those routines, given the circumstances and contexts of authentic classroom teaching and learning.

How is a Core Practice Approach Distinct? Some advocates of core practice approaches openly point to similarities with earlier, now mostly abandoned, notions of practice-based teacher education (Forzani, 2014; Grossman & McDonald, 2009; Zeichner, 2012). After all, some teacher training efforts as early as the 1840s consisted of an approach grounded in opportunities to practice skills of instruction (Forzani, 2014; Fraser, 2007; Ogren, 2005). As for the identification of teaching practices that could inform a redesign of teacher education, The Commonwealth Teacher Training Study of 1929 surveyed thousands of teachers in an attempt to catalog "every move a teacher might make" (Forzani, 2014, p. 363).

Although the Commonwealth study's findings did not result in significant application to teacher education, a later iteration of deconstructing teaching practice was reborn in a widely impactful movement known as competency-based teacher education. This approach intended to use student performance and teacher behavior to identify the practices of effective teachers and, in turn, provide a rubric to guide teacher education (Ryals, 1972). Competency-based approaches

18

to teacher education tended to view teaching as a collection of atomized practices that could be identified, replicated, and performed at varying levels of competency (for a review of this work see Valli & Rennert-Ariev, 2002). Training, therefore, focused on assessable competencies and programs of study were organized into modules that developed these competencies (Forzani, 2014).

Most research on microteaching practices in teacher education in the 1970s and 80s suggested that teachers trained in this way struggled to reassemble atomized practices in classroom teaching contexts (Cruickshank & Metcalf, 1990; Grossman, 2005). That is, many TCs who became proficient in particular moves or practices in decontextualized settings did not use the competency when faced with actual students. Many of the same ideas expressed in competency-based teacher education have reemerged in recent years via efforts to develop generic teaching techniques that "put all students on the path to college" (Lemov, 2010). The apparently long tradition of practice-based teacher education should lead the field to ask whether core practice approaches are any different.

Forzani (2014) has argued that one way to understand core practice approaches is to see the reform as building on but departing "from earlier efforts to prepare teachers for practice" (Forzani, 2014, p. 359). That is, rather than a different approach, core practice approaches might better be understood as a return to some of teacher education's discarded past (Grossman, 2005). Rather than novel, core practice approaches are better understood as "a part of a long trajectory of attempts to build teacher education into powerful preparation for practice" (p. 358).

While core practice approaches are similar to past efforts, Forzani (2014) points to three distinguishing characteristics of efforts to design teacher education around core practices. First, teaching is conceived of as an interactional exchange in which student thinking is a primary

resource. As compared to competency-based predecessors, core practice approaches focus on eliciting the knowledge of students and building on those understandings.

Second, teaching practices in a core practice approach are intended to accomplish situated instructional goals. Instead of a generic practice with specific moves that could be implemented across contents, core practices are more commonly approached as "a vehicle for accomplishing particular instructional goals, in relation to specific content" (Forzani, 2014, p. 365).

Third, in a core practice approach, teaching is viewed as inherently improvisational. Where competency-based approaches tended toward a mechanistic replication of expert practice, core practice approaches expect that TCs will experiment and engage in "inventing educative responses" (Lampert et al., 2010, p. 135).

Together, these three distinctions emphasize a view of teaching as an incredibly complex activity grounded in inquiry, interaction, and context. Learning to teach, as such, requires much more than mastering a set of skills. Core practice approaches suggest that learning to teach requires "repeated practice managing novel teaching situations" and the development of particular "skills such as listening, interpreting, and managing instructional discourse" (Forzani, 2014, p. 365). For a variety of reasons, these skills are not commonly found in American classrooms, although they are commonly taught in teacher education (Cuban, 1991; Forzani, 2014).

How TCs understand a core practice. A significant number of articles and book chapters published in recent years outline specific core practices of teaching and describe potential approaches to the implementation of these CPs in teacher preparation methods courses (Ball & Forzani, 2009; Grossman et al., 2009; Kazemi et al., 2014; Lampert, 2009; Lampert et

al., 2013; McDonald et al, 2013; Windschitl et al., 2012). Evidence that would tell the field how TCs understand these practices, however, remains largely anecdotal.

For example, Ball, Sleep, Boerst, and Bass (2009) described a process then underway at the University of Michigan to improve preservice teaching of mathematics through an approach to methods courses that utilized CPs of mathematics instruction. While the authors described changes to the program, the article did not report empirical evidence of changes in TC learning during the period of reorganization.

In another article, Kazemi et al. (2009) similarly described the effort to transform mathematics education at The University of Michigan from a program focused only on analyzing student thinking to a program that uses instructional activities as a vehicle to improve TCs' ability to elicit, respond to, and further the thinking of students. Although the authors imply improvement, no empirical evidence demonstrates changes in TC thinking or performance. Others similarly imply improvement without providing a detailed explanation of the evidence for that improvement (Ghousseini, 2009; Hatch & Grossman, 2008; Kazemi et al., 2014; Lampert & Graziani, 2009; Lampert et al., 2010).

Two categories of literature. Available research on core practice approaches can be put into one of two categories. First, there is conceptual literature designed to deconstruct teaching into specific practices, as noted above (e.g., Ball & Forzani, 2009; Fogo, 2014; Grossman & McDonald, 2009; Kazemi et al., 2009; Windschitl et al., 2009). Some authors have advocated for CPs in general—those practices that most subject area teaching would hold in common (e.g., Ball & Forzani, 2009; Grossman & McDonald, 2009), others have narrowed the grain size to CPs specific to subject areas (e.g., Kazemi et al., 2014; Windschitl et al., 2012).

The second category of literature is focused on the implementation of such core practices in teacher education courses (e.g., Lampert et al., 2010; Ghousseini & Sleep, 2011; Kazemi et al., 2014). Authors of this body of literature explain specific pedagogical strategies underway in teacher education that are intended to cultivate TC capacity for certain core practices. At present, therefore, the literature on core practice approaches in teacher education is primarily theoretical or descriptive (identified exceptions are explored below).

In recent years, a few studies have suggested the promise of particular core practice approaches for the development of TC expertise. For example, teacher educators from the *University of Michigan Mathematics Consortium* (Lampert et al., 2013) examined the interactions of teacher educators and novice teachers during methods course CP rehearsals. The researchers found that the structure of CP rehearsals in the methods course provided the types of opportunities novices need to engage in challenging practice. This evidence suggests that the CP approach used by these teacher educators may indeed align with research on the development of professional expertise, which suggests that repeated exercise on targeted approximations of challenging practice coupled with targeted feedback and reflection can improve practice (Ericsson, 2006).

While the above study suggests potential for TC development, evidence concerning TCs' understandings related to core practices remains largely anecdotal. However, one group of science educators at the University of Washington compared a set of TCs who participated in communities infused with discourses and tools connected to CPs of science teaching to a set of TCs who participated in communities supported by discourses and tools anchored in more traditional, less ambitious teaching practices (Thompson et al., 2013). The researchers studied how the TCs developed instructional repertoires and found important differences in how these

teachers engaged (or did not engage) in the CPs. This empirical investigation into teacher understandings related to CPs of science teaching provides valuable insight into whether, how, and why early career expertise develops in TCs by suggesting that the development of particular tools and communities of practice can influence the trajectory of TC expertise in the CP.

What do we need to know about a CP approach? In a field historically marked by a tendency to implement faster than we can understand (Grossman, 2005; Zeichner 2005a), it is vital that researchers learn more about how candidates develop target practices in the context of core practice approaches. As the field moves ahead, research needs to reach beyond theoretical analysis and description and into empirical investigations of TC understandings in such contexts. Particularly, the field needs to know more about how candidates understand target practices in light of specific core practice approaches. To partially address this gap, this dissertation was designed to study teacher candidates who were learning a specific core practice. In the next section of this literature review, I turn to *eliciting, interpreting, and responding to student thinking*, the core practice selected for this study.

This Study's Core Practice

I selected the core practice of *eliciting, interpreting, and responding to student thinking*. In this section of the literature review, I describe EIR as essential for good teaching and explore each of its interdependent sub-practices. I argue that although research has begun to explore mathematics and science TCs' abilities in this core practice, the field knows little about history TCs' abilities in eliciting, interpreting, and responding.

Why this practice? Good teaching requires a teacher to engage students' prior understandings and anchor new knowledge on existing understandings (Bransford, Brown, & Cocking, 2000; Sleep & Boerst, 2009; Wiliam, 2010). A growing body of research suggests that

teachers' effective use of formative assessments within learning activities can significantly improve student learning (e.g., Black & Wiliam, 1998; Shepard, 2000).

By formative assessment, I mean activities in which teachers engage that provide evidence of student learning, evidence that is used to adapt instruction to meet student needs (Black & Wiliam, 1998). A far cry from test-like summative measurements of knowledge, formative assessments are ongoing and grounded in teaching and learning activities (Bennett, 2011; Coffey, Hammer, Levin, & Grant, 2011; Gipps, 1994; Popham, 2008; Shepard, 2000).

Formative assessment that is conversational in nature, in particular, "produces a qualitative insight into student understanding" (Popham, 2008, p. 6) that can enable a teacher to respond with effective instructional moves (Coffey et al., 2011). If assessment is to occur in the moment-to-moment of classroom instruction, then it is tied directly to teachers' abilities to 1) get students to explain their reasoning, 2) understand that reasoning in light of the content, and 3) follow up in ways that promote learning (Ball, 1993; Bransford et al., 2000; Fanke et al., 2009; Levin et al., 2009; Pierson, 2009; Singer-Gabella et al., 2015). These required abilities suggest that good teaching is a complex, interactional task composed of many practices unnatural for the typical person.

What is EIR? EIR is similar to what others have called "attending to student thinking" (Levin et al., 2009, p. 1) or "leveraging student thinking" (Singer-Gabella et al., 2015). I conceive of eliciting, interpreting, and responding as a single teaching practice made up of distinct, but entangled, sub-practices. For the sake of analysis, I find it helpful to distinguish the three parts from EIR practice as a whole.

Eliciting student thinking requires a teacher to establish a generative problem space or launch a learning activity that promotes shared investigation of student understandings of the

content (Lampert, 2001; Singer-Gabella et al., 2015; Sleep & Boerst, 2012; Stein et al., 2008). During such an activity, a teacher has the opportunity to interpret and assess understandings related to target instructional objectives (Kennedy, 2005; Lampert, 1985; Sleep & Boerst, 2012). Teacher responses, or next instructional steps, include the extent to which a teacher "takes up" student thinking in the learning activity. In the following review, I consider each aspect of EIR and then unpack each of the sub-practices as evidenced in an example of classroom discourse (VanSledright, 2002a).

What is Eliciting? Eliciting begins when a teacher initiates reasoning by inviting students to articulate their understandings about the topic under study. Not just any question, problem, or prompt, however, counts as eliciting student thinking. In eliciting, a teacher establishes "a problem space in which students' ideas are made visible and elaborated, and so brings student thinking to the public forum of the classroom" (Singer-Gabella et al., 2015, p. 6). The productiveness of the task, questions, and prompts depends on their "generativity" (Singer-Gabella et al., 2014, p. 6) or the capacity of the problem to promote shared investigation (Lampert, 2001; Singer-Gabella et al., 2015; Stein et al., 2008).

Mathematics and science education researchers have delineated characteristics of generative tasks that operate as the soil for eliciting student thinking. In mathematics, generative tasks position students to reason about mathematics, invite diverse solution strategies that reflect a range of understandings, and capitalize on learners' prior knowledge while pushing them toward new ways of thinking about the mathematics at hand (Franke et al., 2009; Singer-Gabella, 2015).

In science education literature, eliciting student thinking has been depicted in similar, though science-specific terms. The teacher initiates instruction by encouraging "meaning-making

by pressing for a connection between scientific phenomena and students' lived experiences and knowledge" (Thompson et al., 2013, p. 581). Eliciting is focused on sense-making discourse and, in this case, links classroom activity to the development of students' scientific ideas (Thompson et al., 2013). Generativity in science classrooms, therefore, is marked by similar goals as mathematics but is unique in the ways that students reason about scientific ideas grounded in the methods of scientific disciplines (Coffey et al., 2011).

Common to both mathematics and science is the idea that a generative task will elicit disciplinary reasoning and sense-making that connects content with prior understandings. The same should be true of eliciting in history classrooms. As history is a school subject grounded in a distinct discipline, it follows, that characteristics of generativity must be articulated before eliciting is identifiable. Because I did not find a comparable framework for assessing generativity in history classrooms in the research literature, I adapted Singer-Gabella et al.'s in accordance with principles of historical thinking.

Eliciting student thinking begins with the establishment of a generative problem space through the launch of certain activities, questions, and prompts. This initial eliciting creates the soil in which student reasoning is able to grow and the foundation for further eliciting throughout the lesson (Singer-Gabella, 2015).

Eliciting as conceived here, begins with the launch of the problem space in the moment-to-moment interactions of classroom teaching. In addition to initial eliciting (i.e., launching a problem space), eliciting can also include attempts that the teacher makes to initiate new lines of reasoning throughout a lesson. Such a scenario demonstrates the difficulty of disentangling EIR. While I recognize that a teacher who responds to a student idea may continue to elicit student

thinking, for analytical purposes, I have coded eliciting only as teacher-initiated attempts to bring student thinking into the public space.

What is Interpreting? If eliciting student thinking invites disciplinary reasoning during the course of instruction (Coffey et al., 2011), then interpreting is the teacher's attempt to make sense of that reasoning and determine instructional next steps (Singer-Gabella, 2015). In this way, interpreting serves as the hypothetical link between elicited student thinking and a teacher's observable response.

Interpreting, in a broad sense, consists of a part of the teacher's cognitive activity in the brief moments between student reasoning and instructional response. Interpretation is complex and highly idiosyncratic because it is influenced by a host of relatively inscrutable factors including teacher beliefs, teacher knowledge, personal history, and situated circumstances and because it is not necessarily accompanied by any observable behavior on the part of the teacher (Kennedy, 2005; Richardson, 1996; Robertson, Atkins, Levin, & Richards, in press; Stigler & Hiebert, 1999; Van den Berg, 2002). Instructional decisions are mediated by interpretation and depend on teachers' "meanings–that is, the meanings that teachers ascribe to the events they see in their classrooms" (Kennedy, 2005, p. 33).

Interpreting is relatively difficult to assess or even describe because the researcher is rarely able to access what the teacher sees or thinks, and thus is unable to definitively determine the teacher's intent (Robertson et al., in press). As such, much of what is empirically known of interpreting is actually inferred from teachers' instructional responses (Robertson et al., in press). That is, both eliciting and responding are empirically observable in a classroom and, thus measurable, and even quantifiable, under some circumstances.

So, how can one get at teachers' *meanings* and thus understand decisions they make in their classroom practice? Some researchers posit that interpretation can be inferred based on a teacher's instructional response (Robertson et al., in press). Others suggest that interpretation should be accessed by in-depth interview (Singer-Gabella et al., 2015; Sleep & Boerst, 2012) or stimulated recall interviewing (Kennedy, 2005). Still others admit that interpretation is a black box of sorts, that it is an impossibility to understand someone else's thoughts, and choose not to address its myriad unconscious influences. Because of the difficulties of identifying interpretation, I delineate interpretation as TCs' articulated meanings of instructional interactions as reported in the post-observation debrief interviews.

Uncovering interpretation is so hard to do that teacher educators at the University of Washington are experimenting with making interpretations public for interns (Kazemi, 2015). In one example, Kazemi periodically stopped to think aloud for her interns while teaching a third grade mathematics class. Making challenging instructional decisions public through these "teacher timeouts" provided a second layer of 'text' for interns to consider in addition to the observable classroom discourse (Kazemi, 2015). To fully address interpretation, researchers need just such a "teacher timeout" in order to explore what a teacher recognizes in the classroom and why he/she chooses to respond in certain ways. Of course, classrooms are not set up to accommodate such inquiry and researchers are left to sort through a black box between what a student says and how a teacher responds.

Ideally, a TC will notice generative student thinking and assess that thinking according to target objectives and a framework for student learning (Singer-Gabella et al., 2015). However, this clean-sounding, step-by-step process happens in seconds and is influenced by a host of

relatively inscrutable factors, some of which have nothing to do with the student thinking (Kennedy, 2005; Lampert, 1985).

What is Responding? If eliciting student thinking invites student reasoning during the course of instruction (Coffey et al., 2011), and interpreting is a teacher's attempt to make sense of that reasoning, then responding consists of the instructional moves that a teacher makes to engage in next steps (Singer-Gabella et al., 2015). This study defines "responsiveness" as "the extent to which teachers 'take up' students' thinking and focus on student ideas in their moment-to- moment interactions" (Pierson, 2008, p. 25). Responding requires instructional moves designed to build from student current understandings to more sophisticated understandings. Responding is different from eliciting because student thinking initiates a teacher follow-up. Rather than the student responding to the teacher's initiation (i.e., eliciting), in responding, the teacher responds to the student(s).

Responses to student thinking, or "next steps" as Singer-Gabella et al. (2015) call them, benefit from pre-planning but typically require significant spontaneity (Kennedy, 2005). In response to student thinking, a teacher can take a number of directions that may be more or less predictable. For example, in a mathematics classroom, a teacher might name ideas or strategies, re-present student solutions, offer new problems to condition or constrain a student strategy, or provide or remove scaffolds (Singer-Gabella et al., 2015). Such instructional responses require attention to the substance of students' mathematical thinking and "in the moment" interpretation of that thinking.

Two aspects of teacher response, questioning and feedback, have received the most attention in the research literature. In response to student reasoning, a teacher may ask a question or facilitate questioning in order to clarify student thinking. For example, one teacher might ask,

"Why do you say that" while another teacher would respond "Do you have evidence to support that answer?" (Monte-Sano, 2011a, p. 265). Still a third teacher might come back with, "Does anyone want to try to build on the argument that Mary is making?" (Reisman, 2012a). Research suggests that the nature of teachers' questions can powerfully influence the discourse of a classroom and, thus, student reasoning (Black & Wiliam, 1998; Coffey et al., 2011; Franke et al., 2009; Nystrand, Wu, Gamoran, Zeisler, & Long, 2003).

Another response move a teacher can make is to provide specific guidance on how to improve (Black & Wiliam, 1998; Coffey et al., 2011). For example, one teacher might respond to a student comment with "Good job" (Mont-Sano, 2011a, p. 268). Another teacher may say, "There was some very good thinking when you said…" Ashby, Lee, & Shemilt, 2005, p. 99) while a third responds, "Let's look at the thing Sonny was saying because…" (p. 135).

Going beyond simply distinguishing between questions and feedback in her study of 13 mathematics teachers' responsiveness, Pierson (2009) organized instructional responses according to four categories: No/Low, Medium, High I, and High II. No/Low response included instructional moves that merely evaluated, acknowledged, or brushed off the student thinking. Medium responses were follow-ups that hinted at correct answers or reformulated the student reasoning with only vague connections to the actual student statement.

High I responses put teacher reasoning on display in response to the student's idea. In such a response, the teacher typically took over the student's idea and made teacher thinking the focus of the discourse. High II responses, on the other hand, put student reasoning on exhibition. These follow-ups were characterized by teacher exploration into student reasoning through such moves as probing a student answer, inviting others to make sense of a student's idea, and asking

a clarifying question. Pierson (2009) found a significant correlation between High II responses and student learning in mathematics.

EIR in action. VanSledright's (2002) study of his own 5th grade class includes classroom exchanges between student and teacher. Below, I present a slice of dialogue from this larger case study in order to demonstrate EIR in action. The dialogue began after students read several primary sources that shed light on what happened during the first winter at Jamestown that caused pioneers to starve in large numbers.

As they read, students filled out a chart that prompted *Questions Historical Detectives Ask To Solve the Mysteries of the Past.* After trying to 'solve the mystery' in small groups, VanSledright (2002, p. 42) invited each group to share their argument for what happened during the starving time. He started by asking Table 5 about their argument but the discourse below begins shortly after:

> Jeffrey: Table 1 thinks that we trust this document [the modern version of John Smith's comments on the Starving Time; "... [W]e did not plan well, did not work hard, or have good government" [holding up and then reading from the document].
> Dr. V: … I want to ask you, so they got lazy and didn't organize themselves and the food was gone and they didn't know what to do to get more food, so they just sat around and starved? [Table 1 students nod.] Interesting. Group 3?
> Ben: But it could be that they were lazy or that [Captain] Percy ate all the food or maybe they had a war with the Indians....
> Dr. V: So which do you think it is? Lazy or Percy the glutton or war with the Powhatans?
> Ben: Well, one document says that the Indians fought them and starved them out [Hakim account], and another says that they were lazy [John Smith's account].
> Dr. V: We have conflicting clues. One says the Powhatans were friendly and they brought corn. John Smith said that. He talks about that in one of the documents. Another document said there was an Indian war with the settlers that kept them from getting their food. So which was it?
> Ben: We're not sure.
> Dr. V: … Table 2? What's your position?
> Brittney: The Native Americans could have been thinking ... they would get 10 times more food if they took what the settlers had, but maybe the settlers were just lazy....
> Dr. V: So what's your argument here–so what do you think happened?
> Brittney: [bashfully] Okay, war.

Dr. V: So they were starved to death by the Powhatans? [Students at Table 2 nod.] Okay.

Eliciting in the example. Eliciting is clearly evident in the VanSledright (2002) example. Before the lesson began, VanSledright designed a historical problem with an essential question ("What happened during the starving time?"), assembled accessible sources, and created a graphic organizer that would guide students through targeted questioning of the texts. This strategic design work supported VanSledright when he launched the investigation and elicited student thinking in the classroom.

During the classroom interactions, VanSledright elicited student thinking by positioning students as "historical detectives" (p. 1096) in search of an evidence-based solution for a mystery. His question, "What's your argument for what happened?" prompted student investigation of sources with the support of *Question*s graphic organizer. The same essential question served as the foundational elicitation for the class discussion, an elicitation that VanSledright returned to time and again. As the dialogue demonstrates, VanSledright's elicitation is rich soil for student reasoning to grow.

Interpreting in the example. Because, in the example, VanSledright is both teacher and researcher, readers have a unique opportunity to understand the interpretations that VanSledright has chosen for readers to hear. In the research report, VanSledright explained the lesson design, including his rationale for the selected texts, the *Questions* chart, and how the activity fit into the larger content and disciplinary goals. Thereby, VanSledright's reader understands that the discourse occurred in a designed setting intended to "dislodge" students' belief that textbooks are able to convey "what happened" in the past.

Because VanSledright (2002a) does not tell the reader about his responses point-by-point, one is left to make inferences based on his stated intentions and general analysis in the report.

When he pushed students to take a position, he felt that their comments indicated "weakly grounded but wonderfully imaginative interpretations rather than strong, evidence-based argumentation" (p. 1103).

Many of VanSledright's responses suggest his intention to introduce students to a tension between history as interpretation and history as "anything goes" relativism. His responses to students' arguments pressed them to look closely at the evidence, an exhortation that reflects his hope that students would "point to [evidence] specifically in defense of their arguments" (p. 1099). Given Britney's "bashful" answer, one can guess that VanSledright's tone must have betrayed some of the frustration he was feeling as he recognized students were not moving toward his primary intention.

VanSledright does not explain why, in response to Jeffrey, he started to ask a question but instead stopped and simply clarified Group 1's argument. Neither does he tell the reader why he decided to move on from particular student comments when he did. VanSledright does not tell the reader how he processed Britney's comment or why he accepted her argument when he did, rather than pushing her to provide evidence. Even with VanSledright's unpacking in the research report, the reader is left with dozens of additional questions about his interpretations. What did he notice? What he was thinking? What factors influenced that thinking? An interviewer could question VanSledright on myriad other aspects in an effort to understand the instructional decisions he made during the short dialogue.

Responding in the example. At several points in the short classroom dialogue above, VanSledright responds in ways that make student thinking public and encourage deeper reasoning. His first response is to expand on the student reasoning by reformulating the argument and checking with the group for confirmation. After further discourse, he interjects again but this

time he asks a student to take a position and succinctly revoices three arguments that other students contributed saying, "So which do you think it is? Lazy or Percy the glutton or war with the Powhatans?" (p. 1098). VanSledright's response demonstrates the entanglement of eliciting and responding but for analytical purposes I regard such moves as responding because they were initiated by student thinking.

Additional responses summarize and revoice student reasoning in argument form and further press students to argue a position in light of "conflicting clues" (p. 1098). Despite the fact that the discourse moves back and forth in a teacher-student-teacher pattern, the brief dialogue clearly demonstrates the way that VanSledright's responses promote student reasoning and evaluation of student reasoning. Rather than "funneling" (Wood, 1998) or "taking over" students thinking (Pierson, 2009), most of his responses in this clip are characterized by further exploration of student reasoning.

At certain points, VanSledright challenged student arguments but at other times he simply revoiced the arguments. In the exchange, he never responds with silence and his responses do not encourage talk across groups, although either could be effective. What is directing these decisions and how does VanSledright calculate his instructional responses? The answers to these questions are hidden in VanSledright's unexplained interpretations.

As this example illustrates, the "I" of EIR is never empirically resolved in any final sense. Rather, every new question about a teacher's interpretation opens a metaphorical hallway full of doors, each door leading to new passages of possible inquiry. Thus, my decision to name and bound interpretation at the articulated meanings of instructional interactions in which student statements are elicited.

Together the sub-practices outlined above form the large-grain practice of EIR, the core

practice targeted in this study. As previously mentioned, these sub-practices are often entangled,

as demonstrated in the VanSledright (2002a) example when VanSledright responds to Ben's

comment with:

> We have conflicting clues. One says the Powhatans were friendly and they brought corn.
>
> John Smith said that. He talks about that in one of the documents. Another document said
>
> there was an Indian war with the settlers that kept them from getting their food. So which
>
> was it? (p. 42)

VanSlredright is clearly responding by uptaking Ben's interpretation in light of an earlier

student's interpretation. But, is he not also eliciting student thinking about corroborating across

conflicting accounts in history?

Although I break out the three sub-practices for the purpose of analysis, each operates in

interdependent concert to promote powerful learning opportunities for students and powerful

learning opportunities for TCs, just the type of practice envisioned for a core practice approach.

But, should a practice as difficult as EIR be reserved for more mature teachers, rather than

expected of teacher candidates? That question is at the heart of this dissertation.

Can Teacher Candidates EIR? It has long been suggested that novice teachers cannot

pay attention to student thinking (Jacobs, Franke, Carpenter, Levi, & Battey, 2007; Shavelson,

2006; Sherin & Han, 2004). Some evidence suggests that beginners naturally focus on their own

performance rather than that of students (Kagan, 1992) and cannot implement strategic questions

or elicit student understandings (Hogan, Rabinowitz, & Craven, 2003).

However, a growing body of work suggests that, with support, teacher candidates can

begin to notice and interpret student thinking (Coffey, Edwards & Finkelstein, 2010; Kazemi et

al., 2009; Levin et al., 2009; Levin & Richards, 2010; Monte-Sano, 2011a; Singer-Gabella et al., 2015; Windschitl et al., 2011). For example, some novice science teachers were indeed able to attend to student thinking from early stages of their careers (Levin et al., 2009; Thompson et al., 2013) and some mathematics TCs were able to notice generative aspects of student thinking (Franke et al., 2009; Singer-Gabella et al., 2015; Sleep & Boerst, 2009). Limited research suggests that history teacher candidates also can notice students' disciplinary thinking in history classes (Monte-Sano, 2011a).

The findings from such research fly in the face of stage-based notions of teacher development that suggest novices are fundamentally not able to pay attention to student thinking, much less respond to it. Responding is reportedly difficult even for teachers with many years of experience (Alexander, Osborn, & Phillips, 2000; Banilower, Smith, Weiss, & Pasley, 2006; Fennema, Franke, Carpenter, Ansell, & Behrend, 1998) and purportedly beyond the capability of teacher candidates and novice teachers (Heritage, Kim, Vendlinski, & Herman, 2007).

Recent research, however, suggests that such stage-based notions may not be helpful for understanding what teacher candidates are capable of doing in this arena. Mathematics education researchers, in particular, have now demonstrated evidence that some novices can elicit, interpret, and respond to student thinking when teacher training emphasized such practices (e.g., Coffey et al., 2010; Kazemi et al., 2009; Levin & Richards, 2010; Singer-Gabella et al., 2009). Similar findings in science education research corroborate those in mathematics (Levin et al., 2009; Levin & Richards, 2010; Windschitl et al., 2011).

Less attention has focused on why some novices are able to EIR while others cannot, do not, or will not. In a study of teacher candidates whose attempts to leverage student thinking fell short, Singer-Gabella et al. (2015) suggested a number of interrelated factors that may impact

TC's capability to engage in EIR including 1) discrete knowledge, 2) managerial readiness, 3) epistemic beliefs, and 4) the interaction of will and skill.

External factors surely also support or constrain a TC's ability to EIR. For example, Windschitl et al., (2011) found that adaptable tools of instruction, created by the teacher educators, better enabled candidates to engage in practices that required attention to student thinking. Monte-Sano's (2011a) findings suggested that mentor teachers with traditional views on teaching-learning may constrain TCs' capacity to develop or demonstrate attention to student reasoning. Singer-Gabella et al. (2015) suggests that even when the skill and the will is available to a TC, the context– "social norms, policies, structures"– inform and constrain what is necessary for a teacher to develop a capacity in EIR.

Thus, research suggests that although EIR is challenging for most teachers, some TCs are able to engage in this practice if supported. The field has learned a great deal about mathematics and science TCs' ability to EIR in recent years but very little about history teachers' (c.f., Monte-Sano, 2011a). To narrow this gap, this study explores candidate ability to EIR across the course of teacher preparation.

Thinking about Evidence in History

Because students think about many things, I narrowed my interest in EIR to a particular type of student thinking essential to historical discourse. Research demonstrates that working with historical sources in particular ways can improve students' reading, critical thinking and writing skills but the field knows little about TCs' attempts to attend to such thinking.

In this section of the literature review, I explain my rationale for choosing to focus on TCs' EIR of student thinking about evidence in history. I first explain this study's conception of history, what it means to think historically, and how such thinking differs from the history

usually practiced in schools. Next, I turn to research on student historical reasoning and explain why I chose to narrow the scope of this study to TCs' ability to attend to evidence in history. Before proceeding to the methods chapter, I review what little we know about TCs' ability to elicit, interpret, and respond to evidence in history.

History and historical reasoning. History is a method "of selecting, analyzing, and writing about the past… something that is done, that is constructed, rather than an inert body of data" (Davidson & Lytle, 2004, xviii). In this sense, history is not simply a retelling of the past. Nor is history synonymous with the past. Rather, it is the product of a particular type of investigation (Holt, 1990; Mink, 1987; VanSledright, 2004). While histories are readily available, the past is "immense, infinitely polysemous, sublime and gone" (Seixas, 2000, p. 27).

Consequently, history education researchers make a distinction between learning history as a fixed body of information and learning history as an evidence-based exploration of contested arguments about the past (Lee, 2005; National Center for History in the Schools, 1994; NCSS C3 Framework for State Standards; VanSledright, 2004; Wineburg, 2001). VanSledright and Limon (2006) described three helpful categories for conceptualizing what it means to "know" history: first-order knowledge, second-order knowledge, and historical reasoning (VanSledright & Limon, 2006).

First order knowledge is substantive knowledge and comes from "who, what, where, when, and how questions" (VanSledright & Limon, 2006, p. 547). For example, in 1776 George Washington crossed the Delaware. Second order knowledge is also substantive but pertains to particular "concepts and ideas that investigators impose on the past" (VanSledright & Limon, 2006, p. 547) in an effort to make sense of it. For example, the 'decline of the Roman Empire' imposes historical concepts of progress and decline on a series of events from the past.

Historical reasoning, the third type of knowing in history, is a procedural type of knowledge and is the key to understanding history as a discipline, a method, or way of knowing, that sets it apart from such work as journalism or writing fiction (VanSledright & Limon, 2006). At the heart of historical reasoning is the analysis of historical artifacts and accounts, both primary and secondary (Seixas & Morton, 2014; Wineburg, 2001).

This analysis is characterized by the interrogation of evidence, weighing of conflicting accounts, consideration of author perspectives, and construction of arguments (primarily written) based on evidence (Monte-Sano, 2008; VanSledright, 2004). Where historical reasoning acts in concert with first and second-order knowledge about the past, first and second-order knowledge is dependent upon historical reasoning. Without evidenced-based arguments about the past, we would know little that could be counted as historical fact or concept (Cochran, 2009; Lee, 2006; VanSledright & Limon, 2006; Wineburg, 2001). This procedural knowledge that directs how to "research and interpret the past," (VanSledright & Limon, 2006, p. 547) is the domain of historians and not the domain of most school students.

School history. Where investigative processes and argument-building mark historians' work, "school history" in the US is marked by memorization of prepackaged narratives (VanSledright, 2008). School history is most often an institutionalized exercise in memorizing a large number of historical facts, usually for the purpose of internalizing a narrative of American progress or, conversely, internalizing a critical narrative from the perspectives of marginalized peoples (Cuban, 1991; Lee, 2005; VanSledright, 2008).

As such, research suggests that successful history students master an agreed-upon body of facts and narrative but are unable to evaluate or decide whether certain narratives are more or less valid than other narratives (Bain, 2005; Seixas, 1996; Wineburg, 2001). When they are faced

39

with conflicting accounts of the past, students typically determine one true and the other false (Lee, 2005). Such findings led one educational psychologist to compare high school history students to those who could discuss separate scenes and characters from the play *King Lear* "but do not know what a play is" (Shemilt, 1983, p. 15). In other words, recollection of details about an account of the past without any understanding of the problems, questions, and criteria involved in constructing that account should be called something other than 'history' (Lee, 2005; VanSledright, 2002).

Omnipotent, godlike narratives about "what happened" in the past bear little resemblance to the contested terrain of the historian (Cuban, 1991; Lee, 2005; Seixas, 1996; VanSledright, 2004; Wineburg, 2001). If history is actually a way of knowing the world "through which we organize the residua of the past into a form meaningful to us in the present" (Seixas, 1996, p. 777), then students are not learning history. Rather, students are engaged in an exchange of information, an acquisition of first and second-order knowledge about the past, without an understanding of the process for arriving at that knowledge.

Some might argue that school history should not necessarily resemble the work of historians. "What's wrong with history as an agreed-upon narrative, at least for beginners?" such critics might ask (e.g., Gitlin, 1995; Ravitch, 1990; Schlesinger, 1992). Such attitudes have sustained a relatively consistent approach to teaching and learning in school history over the last century, whether in elementary school or in advanced high school courses (Cuban, 1991; Goodlad, 1984; VanSledright, 2002; Wineburg, 1991b; Wineburg & Martin, 2004).

Teachers generally transmit knowledge about the past through lecture and question students about the content of textbooks (Cuban, 1984; Cuban 1991; Downey & Levstik, 1991;

Goodlad, 1984). Students, therefore, spend most of their time in history classes passively listening and reading textbooks or doing worksheets that emphasize the recall of historical facts (Downey & Levstik, 1991; Goodlad, 1984; Lapp, Griggs, & Tay-Lim, 2002). Enthusiasm in recent decades for the use of primary sources and the availability of resources does not appear to have significantly changed pedagogical or assessment practices (Fickel, 2006; Grant, 2003; 2006; Van Hoever, 2006).

Despite the prevalence of common approaches to teaching and learning 'history,' research does not suggest that a focus on content, at the expense of method, is effective in facilitating student learning, even for learning of facts and concepts (first and second-order knowledge). In fact, consistent research over the last century suggests that students do not retain information conveyed in history classes (Lee & Weiss, 2007; Ravitch & Finn, 1987; VanSledright, 2008; Wineburg, 2004), are bored by memorizing prepackaged narratives (Rosenzweig, 2000), and feel marginalized by the "selective memory" (Kammen, 1989, p. 145) of these historical narratives (Epstein, 2000; Hawkey & Prior, 2011). Consequently, as students develop an "encyclopedia epistemology" (VanSledright, 2002b, p. 76) of history by memorizing prepackaged narratives about the past, they simultaneously fail to grasp the content the curriculum was designed to deliver. In contrast, research across subject areas suggests that students' ability to engage in disciplinary discourse, or reasoned dialogue based on evidence, is clearly connected with the development of deep understanding (Bransford, Derry, Berliner, & Hammerness, 2005; Engle, 2011; Leinhardt & Steele, 2005).

The Value of Historical Reasoning. In history, when teachers see their role as facilitator of student reasoning rather than repository of knowledge, students can deepen their understandings significantly (Bain, 2005; Grant, 2001; Grant, 2003; Shemilt, 1983; Wineburg,

2001). It has long been argued that students, especially those in lower grades are not capable of the nuances and challenges required for disciplinary literacy in history (Elton, 1970; Hallam, 1967). However, study after study over the last three decades has shown that this assumption is false and suggests that even young children are capable of elemental forms of the type of thinking required for historical work (for reviews of this research see Barton & Levstik, 2004; Grant, 2003; Levstik & Barton, 1997). If well equipped, students in late high school are largely able to engage in authentic forms of history like novice historians, albeit in less sophisticated ways (Bain, 2005; Holt, 1990; Reisman, 2012a; Shemilt, 1983)

Multiple researchers have documented improvement in historical reasoning and argumentation of students when teachers approach reading and writing in history as a disciplinary activity (De La Paz, 2005; De La Paz & Felton, 2010; Monte-Sano, 2008). Students in classrooms that represent history as an interpretive activity based on evidence can learn to use evidence as a basis for their reasoning (Ashby et al., 2005; Bain, 2005; Lee & Dickinson, 1984, Monte-Sano, 2008). With further support, students can even learn to write with evidence-based claims that are grounded in historical contexts (Monte-Sano, 2008).

In recent years, promising results emerged from research that looked at student performance in light of curricula designed to facilitate the type of reading, writing, and thinking, formerly reserved for historians (e.g., De La Paz et al., 2014; Monte-Sano, 2008; Monte-Sano, 2011b; Reisman, 2012b). These findings suggest that not only can students learn to reason historically; the process can have a positive impact on their reading, writing, and thinking skills. Although some studies have identified and explored TCs' attempts to promote historical reasoning (e.g., Monte-Sano & Cochran, 2009; Monte-Sano & Harris, 2011), the breadth of interest in these inquiries limited identification and description of the historical reasoning the TC

42

promoted in the classroom. I narrowed the scope of student thinking to a foundational point in historical reasoning – thinking about evidence in history.

Why Focus on Evidentiary Thinking? This study could have examined TCs' abilities to EIR student thinking about any number of disciplinary concepts or skills (e.g., Lee, 2005; Seixas & Morton, 2014). But, given the common problems in the way that students think about history, "developing students' interpretive and evidence-based thinking is foundational to advancing their disciplinary understanding" (Monte-Sano, 2011b, p. 261). As Lee (2005) noted, "the concept of evidence is central to history because it is only through the use of evidence that history is possible" (p. 54). I chose thinking about evidence in history as the focus for teacher candidate EIR, therefore, because it is an important foundation for all aspects of historical knowledge (1st order, 2nd order, and historical reasoning). If a TC cannot attend to a student's understanding of evidence in history, the former is not likely to attend to aspects of knowing in history beyond content acquisition.

Thinking about evidence. If teachers must help their students think about evidence in disciplinary ways (Reisman, 2011b), then one must understand how historians think about evidence. Sources can provide intentional evidence (e.g., memoirs) or unintentional evidence (e.g., pottery) when historians question the sources in particular ways (Lee, 2005; Levesque, 2008). Wineburg (1991a) identified three heuristics that historians use when they analyze texts for evidence.

First, historians engage in sourcing, a critical form of literacy that allows them to explore the type of document, the author, the author's point of view, and the date. Each aspect of sourcing can provide historians important information for understanding the evidence

(Wineburg, 1991a). If students are to think about evidence like historians, they must learn to source artifacts from the past (Reisman, 2012b; VanSledright, 2004, Wineburg, 2001).

Historians engage in a second heuristic called contextualization (Wineburg, 1991a). Historians treat the past, from which the document came, as more unfamiliar than familiar. They locate the residua of history in a particular context by learning about, and empathizing with the people of that time and place (Mink, 1987). Contextualizing a document helps historians avoid presentism, or anachronistically seeing the past only through one's present lens (Wineburg, 2001). For students to think about evidence like historians, they must approach sources like artifacts from a "foreign country" (Lowenthall, 1985) that require unique investigation in order to appropriately interpret (Reisman, 2012b).

A third heuristic historians use is corroboration (Wineburg, 1991a). Rather than try to identify an account of 'what happened,' historians interrogate many sources to build a case, much like a detective (Wineburg, 2001). They compare and contrast, moving back and forth between documents in order to ground their interpretations about the past in the available evidence. If students are to think about evidence like historians, they must learn to corroborate available evidence and test all arguments against new evidence (Reisman, 2012b; VanSledright, 2004, Wineburg, 2001). While some researchers make a distinction between "historical thinking" and "historical reasoning," throughout this dissertation I used these terms as synonyms and as a short hand reference to refer to the heuristics of historians described above.

Eliciting Student Thinking About Evidence

A number of studies have explored TCs' disciplinary knowledge and related practice (Bohan & Davis, 1998; Fehn & Koeppen, 1998; Monte-Sano & Harris, 2011; Seixas, 1998;

Yeager & Davis, 1995). Some researchers have narrowed this focus on TCs' attention to students' disciplinary thinking, although such research is rare (Barton, McCully, & Marks, 2004; Monte-Sano & Cochran, 2009; Seixas, 1994). Although some of these studies did not investigate TCs in authentic classroom teaching (Barton et al., 2004; 2009; Seixas, 1994), researchers found that TCs were able to notice important aspects of the substance of students' disciplinary thinking. These findings are significant because they align with those in other subject areas that suggest that novices can attend to student thinking when given proper support (e.g., Levin et al, 2009; Sleep & Boerst, 2012; Windschitl et al., 2013).

A few studies have included examination of TCs' ability to notice and respond to student disciplinary thinking in actual classroom settings. Monte-Sano and Budano (2013) included "attending to students' ideas about history" as one of four pedagogical content knowledge (PCK) categories that they constructed and used for analysis of two TCs[1]. Both teachers in this study were able to identify student disciplinary thinking during interviews and course assignments but once in the classroom, the TCs either ignored student thinking or responded with "stock phrases such as 'How do you know?'" (Monte-Sano & Budano, p. 192).

Despite findings that the two teachers grew in their abilities to attend to students' ideas about history, the research report did not provide enough detail for the reader to deduce the aspects of student thinking about which history TCs were able to notice and respond. Given the breadth of PCK, Monte-Sano and Budano (2013) offered an important but relatively broad investigation of TCs' abilities to attend to students' thinking about history, as compared to the intention of the present study.

---

[1] The researchers then followed the teachers into their first two years of teaching

In another article, Monte-Sano (2011a) described how two TCs learned to identify significant student reasoning in the classroom to varying degrees while a third TC showed little attention to student thinking about historical interpretation and evidence. Both of the TCs who developed an ability to attend to student thinking struggled to provide instructional responses when thinking was made public. The study is rare in its attention to history TCs' abilities to "uptake" student thinking. The authors' articulation of what counted as target student thinking (interpretive and evidentiary historical thinking) and what counted as an instructional response is also rare for the field.

The studies noted above (Monte-Sano, 2011a; Monte-Sano & Budano) are also noteworthy because the authors consider TC ability to notice and respond to student thinking in light of a methods course that specifically focused on the facilitation of interpretive and evidence-based thinking in students (see also Monte-Sano & Cochran, 2009; Monte-Sano & Harris, 2011). Although not necessarily a core practice approach, in that the authors did not report a cycle of enactment and investigation, these studies looked for particular evidence of impact in TC teaching performance based on an approach taken in the methods course.

What do we need to know? Like the broader body of literature, limited research in history education suggests that some teacher candidates can elicit, interpret, and respond to student thinking while others cannot, or will not. We need to know more about how history TCs exposed to EIR in methods courses attempt this practice, how they understand what they are doing in these attempts, and why some simply do not attempt it. Furthermore, we need to know whether and how core practice approaches may be valuable for the development of targeted skills such as EIR.

In this chapter, I argued that history education needs research that examines teacher candidates' abilities to elicit, interpret, and respond to students' thinking about historical evidence in light of a core practice approach. I first provided an overview of core practices in teacher education, reviewed what research suggests about how TCs develop core practices, and argued that the field needs more research to explore how candidates understand target practices in the context of a core practice approach. With a core practice approach outlined, I then examined research on the core practice selected for this study.

Based on the review, I argued that we know little about history TCs' capabilities to elicit, interpret, and respond to student thinking. In the final section of this literature review, I narrowed to the student thinking prioritized in this study. Although we know that historical reasoning can lead to improved reading, critical thinking and writing skills, we know little about how teacher candidates learn to facilitate such thinking. In light of a core practice approach, the intersections of the gaps noted above suggest the need for research that examines TCs' abilities to elicit, interpret, and respond to student thinking about historical evidence. The next chapter explains the methods I used to engage in such a study.

Chapter 3: Research Design and Methods

In Chapter 2, I argued that teacher educators need to know how teacher candidates understand core practices, particularly the practice of eliciting, interpreting, and responding to student thinking about historical evidence. In this chapter, I describe the design of a study intended to address this gap by answering the following research question:

How do the teacher candidates elicit, interpret, and respond to student thinking about historical evidence during their teacher preparation program?

- How do they engage in these practices at the outset of their program?

- In what ways do their capabilities to engage in these practices change during their program of study?

In what follows, I describe each of my methodological choices for this study. First, I describe my approach to case study methodology. Next, I explore contextual matters including setting, study participants, and embedded cases. I then explain aspects of the study's design including the 1) research questions; 2) conceptual framework; 3) units of analysis; 4) data collection; and 5) data analysis. Finally, I consider issues of validity and generalization.

Approach to Case Study

Because the purpose of this study was to examine teacher candidates' abilities to elicit, interpret, and respond to secondary students' thinking about historical evidence in light of a methods course designed to develop these abilities, I needed a methodology that explored teacher candidate (TC) ability over time without losing the contexts in which learning occurred. I chose case study methodology because it allowed me to "retain the holistic and meaningful characteristics of real-life events" (Yin, 2003, p. 3) — as in the dynamic process of learning to

teach in the midst of complex situational and relational factors. That is, I chose case study methodology because it allowed me to focus on a phenomenon but retain the complex situational and relational factors present in the dynamic process of learning to teach. To examine the phenomenon of TC eliciting, interpreting, and responding to student thinking, I designed an exploratory multiple case study (Yin, 2006). The cases were bounded at the level of the individual teacher candidate and served as a window into the primary phenomenon of interest.

Context

From a teacher's apprenticeship of observation, to experiences in methods courses, to field placements and student teaching, research suggests that contextual influences play an important role in the process of learning to teach (Clift & Brady, 2005; Grossman, Valencia, Evans, Thompson, Martin, & Place, 2000; Lortie, 1975; Valencia, Martin, Place, & Grossman, 2009). Below, I outline a few of the many factors that made up the context in which the TCs I studied learned to EIR.

The college. I conducted this study as a doctoral student at a large state research university in the mid-Atlantic region of the United States. At the time of this study, a number of faculty in my College of Education were involved in multiple and overlapping dialogs about the use of core practices as an organizing feature for both methods course curriculum and program redesign. Summer workshops offered faculty and instructors an opportunity to work together on course redesigns based on principles of the core teaching practices described in Chapter 2. The conversations and partnering opportunities available through this emerging work played an important part of the context of this case study because they offered a community of practice for developing the methods course approach.

The program. Participants in this study were drawn from a master's certification program, an operation of the college of education in which I was enrolled. TCs were accepted into the program based on their bachelor's degree grade point average, Praxis I and II scores, academic background in social studies, and experience with children. TCs were simultaneously enrolled in courses and worked in year-long internships in local public schools. In addition to social studies courses and practicums, TCs took courses on literacy, diversity, adolescent development, teaching as a profession, and action research. At the time of the study, the social studies strand of the program was in a time of transition and was unofficially under the direction of the field placement coordinator (also the Methods II course instructor).

The course. Social Studies Methods II (Methods II) was the second of three required methods courses for students seeking social studies certification. The first of these methods courses was a prerequisite and was offered in the summer of 2014 for those students who did not take Methods I as an undergraduate. The methods courses were designed to support TCs in learning how to frame history as inquiry, use historical documents, understand developmental thinking in history, and develop instructional and reflective proficiencies concurrently with experience in the internship.

To a lesser extent, the courses address social studies subjects other than history. Four of the 13 TCs enrolled in Methods II took the prerequisite Methods I course in the summer of 2014, where I served as the instructor. The remaining TCs, all graduates of the undergraduate program at the same university, took Methods I as undergraduates. The undergraduate course did not provide the foundation for the subsequent spiral of disciplinary thinking and literacy skill instruction that I offered students in the summer course. Given the importance of the course in this study, I considered adding a research question in order to highlight the importance of the

teacher preparation instruction. Concerned that the addition of such a research question would obscure the study's focus on the phenomenon of interest, I decided to instead dedicate Chapter 4 to the explanation of the course and its details.

Study Participants

In order to explore TC eliciting, interpreting, and responding to student thinking about historical evidence in light of a core practice approach, I considered several sections of preservice social studies methods courses that I had access to as a graduate student instructor. I limited my sample to TCs in the Fall 2014 section of Methods II, all 13 of whom were willing to participate in my study. Methods II provided the disciplinary depth, curricular content, and field context to enable the generation of rich data to explore the study's research questions.

Embedded case selection. Following the lead of others in the field (e.g., Grossman, Smagorinsky, & Valencia, 1999; Monte-Sano & Budano, 2013; Smagorinsky, Cook, & Johnson, 2003), I chose several candidates for in-depth analysis in hopes of providing a meaningful picture of the phenomenon of interest, while still allowing the study to remain manageable in terms of time and resources. I began case selection by considering all participants in the methods course, which included a body of 13 master's level students, placed in four school districts, and working with 18 different mentor teachers (See Appendix A).

The best case studies come from "information-rich cases," or cases "from which one can learn a great deal about issues of central importance to the purpose of the research" (Patton, 1990, p. 169). Because this study explored a phenomenon that required student thinking about evidence in history, I first eliminated candidates placed in non-history placements. I also eliminated two candidates placed in seventh grade World History and Geography classrooms because I was concerned that the curriculum lacked the content rigor of a typical secondary

history class. I sorted the seven remaining candidates by two key criteria that research on teacher education suggests can be influential for beginning teachers: disciplinary understanding and placement context (Adler, 1991; Ball & Cohen, 1999; Goodlad, 1984; Grossman, Wineburg, & Woolworth, 2001; Koeppen, 1998)

Disciplinary understanding. Research suggests that teaching effectively with primary sources and historical concepts requires a deep understanding of the discipline of history, understanding that includes an awareness of how knowledge is constructed in history (Ball & Cohen, 1999; Bransford et al., 2005; Grossman et al., 2001; Wineburg, 2001; Wilson & McDiarmid, 1996). Those teachers and TCs who understand that history is an interpretive investigation that requires analysis of evidence have at a minimum a framework for organizing a conception of teaching historical thinking (Grant, 2003; Hartzler-Miller, 2001; van Hover & Yeager, 2003; Yeager & Davis, 1996).

To assess initial TC disciplinary understanding, I conducted a pretest (Appendix G) of all 13 candidates using an assessment adapted from Nokes, Dole, and Hacker (2007). The assessment provided the opportunity to demonstrate the heuristics of a historian and demonstrate disciplinary understanding through demonstration and explanations of historical method. Pretests were scored according to Nokes et al.'s (2007) rubric for disciplinary understanding and results were corroborated against TC questionnaires and coursework.[2] I assessed extensiveness of disciplinary understanding relative to other TCs in the study (See Appendix B).

*Placement contex*t. Research suggests that placement context and mentor are prime influencers of TC performance during student teaching (Adler, 1991; Goodlad, 1984; Koeppen, 1998; Wilson, Konopak, & Readence, 1994). I expected that the mentor teacher's emphasis on

---

[2] One limitation of this analysis approach was an emphasis on heuristic count rather than content of historical explanations. To corroborate the scores, I assessed the pretests against an entirely different rubric (Seixas & Morton, 2014) and found similar patterns among the TCs.

evidence in history and historical reasoning would be an important factor in the candidate's use

of the same. Because I had no external measure to assess richness of classroom context for

historical reasoning, I assessed the contexts in relation to one another, using three sources:

coursework, program coordinator report, and candidate questionnaire.

I assessed each TC's context relative to one another by grouping contexts according to

the following categories: 1) mentor never uses historical reasoning, 2) mentor includes some

aspects of historical reasoning but with limitations (e.g., teaches only skill and not inference), or

3) mentor regularly uses historical reasoning. I created a table with each candidate's context

rating and descriptions for the rating.

Final selection. Using my two key criteria, I created a 2x2 table in order to classify each

of the seven candidates according to 1) richness of context for historical reasoning and 2)

candidate disciplinary understanding.[3] Because I sought the cases with the greatest range in the

two categories described above, I selected Kenra and Sally who demonstrated the two lowest

scores in heuristics on the pretest when compared to other potential participants.

---

[3] Two qualifications are in order here. . First, Kendra changed mentors in October from a government placement
into history placement and was assigned a mentor noted for his use of historical reasoning. Second, Craig's context
was challenging to assess. His mentor teacher emphasized historical reasoning in the on-level classes but de-
emphasized it in the inclusion classes. Because Craig's mentor seemed to understand and value historical reasoning
in instruction, I rated the context highly.

	Rich Context	Low/Medium richness of context
Extensive disciplinary understanding	**Craig** (Grade 9, US history inclusion)	**Gabby** (Grade 9, US history, gifted)
Average/Limited disciplinary understanding	**Kendra** (Grade 8, US history)	**Sally** (Grade 9, US history)

Table 3.1: Case Selection 2x2

Craig's context was reportedly rich for the regular education classes, Craig was teaching ESL and inclusion classes. Because Craig's mentor reportedly met the criteria for rich context in regular classes, I decided to use Craig in the upper left quadrant of the 2x2. Gabby, the other candidate with extensive disciplinary knowledge, was reportedly placed with a mentor who did not pursue historical reasoning as a goal in class. However, the district curriculum included historical reasoning skills across secondary grade levels, which means students probably encountered historical reasoning at some point prior to the class. Thus, although Gabby's mentor teacher did not pursue historical reasoning, the context would likely be more amenable than reports of the mentor's teaching suggested.

In the absence of the ideal, I chose to select the two candidates with extensive disciplinary understanding and place them in the two upper quadrants of the 2x2. The distinctions between the two candidates selected for the lower quadrants were even more apparent. Because I prioritized the criteria described above (i.e., disciplinary understanding and richness of context), the participants' placements were somewhat diverse. The four selected taught in four different districts but with the exception of Kendra, all taught ninth grade US

history. While Kendra and Sally taught on-level courses, Craig taught inclusion ESL, and Gabby taught gifted/talented. Although the varying levels initially caused me to reconsider Gabby and Craig, further investigation revealed that the TCs used the same lesson plans, materials, and assessments in these ESL and gifted/talented classes that were used in the regular classes.

Research Design

In this section, I discuss my study according to five components of a research design adapted from Yin (2003): 1) the question; 2) the conceptual framework; 3) the units of analysis; 4) the data sources; and 5) the data analysis.

Research question. The research questions functioned like a signpost in this study, frequently pointing me back to the central matter at hand: What is it I want to know here and what will specifically direct me to that knowing (Yin, 2003)? My question and sub-questions were designed to point to the substance of how TCs elicited, interpreted, and responded to student thinking about historical evidence across a program of study. The form of the questions helped point to specific manifestations of EIR 1) in individual TCs, 2) across time, and 3) across TCs.

Theoretical framework. In Chapters 1 and 2, I described an EIR framework that I constructed from the research literature and detailed its constituent parts. In this study, the EIR framework served as both lens and bridge. It was a lens in the sense that it helped me determine where to look for the phenomenon, by prompting exploration of certain aspects of TC teaching and understanding. Conversely, this framework helped me avoid a focus on interesting aspects of TC development not relevant to the phenomenon. Because the EIR framework was made of building blocks from the research literature, it eventually served as a bridge to link the findings

of the case study back to gaps in the research literature, where such analytic generalizations were appropriate (Merriam, 1998; Yin, 2003).

Units of analysis. A unit of analysis is the point that clarifies the case study's beginning and end (Yin, 2003). Topics like teaching are difficult to define and bound within a unit of analysis. When I envisioned the end of this study, I wanted to be able to "say something about" (Patton, 1990, p. 168) TCs' ability to elicit, interpret, and respond to student thinking about evidence. For this reason, I selected two units of analysis.

The first unit of analysis was individual TC's experiences with EIR at multiple points across the program of study. The second unit of analysis was an examination across the cases (four TCs) in order to explore patterns among the TCs' understanding of the core practice. This second unit of analysis was intended to explore how the collection of cases could lead to understanding about TC development in EIR.

Data Collection

I collected data in order to shed light on the cases, including both the phenomenon of interest and the contexts in which I studied that phenomenon. I used the main research question to determine what potential sources of data might serve to answer each sub-question (Yin, 2003). Table 3.2 serves to demonstrate the link between data and the research question/sub-questions the data were intended to address (data related to contextual matters included in distinct rows).

Data	When collected?	Relevant sub question	How analyzed?
Observation Notes from observations 0 & 1	Oct 2014 – Nov 2014	• How doTCs elicit, interpret and respond to student thinking about historical evidence at the outset of their program? (SQ1)	• Initial coding • Within & across pattern coding • Testing propositions, searching for contradictory evidence
TC Pre Assessments	Aug 2014 – Apr 2015	• How doTCs elicit, interpret and respond to student thinking about historical evidence at the outset of their program? (SQ1)	• Multiple analytic passes • Within and across pattern coding • Testing propositions from developed from other data, searching for contradictory evidence and explanations
Coursework related to EIR: Lesson plans, rehearsals, Case talk-through	Sep 2014 – May 2015	• How doTCs elicit, interpret and respond to student thinking about historical evidence at the outset of their program? (SQ1)	• Multiple analytic passes • Testing propositions from developed from other data, searching for contradictory evidence and explanations
Observation Debrief from observations 1	Oct 2014 – Dec 2014	• How doTCs elicit, interpret and respond to student thinking about historical evidence at the outset of their program? (SQ1)	• Within & across pattern coding • Testing propositions developed from observation analysis, searching for contradictory evidence
Observation Notes from observations 2 – 4 (depending on TC)	Jan 2015 – May 2015	• In what ways do the teacher candidates' capabilities to elicit, interpret and respond to student thinking about historical evidence change during their program of study? (SQ2)	• Initial coding • Within & across pattern coding • Testing propositions, searching for contradictory evidence
Observation Debrief from observations 2 – 4	Jan 2015 – May 2015	• In what ways do the teacher candidates' capabilities to elicit, interpret and respond to student thinking about historical evidence change during their program of study? (SQ2)	• Within & across pattern coding ▪ Testing propositions developed from observation analysis, searching for contradictory evidence

Table 3.2: Link between data and research questions

EIR data. The main thrust of data collection focused on evidence of TC ability to EIR at

targeted points throughout the period of study. Data collection related to TC ability to EIR was

divided into two basic categories based on the research sub-questions: *outset* data and *later* data. The first sub-question calls for evidence of an ability to EIR at the outset of the program.

In the ideal study, TCs would provide a tidy initial benchmark on the first day of the program by teaching a lesson, giving an interview, and engaging in a think-aloud. In consideration of TCs' interests, their students' interests, and the practicalities of school life, I instead interpreted "outset of their program" in broad terms. Candidates completed the pretest prior to the beginning of the Methods II course but not prior to the program (undergraduate Methods I was a prerequisite). "Outset" data included course-required lesson plans, rehearsals, and observations that occurred during the first three months of the course. Observation debriefs were voluntary and only included TCs chosen as embedded cases. Consequently, data collected at "the outset" included data bounded by the first three months of the course.

Later EIR data, which served to benchmark analyses of change, included data collected during the remainder of the study (December 2014-May 2015), most of which were not related to course assignments. The post-test was offered in late April. Below I describe the data in more detail.

Data Source 1: Observation notes. The research questions required evidence of TCs' ability to EIR.[4] Although I considered lesson plans and rehearsals as informing my assessment of TCs' EIR, I felt that observations of classroom teaching offered the best opportunity to see what TCs actually did. I observed each candidate's classroom teaching at least four times between November 2014 and May 2015. I asked TCs to notify me when they would be using primary sources in their internship teaching but I did not mention my interest in historical reasoning or EIR (Observed lessons and topics listed in Appendix F).

---

[4] I followed similar procedures as those described here in the observation notes I made while viewing course session videos.

In my observation notes, I tried to "capture the slice of life" (Bogdan & Biklen, 2007, p. 120) in a way that could later inform both understandings about the phenomenon of interest and contextual details of the case.[5] Given the challenges of time, I tried to tailor the focus of my observations enough to capture teacher-student exchanges but maintain attention on the broader learning experience in each classroom.

I initially created an observation protocol to capture details about the use of primary source documents, specific historical heuristics, hypotheses about interpretations, and observable responses. When I conducted my first observation, however, I found that the protocol was not particularly helpful and I adapted my approach mid-lesson. From that point on, I focused on classroom dialog during lessons and especially prioritized exchanges that included the TC and were related to evidence in history.

In an effort to address restrictions in this focus, I always made additional descriptive notes immediately after the lessons and tried to fill in any gaps (Stake, 1997). When reviewing observation notes, I returned to the descriptive notes to fill in details related to empirical examples of elicitation and responding, and hypotheses about interpretation.

Although frequently embedded within my descriptive observation notes, post-observation reflective notes and memos allowed for a more subjective perspective on the evolving project (Bogdan & Biklen, 2007). These notes included many of the aspects avoided in descriptive observation notes such as my prejudices, ideas, hunches, mistakes, impressions, and feelings. I marked the reflective portions of the notes with an annotation rather than keep these reflections separate from the descriptive notes because the two sets of notes contained significant crossover.

---

[5] Although most observations were conducted in-person, limitations of time and space required that I conduct some observations by review of classroom recording (e.g., several of the TCs' EDTPA submissions were used for observation data).

Data Source 2: Interviews. My approach to data collection depended partly on interviews because interviews offered a targeted opportunity to explore the meaning making of preservice teachers. Post-observation debrief interviews provided secondary evidence to confirm or challenge my understanding of the elicitations and responses I observed in class. I drew on debriefs interviews to collect data on TC interpretation.

Although questions were open-ended, the interviews were deductive in the sense that I followed lines of questioning that focused on aspects of EIR. I used an audio device to record and backup all interviews and transcribed all interviews so that I could focus on the conversation in the present and later return to precise language and my interpretations (Bogdan & Biklen, 2007; Kvale, 1996).

Post-observation debriefs. After each observation, I interviewed the TC in order to test my empirical observations and explore TC interpretations. I designed an interview protocol that began with a broad understanding of the TCs' vision and goals and eventually narrow to specific exchanges (Brenner, 2006). I piloted the interview protocol on two separate occasions with TCs before using it to collect data on case study participants (Bogdan & Biklen, 2007).

The interview protocol included standard questions about vision for the lesson in order to elicit a clear articulation of what the TC was trying to accomplish during the lesson and what student thinking he/she envisioned (see Appendix C for entire protocol). With each question, the protocol narrowed to explore the candidates' perceptions of student thinking in the lesson. The final portion of the protocol focused on specific exchanges with students.

I always took at least 30 minutes between the end of the lesson and the beginning of the interview to make strategic decisions about the protocol and select particular exchanges that I

wanted to explore during the debrief. I prioritized classroom discourse related to evidence in history and was especially curious about exchanges that included multiple teacher responses.

Managing assumptions and power. My knowledge of the subject matter and my relationships with each of the interviewees were, at different times, both assets and obstacles. I tried to strike a balance between "generous listening" and informed questioning, particularly when conversations departed from the planned line of questioning. Rather than jumping to conclusions about shared meanings, I checked many of my interpretations by restating what I heard, asking the question again in another way, or asking for examples (Kvale, 1996; Patton, 2002).

I was in a unique position because I was co-instructor of the course but I did not grade any coursework. Even with this stance, matters of power and influence were potentially problematic for the richness of data available through research. Despite my attempts to develop a peer-like relationship, I sometimes felt as though candidates wanted to meet my expectations by providing certain answers. By focusing attention on TC experience and understanding, rather than the right or wrong answers about teaching, I tried to reduce the status conventions typically present in a teacher-student conversation (Bogdan & Biklen, 2007).

From the beginning of the course and in each research-related interaction, I reminded participants that I was not the teacher of record in the course and that I was not involved in grading or consulting on grading any course activities. I also regularly reminded them that I would not be reporting back to Meredith, the course instructor, about any of the details of our conversation. In time, I cultivated a friendship with my participants that did not feel hierarchical to me.

Data Source 3: Coursework documentation. Since course artifacts are a source of specific details that can corroborate or correct information from other sources (Yin, 2003), I used the coursework mostly to triangulate evidence from other sources. The course instructor and I designed several methods course assignments that targeted EIR and provided evidence of EIR understanding.

Artifacts created for course assignments, especially when these assignments are graded, can be problematic sources of data because they may reflect the course or professor's expectations rather than self-revealing narratives about a TC's actual views. I followed Bogdan and Biklen's (2007) recommendation to use subject-produced data as part of a study in which the major thrust is participant observation or interview, rather than documentary evidence alone. The subject-produced data sources helped me select cases and triangulate findings noted from other data sources.

Document Source 4: Pretest: All TCs took a pretest prior to the beginning of the course adapted from Nokes, Dole, and Hacker (2007) and De La Paz, Monte-Sano, and Felton, (2010). The same test was taken again nine months later to assess aspects of TC growth. The purpose of the pretest was to provide benchmarks of TCs' development in disciplinary understandings and pedagogical content knowledge, specifically related to teaching with historical reasoning.

In addition to the TCs' disciplinary understanding, the pretest was designed to assess aspects of TCs' abilities to EIR. For example, two questions provided a student answer in response to a question about primary source documents. The TC explained what they noticed in the students' thinking (interpreting) and what they would do next with that student (responding).

The final section of the pretest invited TCs to outline a lesson they could use with their students that included one or more primary source documents. The prompt asked the TCs to

include activities or goals for facilitating students' historical thinking, reading, and writing. Though limited in its capacity to assess abilities in eliciting, this lesson plan provided a sense of how candidates might approach task design and eliciting student thinking.

To supplement the pretest data on eliciting, I also administered a case study assignment that provided TCs a hypothetical context and asked them to decide on the three most important questions to ask the students in order to recognize how they think about evidence in history. In addition to explaining why they chose these questions, TCs also designed an activity that could determine how the students understand *how we know what we know* about the past.[6]

Context Data. In addition to the data I collected directly related to EIR, I collected data in order to document the contexts in which TCs were learning to teach. These data included TCs' mentor interviews, placement expectations negotiation worksheet, opening week observations papers, and a student inventory assignment. TCs also wrote regular reflections throughout the semester that frequently included descriptions of their placements, students, mentors, and connections between teacher education settings.

I also kept context notes when I visited TCs' schools for observations. In addition to these data intentionally related to the TCs' contexts, I frequently noted contextual details in the interview transcripts, course video transcripts, email communication with candidates, and email communication with the methods course instructor.

Data Analysis

Analysis of data for this project was designed in accordance with the research question conceptual framework and organized by a process of data reduction, data display, and iterative

---

[6] Unfortunately, I only conducted this exercise with TCs in history placements. At that time, Kendra was still in a government placement. Therefore, I have the case study assignment data for three of the four TCs in my study.

memoing of conclusions (Miles & Huberman, 1994). Below I describe the data analysis processes I used to explore TCs' EIR in this study.

Eliciting analysis. Because eliciting student thinking about evidence in history requires a generative historical task, I began analysis of eliciting by first considering the generativity of the task that I observed in the lesson. Based on my review of EIR and history education literature, I determined that eliciting student thinking in history is typically dependent upon the establishment of a generative problem space through the launch of certain activities or tasks.

This generative task creates the soil in which student reasoning is able to grow and the foundation for eliciting throughout the lesson (Singer-Gabella, 2015). Therefore, the first step of the eliciting analysis protocol was, "Was there a generative problem space?" This required a clear definition of a generative task in history class (See Appendix D for full protocol and codes).

In order to identify generative tasks in history, I drew on history education literature to adapt characteristics of generativity articulated by Singer-Gabella et al. (2015). As noted previously, historical arguments require "turning a source into evidence," (Seixas & Morton, 2013, p. 46) or making inferential conclusions based on the sourcing, contextualization, and corroboration of relevant artifacts (Holt, 1990; Lee, 2005; Monte-Sano, 2008; Peck & Seixas, 2008; Seixas, 2006; VanSledright, 2004; Wineburg, 1991a). In this light, a generative task in history instruction should:

1) position students to reason about evidence in history, rather than simply identify and memorize a historical narrative. That is, students will examine evidence, make inferences based on sourcing and contextual information, and/or corroborate across sources of evidence.

2) invite diverse interpretations and argument-building grounded in available evidence, rather than funneling students toward a common answer.

3) engage learners' prior knowledge while pushing them towards new ways of thinking about the historical reasoning at hand.

Designing a generative task alone, however, did not count as eliciting. To bound eliciting for this study, I demarcated eliciting as a teacher move done in the moment-to-moment interactions of teaching, as opposed to design work completed prior to teaching. The design of the task is a separate, but related, teaching practice that may include the design of essential questions, historical texts, graphic organizers, and any other classroom materials that support the establishment of a generative problem space.

Eliciting in this study began with the launch of a problem space in the moment-to-moment interactions of classroom teaching. In addition to initial eliciting (i.e., launching a problem space), eliciting also included attempts that the teacher made to initiate new lines of reasoning throughout a lesson. Such a distinction raises the difficulty of disentangling EIR. While I recognize that a teacher who responds to a student idea may continue to elicit student thinking, for analytical purposes, I counted eliciting only as teacher-initiated attempts to bring student thinking into the public space.

With the boundaries of generativity and eliciting clarified, I then identified occasions of eliciting in each set of classroom observation notes and asked, "What kind of historical reasoning does the elicitation position the students to do?" I coded each occasion of eliciting according to deductive coding categories adapted from Wineburg (1991a) and Nokes et al. (2007). These coding categories helped me to identify incidences of eliciting and name them according to the type of historical reasoning that the teacher elicited (Wineburg, 1991a). I did not assume that

types of thinking that I observed captured the extent of thinking that TCs elicited during the duration of their internships. Rather, I approached the thinking elicited during the observations as a window into the TCs' ability to elicit historical thinking that could inform the broader analysis.

A. Eliciting Sourcing	C. Eliciting Contextualization
A1. AU Position Eliciting A2. AU Motivation Eliciting A3. AU Participation Eliciting A4. AU Evaluation Eliciting A5. Date Production Eliciting A6. Document Type Eliciting A7. Evaluation of Document Eliciting A8. Other Sourcing Eliciting	C1. Time or Location Awareness Eliciting C2. Culture or setting awareness Eliciting C3. Biographic Awareness Eliciting C4. Historiographic Awareness Eliciting C5. Linguistic Awareness Eliciting C6. Analogy Eliciting C7. Contextualization Other Eliciting
B. Eliciting Corroboration	**D. Eliciting Justification (use of documents as evidence)**
B1. Direct Comparison Eliciting B2. Direct Contrast Eliciting B3. Claim Uniqueness Eliciting B4. Claim Omission Eliciting B5. Corroboration Other Eliciting	D1. Direct Quote Eliciting D2. General Citation Eliciting D3. Specific Reference Eliciting D4. Use of Doc Other Eliciting

Table 3.3: Eliciting historical reasoning codes

As displayed in Table 3.3, I added a fourth coding category (code D) adapted from Nokes, Dole, and Hacker (2006) called *using documents as evidence*, a category I renamed *justification.* I did not view justification as equal to the categories adapted from Wineburg (1991a) because the heuristic is not necessarily historical reasoning. That is, a student might quote or paraphrase from a document without ever recognizing that document as a historical source. However, I retained the code because I thought it might help characterize a step in the development of eliciting historical reasoning. Although I coded all of the data according to the categories depicted in Table 3.3, analysis demonstrated that the major codes, rather than the sub-codes offered significant insights into the phenomenon of interest. Perhaps with a larger body of data, the sub-codes could point to significant themes.

In time, I developed a protocol that supported the process of eliciting analysis for classroom observation (Appendix D). The protocol led me through a descriptive analysis of a single lesson and the incidences of eliciting in that lesson with the following line of questions:

- Is the problem space (task) generative?

- Do initiations during the lesson invite the articulation of student thinking about evidence in history?

- What kind of historical reasoning does it elicit?

- Do additional elicitations maintain the generative problem space?

Using my answers to these questions, I wrote a descriptive analysis memo that was organized according to empirical observations related to eliciting. Underneath each proposition, I organized data that supported the proposition and data that conflicted with the proposition. With such analytical passes, I eliminated some propositions and revised or added details to others.

Triangulating eliciting propositions. I then turned to the respective post-observation debrief as a secondary source of data for eliciting. I read through the debrief transcript and cataloged any data that supported or conflicted with my preliminary findings from the observation analysis. I asked, "What evidence supports my propositions and what evidence contradicts my propositions?" When I encountered evidence in the debrief that was counter to the findings from the observation, I looked at relevant TC coursework and returned to the observation to reexamine my initial impressions.

After testing each preliminary finding against the debrief data and (when necessary) coursework data, I wrote a memo that included revised propositions and accompanying explanations. For each new observation, I repeated this analytic process and eventually began writing memos on each TC that hypothesized changes over the course of the study.

67

One problem I faced in the longitudinal analysis was that TCs did not engage in classroom teaching until several months into the study. Given the reality that observations could not serve as a primary source of data for this early period, I analyzed pretests and relevant coursework to determine the ability of the candidate to elicit and considered these in light of the earliest observation. To evaluate change over time, I started with later observations and worked backward, noting any observed changes.

Interpreting. Interpreting, in this study, consisted of the meanings that TCs made of instructional interactions and articulated in the post-observation debrief interviews. I made this decision because of the difficulties in fully understanding teachers' interpretations. Because I found no suitable deductive analysis plan in the literature, I approached analysis of TC interpreting through an inductive approach (Goetz & LeCompt, 1984). I created a side-by-side table using my classroom observation notes and the observation debriefs for each respective observation. I organized these exchange tables according to units of teacher-student exchange related to evidence in history with 1) *exchanges between student and teacher* on one side and 2) *interview content about those exchanges* on the other. Thus, I was able to examine and descriptively name each unit of student-teacher exchange and interview data related to the exchange by asking myself, "What led the candidate to respond the way that he/she did?"

For example, one candidate had multiple units coded as "classroom management distraction," indicating that she became distracted from student thinking because of a disruption. After initial naming, I aggregated units into groups based on description and attempted to list how units were similar and different from one another. By this process, I arrived at categories for coding the exchange tables of each TC. Finally, I memoed proposed findings and evidence for these findings for each round of observation/debrief.

Triangulating interpreting propositions. With preliminary findings established for the observations via the process described above, I compared these finding to the ways that each TC interpreted student thinking in the coursework. I asked, "Do the patterns I observed in the observation/debriefs hold when I consider examples of interpretation in the coursework? Is there contradictory data?" When I found contradicting evidence, I went back to the exchange tables in an attempt to resolve the conflict. In evaluating change over time, I looked across the findings and wrote descriptive memos, again working backward through the data.

Responding. This study defines "responsiveness" as "the extent to which teachers 'take up' students' thinking and focus on student ideas in their moment-to-moment interactions" (Pierson, 2008, p. 25). I conducted analysis of responding in a very similar manner to the process I used for eliciting. I began by identifying occasions in which the teacher responded to student thinking about evidence during the classroom observation. For each occasion of responding, I coded it according to deductive coding categories adapted from Pierson et al. (2009). I drew on history education literature to adapt a coding protocol based on the Pierson (2009) to support the responding analysis (Appendix E). This protocol supported a descriptive analysis of the incidences of responding in a single lesson through the following questions:

- Is follow-up responsive to student comment?
- Whose idea is the focus?
- Whose reasoning is on display?

I coded the observation notes according an expanded version of the following categories:

Code	Description
Low/No Responsiveness	Follow-up that is not responsive to student's idea. Moves include evaluating, rebroadcasting, acknowledging, or making a related statement or question.
Medium Responsiveness	Follow-up that is minimally responsive to student; has the form but does not function as responsive. Focus on T's thinking often in a recitation style interaction (S provides basic information T incorporates in her response). Also includes corrective moves and co-opting S response (or peripheral part of it) to make desired point.
High II Responsiveness	Follow-up explores student thinking and allows their reasoning to be the focal point. Uptake in the true sense of the word – responding to and building on a student's idea so that his/her thinking is on display. Includes invitations for students to make sense of one another's ideas; probing S thinking; expanding, clarifying, or giving an example based on S idea; or T asks clarifying question to establish a joint focus of attention.
High I Responsiveness	Follow-up that is responsive to S idea, question, or perceived misconception. The teacher's thinking is on display, but in response to the student's idea. This includes answering student questions and responding to student misunderstandings. T can expand on S comment but takes over S's idea and puts his/her thinking as focus.

Table 3.4: Responsiveness Coding Categories

As with eliciting, I constructed a descriptive analysis memo organized according to propositions drawn from analysis of the observation notes. Under each proposition, I organized supporting and conflicting data from the observation notes.

Triangulating responding propositions. I returned to the post-observation debrief as a secondary source of data for responding. I cataloged any data from the debrief that supported or conflicted with my descriptive propositions. Before writing a final summary description, I memoed revised propositions and explained my rationale for the revisions. I repeated this analytic process for each observation and, over the course of several observations, I began to note changes that I believed I was observing over time.

Again, I faced the problem that TCs did not engage in classroom teaching until several months into the study. Because observations of responding could not serve as a primary source

of data for the early months, I analyzed pretests and relevant coursework to suggest TC responsiveness and considered these data in light of the earliest observation. In developing hypothetical findings about change over time, I started with later observations and worked backward, noting observed changes.

   Cross case analysis. To go beyond the findings of a single case, I looked across the individual cases of my study asking, "how did the teacher candidates elicit, interpret, and respond to student thinking about historical evidence across the period of study?" Like the analysis of the individual TCs, informal cross-case data analysis began as soon as data collection began. My first goal, however, was to understand the patterns and themes of each embedded case before the formal cross-case analysis (Patton, 2002).

   I began by reviewing analytic memos that I wrote throughout the data collection and analysis period on hypothetical differences between candidates' ability to EIR. As I developed ideas of patterns across the TCs, I wrote analytic memos that cited points of data for the propositions. I created checklists that marked similarities and differences between each candidate's coded observations before writing additional analytic memos outlining patterns (Miles & Huberman, 1994). Finally, I reread analytic memos written during the cross-case analyses and I noted patterns of findings and identified supporting and conflicting evidence for these patterns. Throughout this process of constant comparison (Bogden & Biklan, 2007), I met regularly with colleagues to present the evidence for supposed patterns across TCs and to receive feedback.

Issues of Validity and Generalization

   Case study research has long been disparaged because of critiques about the credibility of findings and extent to which cases can be generalized. Where the first issue is a question of how

one can know whether a study's conclusions are valid, the second is a question of whether the findings mean anything beyond an isolated point in time (Maxwell, 2013; Yin, 2003). I explain how I addressed both of these issues below.

Validity. Validity in this study requires only the possibility of testing this account against the world "giving the phenomena that we are trying to understand the chance to prove us wrong" (Maxwell, 2013, p. 123). Because it is impossible to eliminate threats to validity, I sought to acknowledge, document, and capitalize on the relationships between researcher, settings, and participants by frequent memos accounting for the research process. I found that the most straightforward way to deal with validity threats was simply a commitment to integrity. While I could not rid the study of my influence, I did face the data with honesty throughout the course of this dissertation.

I used a number of practical strategies explicitly or implicitly noted throughout this report, which helped me test the validity of my conclusions and expose evidence that challenged those conclusions. The first strategy I used was to collect rich data over an extended period of time. I conducted data collection in order to assess the TCs' capacity in EIR at many points across a nine-month period, well beyond the end of the methods course.

The extensive body of data allowed me to confront and resolve evidence contradictory to my early assumptions. As a natural consequence of the length of this study, I had the opportunity to observe TCs across multiple settings and activities, another strategy that can limit potential for reactivity.

As I explained in the analysis section, I used triangulation to reduce the risk of bias in making chance linkages or failing to notice inconsistencies (Fielding & Fielding, 1986). Analysis of each category sub-practice (eliciting, interpreting, and responding) included the convergence

and corroboration of several sources of primary and secondary data (Yin, 2006). In conjunction with the second strategy, I will make the entire corpus of data available upon request, not only those data that I chose to highlight in this research report. Finally, I employed the support of colleagues and research mentors at pivotal points during the research including the coding of TC data and cross-case explorations (Merriam, 1998).

Generalization. I do not expect to generalize the findings of this study to populations of TCs outside of the context of the cases under study or to argue that all TCs will bear similarities to those in this study. Rather, I hope this study will result in findings that can shed light on the processes under analysis (Maxwell, 2013). That is, the development of EIR (or lack of development) should suggest processes that may operate in similar cases. This study can be generalized, therefore, in as much as its findings highlight processes that might be the same wherever they occur, processes that include variations in settings that result in variations in the findings (Becker, 1990).

Chapter 4: The Social Studies Methods Course

Before I turn to the findings presented in Chapter 5 of this study, I describe the core practice approach used in the teacher candidates' (TCs') methods course. This course was guided by commitments grounded in education literature and articulated in detail in the first two chapters of this dissertation. Namely, Meredith and I were committed to history as an interpretive, evidence-based method of knowing about the past and we were committed to preparing TCs to support students in reasoning, reading, and writing to that end.

Key to such teaching and learning, we believed, was an ability to elicit, interpret, and respond to student thinking about many different aspects of thinking in history. A final commitment was grounded in beliefs about professional learning as articulated in a core practice approach. That is, TCs could learn to teach in the ways that we envisioned if they engaged in the cycles of learning that we designed. In this chapter, I provide an overview of the course and its implementation.

Planning the Course

Can we figure out a way to make argumentative discourse a thread from reading, to discussion, to writing? - Meredith

Meredith and I have been colleagues for years and have worked on a number of projects together. I first approached her about core practice course redesign more than a year before this dissertation study began. We agreed that I would take on the Summer 2014 Social Studies Methods I course to get some experience with secondary methods and then we would do the redesigned course in the fall of 2014.

74

Meredith and I met dozens of times over a six-month period preceding the first class. During this time, we developed a conceptual framework for the course, conducted literature reviews of various teaching practices, and selected a cycle of teacher education pedagogy to guide our design of class activities and assignments. We then drew on this foundational work to redesign the course (as it previously existed[7]), the syllabus, and the assignments around a set of teaching practices and a cycle for investigating and enacting those practices.

While we knew that we wanted to organize the course around practices of teaching, we were initially intimidated by hundreds of potential teaching "practices." The original methods course and social studies program were organized around broad outcomes but we felt we needed to first clarify how we wanted our TCs' students to learn in order to determine the practices that we wanted our TCs' to enact.

We sat in front of a white board one day and brainstormed what students would be doing in the ideal history classroom. We envisioned a history classroom steeped in inquiry and dedicated to building students' ability to 'do history,' or read, write, and discuss by way of historical reasoning. At one point, Meredith wrote on the board, "Can we figure out a way to make argumentative discourse a thread from teaching reading, to discussion, to writing?" To facilitate such a classroom, our TCs would have to recognize teaching as a relational exchange grounded in, and directed by a teacher's understanding of student thinking.

In this way, EIR became an over-arching practice that included a number of practices selected based on the following criteria that I adapted to history education (Windschetl et al., 2012):

---

[7] The course as it previously existed was originally designed by Chauncey Monte-Sano but had been adapted and taught by Meredith for several years. Previous year's course materials included some draft forms of materials that later appeared in Monte-Sano, De La Paz, & Felton (2014).

1. Relevant: The practices must be applicable to the everyday work of teaching, support the learning of all students (e.g., supporting student development in reasoning historically), and be suitable for any secondary level for any historical inquiry.

2. Preparation oriented: The practices must be accessible to novices, teachable, rehearse-able, and enact-able.

3. Coherent whole: The practices must build upon one another and fit into a coherent whole (i.e., the set of practices form a pedagogical core for a unit of historical instruction).

4. Rigorous selection: Rather than breaking practices into atomized units, we sought practices that include the coordination of many micro-practices or "moves."

What teaching practices did we select?

Although we considered many teaching practices, we chose three practices as the infrastructure of the course that were in accordance with the criteria above and required EIR. These practices were 1) designing and using assessments to guide instruction, 2) leading a text-based discussion, and 3) scaffolding argumentative historical writing. While each of the practices above are teaching practices in their own right, all three require an ability to elicit student thinking, interpret student thinking, and respond to that thinking in ways that facilitate targeted student learning.

Figure 4.1: Eliciting, Interpreting, and Responding Framework

How did we articulate the practices?

Once we determined to design the course around these three practices that required EIR, we began the hard work of articulating the boundaries and definitions of each practice. I supported this work by conducting a literature review on each practice in history education research literature and expanded this review where history education literature had little to offer.

For each practice, I reviewed the research literature and kept notes about 1) effectiveness in the practice, 2) challenges or problems related to the practice, and 3) other relevant information about the practice. Prior to meeting with the course instructor for planning meetings, I consolidated these notes into bulleted outlines that served as a "CliffsNotes" for our course planning.

To more clearly articulate the practice for ourselves, we took the additional step of using the literature reviews to write case descriptions of each practice in order to "bring the practice to life." These case descriptions served as a living example of what we found in the literature and forced us to transcend the principles of a practice and actually envision the practice as enacted in a classroom context. We also sent these cases out for peer review to several experts in teacher education and social studies instruction.

Scaffolding the Practices

"Sorry, but we are going to have to find something that the mentor teachers will actually appreciate." – Meredith

While the course was broadly organized around the three practices noted above and EIR as an overarching practice, what made the practices "studyable" (Ghousseni & Sleep, 2011) was use of instructional activities (IAs) and cycles of enactment and investigation. Rather than break up every practice into atomized parts, we maintained the whole of the practice but chose IAs that

provided structured opportunities to study and engage in the practice. In this way, the IAs acted as "containers" for the teaching practice, embodying the key practices, principles, and knowledge required to engage in the practice.

Because our TCs worked across various content areas of social studies in secondary settings, we designed IAs with more flexibility than some of the set activities described in the literature (e.g., Lampert et al., 2010). Both the flexibility and the structure of the IAs we used were points of contention between the instructor, the TCs, and me at various times during the course and its planning. The typical source of contention between Meredith and me was how to adapt a core practice approach to the realities of placement classrooms; I often pushed IAs that would best embody the target practice without recognizing the challenges that TCs would face in enacting the IAs. The primary source of contention with the TCs was related to the fact that several were placed in government classrooms but the course was focused primarily on methods of teaching history. Apart from the obvious disciplinary distinction, additional tensions arose because government was the social studies course that included state-mandated testing.

With IAs, we could design TC learning experiences through an iterative process of representing a practice, deconstructing a practice, approximating a practice for enactment, and reflecting on the enactment (as depicted in Figure 2.1). One of our goals in designing and using IAs was to bound the complexity of EIR within relatively predictable classroom mechanics and, thus, allow for relatively safe approximations of practice and eventual classroom enactments. What remained unpredictable, of course, was the student thinking that emerged when the TC engaged in the IA with actual students.

Thinking up IAs that could make practice studyable was not that difficult. In fact, I had dozens of ideas when Meredith and I began designing IAs. But, when I proposed having each TC

interview several students during class about a primary source document, Meredith said, "Sorry, but we are going to have to find something that the mentor teachers will actually appreciate." Because Meredith knew the mentor teachers and was familiar with the norms of their classrooms, she expected that a "pullout" of a few students would be disruptive for the normal flow of class and, thus, be met with concern from the mentors and the TCs. The challenge of this design work was identifying an IA that could productively be implemented in secondary classrooms in a way that the mentor teachers would see as adding value to the students' educational experience, not just the TCs'. Below I explain the IAs we eventually designed for each practice.

Designing and using assessment to guide instruction IAs. We adapted three instructional activities to serve as the containers for the practice of designing and using assessment to guide instruction: monitoring, one-minute essay, and misconception check. Each of the IAs allowed the TC to engage in a specific means of gathering student understanding data but the IAs left the student task somewhat open for TC adaptation. For example, monitoring is a strategic 'walkabout' in which the teacher compares predicted student responses to actual student responses. As the model teacher walked the classroom, she made decisions regarding when to engage particular students or groups of students in further dialogue, based on the comparisons of student responses to the responses she envisioned.

The one-minute essay is designed to capture a snapshot of student understanding by having students construct a brief response to a targeted question. The design of a question that elicits meaningful responses from students in a brief time span requires rigorous narrowing and skilled prediction (Wiggins & McTighe, 2005). Like one-minute essay, a misconception check can be an efficient means to elicit student thinking and provide a broad record of that thinking

for interpretation and instructional response. The teacher poses a misconception and asks the students to adapt or correct the misconception (Wiggins & McTighe, 2005). While these three IAs have structured aspects, each can be adapted to most contexts, content, and resource capacity. The commonality was that each IA required the TC to design a question or task that elicited student understanding, understandings that the TCs articulated ahead of time and targeted in the IA.

Leading a text-based discussion IA. We also adapted three instructional activities to serve as the containers for the practice of leading a text-based discussion: inquiry discussion, Socratic seminar, and structured academic controversy. Each of these IAs outlined the basic moves for the teacher and left the student task somewhat open for adaptation across contexts.

An inquiry discussion includes an evolving hypothesis, in which students are presented with a basic inquiry question (e.g., Why were Japanese Americans Interned during WWII?), asked to form an initial hypothesis, and then presented with multiple rounds of evidence. With each new round of evidence, students return to the inquiry question to either adjust or defend their evidence-based answer (e.g., Stanford History Education Group: http://sheg.stanford.edu/japanese-internment).

In a structured academic controversy, students work first in teams of two to build a particular argument (e.g., Lincoln was a racist) based on primary sources. Next, the pair explains one side of the argument to another pair of students who were assigned the counterargument (e.g., Lincoln was not a racist). Both groups restate the argument they heard before they are allowed to contest the other side's position. The group of four then drops their assigned arguments and works together to come to a consensus prior to a whole class discussion (e.g., Stanford History Education Group: http://sheg.stanford.edu/lincoln).

With Socratic seminar, students first prepare by reading a common text, answering questions, or engaging in writing. In concentric circles, the inside circle discusses a particular question and the teacher interjects as needed. The outside circle is given the task of analyzing the arguments made by weighing the strength of evidence. At some point, the outside circle rotates to the inside and the discussion continues. Although each of these IAs is different, each offers an opportunity for TCs to engage in leading a text-based discussion by way of eliciting, interpreting, and responding to student thinking (e.g., Metzger, 1998).

Scaffolding argumentative writing IA. We designed a single instructional activity to serve as the container for the practice of scaffolding argumentative historical writing. Because TCs had so little experience with teaching and learning writing, we limited the instructional activity to a flexible model of instruction called cognitive apprenticeship. We used a number of the resources available in Monte-Sano, De La Paz, and Felton's (2014a) practitioner oriented book for this and other IAs.

Cognitive apprenticeship can be used to teach expert thinking or skills to students in a gradual, scaffolded approach. In cognitive apprenticeship, the teacher models the target thinking/skill by making the expert thinking visible, pointing to tools that support reasoning in this new way, and identifying specific strategies used to engage in the target thinking/skill. As students develop a foundational understanding, the teacher continues to support their practice as she moves them toward increasing independence in more challenging forms of the practice (Monte-Sano, De La Paz, & Felton, 2014a). We limited the IA to the first part of a cognitive apprenticeship (i.e., modeling and making strategies visible) because most of the TCs described their students as having limited writing proficiencies and expected to be starting at a foundational point with writing instruction.

The Instructional Cycle

We designed instruction in the IAs based on the McDonald et al. (2013) cycles of enactment and investigation (Figure 4.2), as much as practical realities would allow. The cycle began with collective analysis of an instructional activity model. TCs worked with the teacher educator to deconstruct the model and connect teacher practices to principles of teaching and learning.

Figure 4.2: Cycle for Investigation and Enactment (McDonald et al., 2013, p. 5)

Next, TCs used the analysis to prepare their own enactment of the same instructional activity. They then engaged in a guided practice or rehearsal of the planned enactment and received feedback from peers and/or teacher educators. Next, they enacted the instructional activity in a classroom with real students, watched a video of their enactment, and reflected and/or engage in collective analysis with peers and teacher educator.

Cycle 1: Designing and Using Assessment

"I am seeing teaching that is way up here and student performance way down here. But, then the teacher is just like 'oh well, bad day I guess.'" – TC enrolled in course

Model and deconstruct. The cycle of investigation and enactment for assessment began with analysis of written cases of the assessment IA of monitoring. TCs examined an exemplar and non-exemplar and discussed distinctions between the two. TCs then worked with the instructor to deconstruct the cases using principles of effective assessment from the readings. Together the class constructed a graphic organizer that linked principles of effective assessment with specific moves from the cases. The same principles were then used in subsequent weeks to deconstruct the one-minute essay IA and misconception check IA.

Prepare. Because the assessment cycle fell early in the semester, some TCs were limited in the type of classroom enactment they were able to do. Each TC worked with their mentor to either revise a mentor's lesson or design a new lesson or portion of a lesson that allowed for enactment of one of the assessment instructional activities. That is, TCs worked with mentors to design a lesson portion that included monitoring, one-minute essay, or misconception check as an assessment of student reasoning.

TCs first analyzed the assessments in the original lesson plan and considered how the assessments might be modified or whether an additional assessment(s) could be added to make student thinking more explicit. In preparing the lesson, the TC ensured that at least one of the assessments was a formative assessment from the task pool and that each assessment aligned with principles of effective assessment highlighted in the course.

Enact and reflect. TCs engaged in several activities and rehearsals of assessment in the during methods course meetings. These activities included analysis of student reasoning and

evaluating assessments designed to elicit student reasoning. Building on these experiences and the preparation described above, each TC implemented one of the assessment IAs with at least one class of students. TCs videotaped their enactments and collected evidence of student understanding from the assessments, when assessments included written artifacts.

TCs investigated their practice by individually watching the video of the class with specific attention to student understandings that surfaced during the lesson portion and the student work generated from those assessments. Finally, TCs wrote a paper using examples from class to justify whether the assessment effectively made student reasoning explicit.

What we learned. By the time we began the assessment cycle, many of the TCs were beginning to grasp the relevance of eliciting student thinking in an instructional design that began with student outcomes and worked backward through assessment and task design. As TCs deconstructed models and prepared for classroom enactments, many expressed frustration with not being able to evaluate one assessment IA as the 'best.' For example, several TCs noted that their mentors always used exit tickets to assess students, and that seemed to be sufficient for every lesson. While some TCs recognized that the value of an assessment depended on the student thinking the assessment was designed to elicit, many resisted such nuances.

During the assessment cycle, fissures in the cohort burst wide open. Tensions had been present for some time and were partly attributable to differences in the placement contexts of the TCs. Several felt that instruction in the methods course was only relevant to TCs placed in classrooms with certain norms and student expectations.

At first, frustration manifested as grumbling and eye rolling from some of the TCs when we showed the model IA. When given the chance to speak, three TCs insisted that the IA modeled in class was incongruent with the context of their school and mentors' expectations.

They criticized the amount of time required, the complexity of student thinking the IA targeted, and potential classroom management problems.

One of these TCs explained that the teaching practices she was seeing in this course were designed "for some kind of dream world of magic" and not for her placement where the kids are like "the Children of the Corn." Tom, a TC sitting to her left grumbled, "I am seeing teaching that is way up here and student performance way down here. But, then the teacher is just like 'Oh well, bad day I guess'." On the day we planned to wrap up assessment and begin a new practice, Tom was particularly angry. He explained that he had enacted the assessment IA that day in class, and though it was relatively successful, it was totally incongruent with what students did on a regular basis.

Tom insisted that because of the realities of their classroom placements, none of the TCs were actually enacting the IAs in authentic ways but were instead "just going through the motions" in order to show Meredith what she wanted to see. When the instructor tried to respond, Tom interrupted and shouted her down. The scene was so emotional and the topic so important to the framing of the course that we dedicated the rest of class to talking about the issues that Tom raised and ensured we heard from every TC.

The primary pushback was that the TCs in government courses were seeing teaching and assessment targeted toward the state government exam, rather than deep student understanding. Tom, in particular, observed a tremendous amount of assessment that he characterized as not meaningful for supporting student remediation. Because these unhappy TCs saw their placement as "the real world" and the methods course as an ideal, they initially dismissed the assessment IAs as jumping through hoops.

Apart from the three TCs placed in government classrooms, few other TCs agreed that the IAs or other practices taught in the methods course could not be enacted. Instead, most expressed that the mentor was not necessarily teaching in these ways but encouraged them to try any teaching practices the intern wanted to try (including one government TC). While everyone expressed understanding for Tom's frustration, Sally was the only TC in a history placement who relayed similar concerns about the problematic context of her placement.

The explosion described above was deeply unsettling for Meredith, and even more so for me. I worried that she would abandon the focus of EIR and core practice and back away from the redesign. Meredith and I met numerous times in the week following the incident to decide what was next. She also met individually with the TCs who expressed the most resistance in class in order to more fully understand their perspective. Instead of backing off, we doubled down on the practices approach but worked to make it clear to the TCs that the IAs were, in fact, implementable and valuable.

One of the major changes moving forward was to use preservice teachers for the model IA, whenever possible. In making this tradeoff, we lost the perfection of an ideal model but we sought to address a dilemma we had not expected – some TCs' refusal to believe that preservice teachers could actually engage in these practices in real classroom contexts. The video model we selected for our next practice, in fact, was a preservice teacher leading a discussion in a government classroom at one of the unhappy TC's school.

Despite the contention, all TCs engaged in the classroom enactment of an assessment IA and exhibited increasingly cogent understandings of how to use assessment as an imbedded classroom experience to elicit and interpret student reasoning. As suggested by the discontents, assessment as an embedded and formative task was out-of-line with many of the mentor

teachers' classroom practices. For example, Craig explained that his mentor used the district

formative assessments but simply handed it back to the students with little or no evaluation or

guidance. Very few TCs reported observing models of teachers who strategically elicited and

responded to evidence of student learning. Most of the TCs reported that in their classrooms,

when students did not learn, mentor teachers simply moved on.

Cycle 2: Leading a text-based discussion IA

"Wait, so like what you are doing right now is actually not a discussion. It's really an interactive

lecture, right?" - TC enrolled in course

      Model. We launched the cycle for leading a text-based discussion by showing video

segments of two TCs (from a prior cohort) leading a discussion. The good (but imperfect) model

was a TC who was leading students in an inquiry discussion about why the U.S. invaded Iraq.

The non-exemplar was another TC who was leading a dialogue that required students to "take a

stand" in response to the question "If government instructs you to kill someone then is it

murder?" While both video segments showed high student participation, only the exemplar

demonstrated a group of students engaging with each other's thinking about the substance of

social studies content.

      In subsequent weeks we modeled a structured academic controversy and Socratic

seminar. Each of these three instructional activity models exemplified the foundational principles

of discussion leading that we wanted to teach. Namely, that discussion is an exchange of student

ideas around particular content/text and not simply an interactive lecture or exchange of

opinions.

      Deconstruct. TCs worked with the instructor to deconstruct the models by using a T chart

with the columns "What is the teacher doing?" and "What are the students doing?" This graphic

organizer helped the TCs notice connections between teacher moves and student moves in structured discourse. As students recognized teacher moves, we provided common language with Reisman's (2010) discussion leading moves.

Each week of the cycle, TCs repeated an abbreviated version of this process with a new model. With each deconstruction, the teacher educator highlighted a new discussion-leading move that the model teacher used to elicit and promote response to student thinking (e.g., asking for textual support, prompting counterarguments, uptake, stabilizing the context).

Prepare. Each week, as TCs were introduced to new discussion leading moves, they prepared and rehearsed these moves in a number of ways. One week TCs worked in small groups taking turns practicing the moves as other TCs acted as students who were discussing a primary source. Another week, TCs tried moves in front of the class with a teacher educator playing the role of student. Over the course of the cycle, TCs prepared a lesson segment to enact using one of the three discussion IAs.

In preparing the enactment, TCs anticipated student thinking and planned targeted questions and teacher moves to elicit student thinking. In conjunction with their mentors, TCs also prepared lesson materials and adapted texts for specific classroom needs. Prior to classroom enactment, each TC taught a portion of their discussion lesson to a small group of other TCs who were playing specific (assigned) student roles. Targeted feedback provided during these sessions proved beneficial for classroom enactment.

Enact and reflect. TCs taught the discussion lesson/segment with at least one class of students, videotaped their enactments, and collected evidence of student understanding, when available. Each week of rehearsal included opportunities for TCs to receive feedback from peers and teacher educators on targeted discussion leading moves. After the classroom enactment, TCs

further investigated their practice by watching their enactment video with attention to exchanges that effectively elicited and built upon student thinking. TCs wrote a reflection paper in which they explained whether the instructional materials, plans, and moves were effective in supporting student engagement with each other's thinking about the content.

What we learned. When we started the discussion cycle, most of the TCs admittedly viewed discussion as a type of interactive lecture or exchange of student opinions, successful if more than a few students participated. Deconstruction of models and non-exemplar models were helpful in demonstrating a distinction between engaging conversations and exchanges of student ideas about historical texts.

As early as the first week of the discussion cycle, we noticed that TCs were able to recognize the 'moves' and describe these moves when prompted. For example, when we modeled and analyze several discussion moves in one of these class sessions, TCs had little difficulty pointing out the teacher's actions and explaining how these actions promoted student discourse. When asked to actually engage in the moves in a rehearsal, however, many TCs floundered.

After one rehearsal I observed, Tom gave Gabby structured feedback after she finished leading a mini-discussion with her group. Tom rightly suggested that Gabby did not use many (if any) of the discussion-leading moves the TCs had highlighted during the deconstruction. I was surprised to hear her argue with Tom that she had revoiced student thinking and asked for textual support because she had not actually engaged in those moves. As this example illustrates, it was not surprising for TCs to understand and articulate aspects of teaching that make for a good discussion (as Gabby had) but fail to implement the moves (even unknowingly) when given the opportunity.

During the discussion leading cycle, we also noticed that many of our TCs seemed to compartmentalize their knowledge of teaching and learning. For example, TCs could explain historical reasoning that they intended to target. But, when given spontaneous opportunities to elicit student reasoning by way of discussion, TCs tended to focus on discussion moves that maintained student engagement and ignored opportunities to elicit student historical reasoning.

During one practice exercise, I observed many candidates ask for textual support and prompt counterarguments but did not see TCs incorporate prior comments into subsequent questions (uptake), a move that requires substantial attention to student thinking. By the second week, Francis cleverly confronted me as I asked students questions about leading a discussion. Between my questions, she asked, "Wait, so like what you are doing right now is actually not a discussion. It is really an interactive lecture, right?" Although slightly embarrassed, I was pleased that Francis was grasping the distinction.

When candidates engaged in rehearsals of their planned discussions, many TCs attempted to make student historical reasoning (of classmates) public and build on this reasoning. For example, Kendra led an inquiry discussion based on the question "Who shot first at Lexington Green?" The discussion rehearsal she enacted included many attempts at exploring student historical reasoning skills and discussion moves. At times awkwardly and at times skillfully, Kendra engaged in these overlapping eliciting tasks. More often than not, she ended up in one-on-one student exchanges and struggled to bounce the discussion back to the whole group. However, the way Kendra attempted to elicit historical reasoning and respond to that reasoning was emblematic of the progress we saw across most TCs during the discussion cycle.

Cycle 3: Scaffolding Argumentative Writing IA

"Our students can't write worth shit and my mentor does not teach writing." - TC enrolled in course

Model. We launched the cycle of investigation and enactment for scaffolding argumentative writing not with a model of live instruction, but with a series of model lesson plans from a collection of resources built on a cognitive apprenticeship model (Monte-Sano, De La Paz, & Felton, 2014b). For each of the three lesson plans, the authors provided historical background information, lesson analysis and student writing samples drawn from classrooms in which the lessons were implemented.

Although the Monte-Sano et al. (2014b) text contained valuable instructions for engaging in cognitive apprenticeship, we had difficulty identifying actual cases or video models of cognitive apprenticeship for writing instruction, much less writing instruction specific to history. We settled on a video of an English teacher engaging in and describing gradual release of student writing.

Deconstruct. TCs engaged in a critical examination of Monte-Sano et al.'s (2014a) lesson plan models and the video model. TCs attended to the ways that the model lesson plans demonstrated the gradual release of specific scaffolding tools for general writing skills and historical writing skills. Working together, they created a map depicting the gradual release of responsibility for writing over the three successive model lessons.

Prepare. Using the map as a guide, TCs planned a process for scaffolding argumentative historical writing for their students. This first required an evaluation of students' general and disciplinary writing abilities. TCs then adapted the cognitive apprenticeship process and scaffolds depicted in the models to meet specified student writing needs.

TCs rehearsed and videotaped their attempt at the modeling portion of a cognitive apprenticeship without any students. In this safe but inauthentic approximation of practice, they demonstrated and talked through a general or disciplinary writing practice. The TCs planned this rehearsal in accordance with student ability levels and demonstrated the next stage of writing/thinking that would be required for historical argumentative essay.

Here the TCs practiced delivering a cogent, expert model of the disciplinary and general writing skill they would later teach. While the rehearsal was inauthentic and lacked interaction, it gave TCs a chance to articulate expert practice of a specific skill in a low stakes environment prior to trying it with students. Based on analysis of the rehearsal videos, candidates planned and taught a multi-day writing unit in the spring semester based on a cognitive apprenticeship model of instruction.

Enact and Reflect. Candidates enacted a cognitive apprenticeship instructional approach during the implementation of a writing unit. Depending on student proficiency in the targeted writing skill, TCs engaged in various degrees of demonstration with gradual release of support. Due to constraints in most of the placements, we had to bump the actual enactments into the second semester. Thus, reflection in the course was limited to reflection on the process of planning and rehearsing a cognitive apprenticeship model. All writing lessons were enacted in the spring and included extensive reflection and feedback.

What we learned. As class was ending the week before we began the writing cycle, Meredith told the TCs where we were heading next. One particularly frank TC immediately called out from the back of the room, "Ok, but I'll just let you know now that our students can't write worth shit and my mentor does not teach writing." Several other TCs agreed with the sentiment as everyone packed up.

With this deflating reminder ringing in our ears, Meredith and I sought a strategy for the cycle that might leverage the TCs' belief that their students were poor writers. Rather than arguing that they can write, we took the tack that 'Of course they can't write – we have to teach them how.' This rationale resonated with our students and set up the cognitive apprenticeship IA nicely because TCs were able to conceptualize particular disciplinary and general writing abilities on a continuum from *beginner* to *expert*. Breaking out targeted argumentative and disciplinary writing skills helped debunk the notion that kids can't write and reframed writing in social studies as a long-range, strategic process.

Although we got the motivation right, we got the instruction wrong. More than any other practice, Meredith and I were learning about the practice as we taught it. At one point, I wrote in my field notes that "More than any other time this semester, cognitive apprenticeship has been a concept that Meredith and I have really built while flying the plane." Neither of us had much experience using cognitive apprenticeship as teachers and because few models were available, most of our understanding had come from descriptions in the research literature.

After watching the rehearsal videos, we realized that our TCs did not actually understand the purpose of cognitive apprenticeship. Rather than a demonstration of expert performance and unmasking of expert thinking, the rehearsals exhibited descriptions of graphic organizers, explanations of analogies, and clarifications of instructions. Meredith and I recognized that TCs missed the underlying rationale for cognitive apprenticeship that expert thinking is not something a student can learn to do because it is explained to them. Like learning to hit a fastball or do a pirouette, expertise requires apprenticing the beginner into the skill over time.

I determined to start over at the beginning in class the next week and 'go again' with cognitive apprenticeship –hammering a distinction from modeling or gradual release alone. But, the second round was even less successful than the first. That night, I wrote:

When I did the debrief, I did not hear much talk about expert performance. I suddenly had this terrible feeling that I was looking into a mirror. I was having the exact experience that I was trying to help them not have – how do you help students do something they cannot currently do, and will not be able to do if you just explain it? I had just spent the last 20 minutes explaining and showing examples.

With a few exceptions, my analysis of student reflections on the rehearsals was discouraging because I did not see critique that demonstrated transformed understandings of cognitive apprenticeship. The evidence suggested that the TCs did not differentiate between demonstrating a skill and gradually releasing students in an expert practice through strategically designed scaffolding.

Despite perceiving this failure, I still believe that cognitive apprenticeship was the right IA to challenge our TCs preconceptions that teaching writing in history simply meant giving students opportunities to write. It gave us the chance to demonstrate writing instruction as a process to develop increasingly complex skills over time that depends on a teacher's ability to use long-term planning, and targeted modeling and scaffolding.

Building the Plane while Flying

When I embarked on the course redesign described above, I expected a linear process with challenges that could be ironed out before classes began in September. What I actually experienced, however, was both exciting and terrifying. Tensions began almost as soon as

94

Meredith and I started planning as I pushed for fidelity to the literature and she pulled for relevance to the field placements. Together, we negotiated each PowerPoint slide, every bullet point in our lesson plans, and the minutiae of each assignment. When our expectations clashed, as they sometimes did, we either compromised or argued our way to a solution.

The TCs also pushed back in ways I had not anticipated. Because TCs knew that they would eventually enact the practices in their classrooms, they were not generous about approaches judged as unworkable in their classroom contexts. Where I expected that the relevance of a core practice approach would bring teacher education principles to life, the reality sometimes felt more like awakening Frankenstein. That is, the entire experience from redesign to reflection was interwoven with tensions within and among the project's participants, tensions that sometimes felt like a monster setting upon my tidy dissertation project.

Rather than try to control these tensions, I tried to become a good listener and an adaptive teacher educator. The product of this work was a redesigned course and an experience that transformed my understanding of what I am capable of as a university-based teacher educator. The course, however, served only as the opening of this larger dissertation study.

In the following chapter, I present four TCs enrolled in the course described above. With the findings of Chapters 5-8, I describe how Sally, Gabby, Kendra, and Craig attended to student reasoning about evidence, in light of a course intended to promote eliciting, interpreting and responding.

Chapter 5: Sally

In an early autobiography assignment, Sally described herself as a Type A, shy person who preferred to let others lead. During the time that I knew her, she transcended these tendencies to meet the demands of her program and placement. As a married mother of a one year old, Sally learned to accept the challenges of balancing teaching, graduate school, and family life in a good-humored way.

Sally was "local" in the sense that she graduated from a public high school in a neighboring school district, was enrolled in the 5th-year master's version of teacher education in our college, and had friends and relatives who taught in area school systems. Sally had the most teaching experience of anyone in the program because she served as a long-term substitute in a neighboring district during her yearlong maternity leave. As such, she was confident going into her internship about areas that most candidates worry about, such as finding and creating instructional materials, organizing and setting up a classroom, and implementing the curriculum.

Sally reportedly loved history, loved sharing historical knowledge with others, and saw this passion as her greatest strength as a teacher. Despite her history major, Sally's pretest and early coursework did not demonstrate a particularly strong historical epistemology. Instead, she had a general awareness of the "biased" nature of all sources and tended to talk about history as narrative.

Sally's pretest results suggested that she recognized the need for evidence to support assertions but saw evidence as a straightforward use of textual support rather than interpretive analysis. Perhaps this is why she expressed concern in her self-assessment over how to teach disciplinary thinking skills in social studies.

In Sally's early coursework she often referred to topics such as "critical thinking" and "various learning methods," without clarifying what she meant by these terms. When she described social studies classrooms, both envisioned and observed, she stuck to generalities and was impressed with classes in which students seemed interested and engaged. In early coursework, Sally regularly referred to the importance of primary sources in history class, resources that she felt could make topics more interesting than reading from a textbook.

Context

Sally requested an internship at a particular high school because it would be close to home. Because the area was a tourist destination, Sally expected the students to be similar to those she had seen in the attractive areas of town and was surprised to find that many of her students lived in poverty (45% Free/Reduced Meals) and the students in the classes she taught were quite different than those from the high school she attended.

The district high school housed an International Baccalaureate Program and a fine arts program, in which few of the school's 1700 students participated. The student population was made up of roughly equal parts White, African American, and Hispanic students. This student body included a large Spanish-speaking population, 14% of whom qualified as Limited English Proficiency.

A few years before, the school transitioned to an "honors for all" approach, which put high performing students in the same classes with their low-performing peers. During the first few weeks of the internship, I thought Sally was in a state of shock every time I asked her about how things were going. She was astonished by the range of ability levels in the room, particularly the students' reading abilities. Her classes were co-taught and included many students with IEPs and 504s, and several English language learners.

Sally's classes were all 9th grade "honors" American History and several were co-taught with a special education instructor. In addition to the students with documented accommodation plans, many students in Sally's classes were struggling readers. At one point, Sally suggested that the majority of students "seem to lack the ability to think in an abstract manner." She noted that in the same classroom were students who always finished early and were told to work on assignments for other classes with their additional time.

On almost every occasion I visited, I saw a majority of students who silently tolerated disruptive behavior from a few students who demanded most of the teachers' attention. Conversations that I overheard in the staff room suggested that the scenario was not limited to Sally's classes. On multiple occasions, I heard teachers talk about "terrible classes" and kids who "can't do that work." Despite my impression that Sally's internship was not a healthy environment to become a teacher, Sally always spoke positively about her colleagues and most of her students.

Sally liked her mentor, despite the fact that he was very different from her and very different from what she had requested in her internship application. Where she wanted a mentor who would provide a lot of advice and guidance, Sally's mentor gave her independence. Because he was a giant of a man with a powerful voice and imposing presence, his classroom management strategy was simply to be himself. His charismatic personality seemed to be the fuel that propelled his style of teaching that, while not grounded in historical reasoning, kept most students engaged. He valued writing and used numerous graphic organizers to support students' argumentative writing. Although primary sources were not a staple in every lesson, the mentor liked to use them to practice skills needed for standardized document-based questions (DBQs).

These writing assignments typically depended on straightforward reading of the texts and did not require inferential conclusions based on sourcing or contextual information.

From early in the year, Sally expressed concern that her students struggled to read and analyze even the most basic primary documents and make arguments based on this analysis. When Sally took over the class, she initially struggled to maintain students' focus and had particular trouble when the mentor was out of the room. Sally described some of her biggest challenges as "keeping students engaged" in lessons and knowing when to push her students. She was frequently frustrated that students seemed bored and resistant.

Eliciting, Interpreting, and Responding

In the following section, I describe Sally's ability to EIR over the course of the study. The data that informed these findings include information drawn from the pretest Sally took prior to coursework, course participation in the following months, five classroom observations and six interviews. My first observation of Sally's classroom teaching occurred in mid-October, 2014. My final observation was in late March 2015.

Eliciting at the outset. As described in Chapter 3, I analyzed eliciting according to two features: 1) generativity of the task the TC launched and 2) the specific type of historical thinking the TC elicited. At the outset, Sally did not show evidence of an ability to launch generative tasks, despite the utilization of potentially generative documents and task structures. During the tasks, she elicited thinking about justification and only vaguely elicited thinking about historical evidence.

Outset generativity. Sally chose resources and a task structure with generative possibility, but she had some difficulty in launching a task that positioned students to use their understandings to interpret historical evidence. Early in my observation of Sally's discussion

lesson, I heard her encourage the students to "use … the sources of evidence to support your argument" that the New Deal was a success or failure. As I heard this, saw the historical sources, and observed the essential question on the board, I wondered whether I might see the students transformed into historians before my eyes and hear discourse about historical evidence.

As the lesson took shape, however, I recognized that despite frequent encouragement to "use the evidence," Sally launched the task in a way that limited student reasoning significantly. Instead of positioning the students to build arguments based on diverse interpretations of the evidence, Sally's task required a straightforward reading and explanation of each document.

Each of the New Deal documents had its own space on the graphic organizer for "analysis." After putting half the students into a group that would argue that the New Deal was a success, Sally read the instructions on the graphic organizer:

> You should use Documents A-D to look for your evidence. Then, explain how that
> document supports your argument.

Sally did the same for Group 2, except they had different documents to analyze and would argue that the New Deal was not a success. In the post observation debrief, Sally explained what she meant by "analyze the documents:"

> So for each of the documents … they had a space where they will say 'Document A
> says…' and they would state specifically what the document says. And under that there
> was as spot that would say, 'This supports my argument that the New Deal was a success
> or a failure because…' and they would have to explain it. …Because we've had a lot of
> DBQs ... They will say, like, this map said this or this letter says that. But then they won't
> connect it to 'this is important in what I'm writing about because…by splitting it in half

where they had to do both parts of it, I thought that would maybe help them understand

like, okay, this is important to do it this way.

The analysis task that Sally launched in this lesson included the fundamental aspects of an

argument needed to reason about evidence in history. That is, Sally positioned the students to

identify information from the texts, recognize that evidence as affiliated with a claim, and

provide an explanation to connect the evidence with the claim. In this way, Sally positioned her

students to identify textual data to support claims about the New Deal's success.

However, Sally did not position the students to reason about evidence in history because

she treated the texts as sources of information with only one interpretation, rather than historical

sources that must be 'turned into evidence' by interpretation. Sally's version of analysis did not

position the students to reason about evidence in history because the graphic organizer only

required a matching of information to an assigned claim and an explanation for how that

information supported the claim one way or the other.

The New Deal lesson led me to wonder what exactly Sally meant when she launched the

task with the encouragement to "use evidence." In the debrief, I asked Sally how she would

explain evidence to her students:

> So you have an argument or a claim or something and so you can't just say, I'm saying
>
> this and that's all you say, you need to have something that supports what you're saying.
>
> So that will be your evidence. So if you say the New Deal is a success, you have to have
>
> reasons that you think that it's a success.

Sally rightly understood evidence as data that backs up a claim. But, the launch of her task and

her explanations in the debrief suggested that Sally did not see evidence needed in history as

particularly different from evidence needed for general argumentation. Sally's explanation of

evidence suggested that she saw evidence as something that was 'out there,' available to anyone who could read a document.

After both teams had a chance to "analyze" the documents they were assigned, Sally had each side present the documents that they analyzed to the class and explain how each document supported their assigned claim. But, when it came time for the discussion, all anyone could do was present content from the texts and explain how that content connected to the claim that the New Deal was/was not a success.

The launch of Sally's task was also troubled by the way that she adapted the structured academic controversy *Instructional Activity*. Sally did not require students to examine all documents and argue both sides of the essential question prior to a whole class discussion (as in the IA). Sally's adaptation required each team to read only documents assigned to their argument, documents that she told them were aligned with their assigned argument. Instead of reading and interpreting the documents, therefore, students believed that their four documents supported their assigned claim. Sally explained this decision during the debrief:

> Originally, I wanted the students to look at all eight primary sources and determine which documents would provide the best support for their argument…I decided to scaffold this step …I specifically told group A to analyze Documents A-D and group B to analyze Documents E-H. This way, students would be able to spend more time analyzing their specific documents than trying to determine which documents should be used.

In an attempt to scaffold the task, Sally subverted interpretation of the documents, a decision that was fatal to the task's generativity.

When it came to the discussion, students had no way to rebut or evaluate the other side's argument because interpretation was given and evidence was equivalent to textual information.

Near the end of the lesson, Sally asked students to choose a position between two poles in the classroom that represented success and failure of the New Deal. With the exception of a couple of students, almost everyone flocked directly to the center of the room, saying with their feet that it was both a success and a failure. Based on the discourse I observed, this huddle in the center suggested that students had no means to negotiate between conflicting evidence.

Outset thinking elicited. During the New Deal lesson, Sally elicited thinking about justification, ambiguously raised sourcing, and did not address contextualization or corroboration. At no time in the lesson or debrief did Sally explicitly mention the value of the ways that sourcing, contextualization, or corroboration could influence the interpretation of a source.

Most of the thinking that Sally elicited was thinking about justification (i.e., Noke, Dole, & Hacker's [2007] use of documents as evidence). When students were invited to think about justification, they were encouraged to support an argument with facts or arguments that came directly from one or more of the sources. I delineate justification from the other historical reasoning heuristics because it does not necessarily require a historical perspective.

Although most of what Sally tried to elicit was thinking about justification, on two separate occasions in the lesson I observed Sally ask about the date of Document H. When I asked her about what she trying to accomplish with these questions, she explained:

> I wanted them to see … that this proves that the New Deal was a failure because the little girl wrote the letter…four years after the New Deal went into effect...I just wanted them to kind of say it shows that the New Deal was not successful because people were still really struggling.

This led me to ask Sally whether she had goals specifically related to sourcing. Sally explained that sometimes she points out the source but the goal of the lesson was not sourcing. Instead, she explained that the goal of the lesson was to learn "the right way to analyze primary sources and not just look at it and be like, 'It says this' but to say 'It says this and I think that that means this' or 'This supports my idea that it was a success or failure'."

I was surprised by this response because Sally explicitly pointed students to the date of the document to prompt the inference that the New Deal had not been successful despite many years of implementation. This seemed like sourcing to me, despite the fact that Sally did not explicitly mention it as such. She did not seem to notice that the inference she hoped students would make with the girl's letter required a specific type of reasoning, quite different from the justification reasoning required for the rest of the task.

When I asked Sally about the dates of the other documents in the packet, she told me that the only other date that was important was a graph that showed unemployment statistics over the years of the Great Depression. This referred not to the date of the source but the content of the document. When I specifically raised the issue of authorship and reliability on two of the documents, Sally did not see it as relevant to what she was trying to accomplish.

Despite Sally's apparently limited intention in eliciting reasoning about historical evidence, the documents she selected, the essential question, and the invitation to 'analyze' the documents still offered an opportunity for students to reason about evidence in history. Outside of the single reference to the date of the little girl's letter, however, the eliciting during the observation was limited to thinking about justification.

Interpreting at the outset. I analyzed interpreting according to two features: 1) what the TC reported noticing about student thinking and 2) the factors the TC reported considering for an

instructional response. Across the outset data, Sally noticed the form of student thinking and the level of student participation. The factors she considered prior to her instructional responses were primarily related to aligning students' thinking with her analysis goal.

Noticing at the outset. In the outset data Sally tended to notice two features in classroom interactions: form of student thinking and the level of student participation. Sally noticed the form of student thinking over the substance of student thinking. When I presented Sally with classroom exchanges from the first observation, she reported noticing where students were operating in the two-part analysis process. For the few students who immediately pointed to information in a document and explained how that information demonstrated an argument, Sally interpreted the answers as complete.

Instead of noticing the substance of the student thinking (i.e., how students were thinking with the content), Sally noticed the form (i.e., how they were organizing the content into answers). In this way, Sally seemed to only notice whether the student described the content of a document and explained how that content supported an argument. This tendency to focus on form over substance showed up in the pretest when Sally noticed particular observations that the students made (or failed to make) and noticed the way that the students formulated their answers that were more or less effective for argumentation.

When asked to describe what she noticed about the way a student was using the documents during an interaction in the New Deal lesson, Sally pointed out that the student read and understood that the WPA was providing hot meals "but then it's not being able to do the evidence thing–again, they can read it and say, …it says that they're providing food…okay, how does that show it was successful?...they could do Part 1 but not Part 2." What Sally noticed in the

substance of the student thinking seemed to be constrained by what she noticed about the form of the student's answer.

Even when the substance of the student thinking exhibited flaws, Sally remained focused on form. For example, in the rest of the above exchange, the student explained that the WPA provisions of hot meals showed that the government was meeting peoples needs and thus, that the New Deal was a success. Once the student connected evidence to the claim, Sally determined the discourse was complete. Of course, the provision of hot meals for poor people does not necessarily indicate anything about the success of the New Deal, only that the government was providing assistance.

Similarly, in reference to one student's explanation about Roosevelt's Fireside Chats, Sally said, "When he answers questions, he answers very well and very thoroughly." As I understood it, the student reasoning in the comment Sally referred to above was actually not particularly clear. S1 described how Roosevelt was providing jobs for one quarter of the population and reasoned that more jobs would restart the economy, thus making the New Deal a success. Sally did not appear to notice S1's apparent suggestion that the provision of jobs was a part of the New Deal, not necessarily evidence for its success. Nor did she notice opportunities to explore the reliability of the source of the information that the student was using, given that Roosevelt authored the speech.

While the tendency to notice form over substance was similar across Sally's observation and pretest, another theme I recognized in Sally's noticing was particular to Sally's early observation. When Sally interpreted student thinking in real-time during the New Deal lesson, her interpretation was impacted by an awareness of student participation. Student participation was certainly an important matter to notice, given Sally's explanation that students normally

engaged only in traditional activities of school history such as lecturing, independent reading, and worksheets. However, her attention on participation seemed to be conflated with student engagement in the reasoning her essential question demanded.

In debriefing some of the interactions during the New Deal lesson, Sally expressed relief that students were even following basic instructions. For example, when I asked Sally about what she noticed about a student's treatment of a document, she responded, "Well, like they actually read them because sometimes that's a struggle in our class."

Sally seemed weighed down by her perception that students might be bored and might not want to participate. When I asked her about her goal for the lesson, Sally described her anxiety saying, "Some of them are very reluctant to volunteer information, a lot of them just don't talk… and I wanted to see if I could … establish the discussion … and get them to talk."

Sally described noticing that the discussion was not student-led in the sense that she wanted to "be able to say, okay, so Team A start presenting your argument…and the first person finishes and the other person picks right up instead of me saying, 'Okay does anybody else have anything to add or say'." Sally expressed multiple times how pleased and surprised she was at the student participation in the lesson. She described the way that the lesson ended with a "take a stand" and how "pretty much everybody got up and did it," as if she expected something else.

Factors considered at the outset. Based on Sally's explanations in the early data, the factors she considered before responding were primarily driven by her vision for the particular argument form described above. Sally's interpretation of student thinking in the pretest, early coursework, and discussion lesson observation appeared largely governed by an assessment of whether the student contribution was aligned with the two-part vision of analysis.

When Sally noticed that student answers did not demonstrate the predetermined form, she decided that the students could not do it on their own, and considered ways to "get them to see" how to analyze the documents. This tendency was similar to her interpretations in the pretest. That is, Sally had particular answers in mind that would indicate whether students were accomplishing the two-part analysis. What she described was akin to a mental checklist. First, she would ask herself whether they comprehended the content of the document. Second, she would ask whether they connected that content to an argument.

Similarly in the New Deal lesson, Sally seemed to run each interaction through this preconceived, two-part filter. Did the student identify particular observations about the document and could the student connect the observation and claim with an explanation? When a student did not "get it," Sally considered how to "get them where you want them." Getting students where Sally wanted them required diagnosing where in the two-part analysis the student was stuck and providing support to help them accomplish the analysis.

Given the participation anxieties that Sally held, she weighed instructional decisions in a tension between what the student needed to do and what Sally thought the student would be willing to do. When she interpreted student analysis and considered a response, therefore, she did so under pressure to maintain engagement, participation, and a positive relationship with students.

This tension was evident when I asked Sally to talk about an exchange in which she used hints and prompts to guide one group of students to an answer during the New Deal lesson:

> I mean, this is kind of what happens a lot. Eventually they get where you want them to and sometimes it just takes a lot of leading…like they know a lot more than they pretend that they do and every day we live through that. It's just like they try so hard to not know

the right answer. And, I'm like, 'No, you've got it.' So…if you talk them through it, they

pick up on it pretty quickly.

The specter of non-participation contributed significantly to how Sally saw the lesson and

figured into the factors Sally considered before responding. She described her anxiety:

I was a little worried … a lot of the times they don't talk, like, you can cold call on them

and they're like 'I don't know' and I mean, you can point to the answer and be like, 'just

read this' and they just won't read it. And they're just like, 'no, I don't feel like it, I don't

want to read it'. So I was very worried about doing a full on discussion with this class

because they don't talk.

The possibility that students might flagrantly ignore her instructions suggested that maintaining

student participation was an important factor that figured into Sally's early interpreting.

Responding at the outset. In analyzing response patterns, I considered how the TC

responded to student thinking when students vocalized their thinking about evidence in history.

At the outset, Sally's instructional responses to student thinking about evidence in history were

either evaluative low responses or leading Medium responses.

Sally often responded to student thinking but the response was simply a positive

evaluation (e.g., Yes, Good) or a rebroadcast of what the student said, which functioned as an

affirmation of their answer. For example, in the following exchange, a student asserted that the

New Deal was a success as demonstrated by a picture of people working for the WPA:

S12:   WPA is able to fund these projects and lowering unemployment is just more evidence

that the new deal was successful.

Sal: Good. S17, what were you going to say?

Here a student made a claim based on the idea that the provision of jobs by the government meant that the New Deal was a success. Because he described the content of the document and explained how that content supported his argument, Sally evaluated it as "good," rather than responding in some way that challenged the reasoning that the New Deal was a success because jobs were provided by the government. From that point, Sally moved on to another student's idea.

The second response pattern was to lead the students to a complete answer. That is, many of Sally's responses had the form of exploring student reasoning but the exchange actually served to lead students to a 'right answer.' This pattern was evident in exchanges in which Sally asked students a series of yes/no questions or fill-in-the-blank questions until they arrived at an end that Sally had in mind.

For example, when one student said that she did not understand a passage from Roosevelt's speech, Sally responded by walking the student through a justification that the New Deal was a success.

Sal: Mortgage distress. Do you know what that is?

St: No

Sal: (explains mortgage). So, if he is talking about easing the distress of house payments then who would that help?

St: People?

Sal: People who…?

St: Could not afford to live in a house?

Sal: So, he is talking about easing the distress of farmers and homeowners. Remember what happened to farmers in the dustbowl? When their crops failed they could not afford...

St: Food?

Sal: What else could they not pay?

St: Mortgage?

Sal: Right. So Congress is about to pass legislation…So, that would be your specific

        evidence. So, on this part where it asks how does that support my argument just write

        how those two things you wrote down show that it was successful.

In the exchange above, it was unclear what the student understood about the source, the

historical context, or even Sally's questions. Sally responded to the few student contributions by

reformulating the student's comments, but the connection to actual student thinking was vague,

at best. Sally was doing the thinking in this exchange and she only stopped when she arrived at

the answers sufficient to add to the graphic organizer.

        Sally's low and Medium responses followed a pattern that matched her two-part

"analysis" of documents. When a student knew how to describe the content of the document and

explain how that content was evidence for one side of the argument or the other, Sally responded

with a positive evaluation. For example, one student explained evidence that supported his

argument:

S12:   In the Fireside Chat, I quote, "First we are given opportunity of employment to one

        quarter of a million of the unemployed." During this time of the Great Depression,

        unemployment was very high and many could not provide for their families and could not

        get jobs. If people could get jobs then they would have money and restart the economy,

        proving that the new deal would be successful.

SS: …I can't get all that down

Sal: I just said summarize it…Listen. So, Document A. Please summarize what you said.

111

S12:   A massive increase in employment money would be made to turn around the economy.

Sal: So, Roosevelt provided lots of jobs that helped people make money. It is as simple as that.

When the student was able to provide the content of the document and explain the link to the assigned argument, Sally always moved on.

When a student comment demonstrated an inability to describe the content of the document or explain how content related to an argument, Sally consistently engaged in Medium responses–follow-ups that take the form of a response but are not actually responsive to student thinking–to lead them to a completed answer. For example, the following exchange occurred when a student was confused about one of the documents:

Sal: So who got the better jobs?

S4:   Whites.

Sal: And, who got higher wages?

S4:   Caucasians. And (blacks) got the bad jobs.

Sally: …So, how does that show that the New Deal was a success? What you just wrote down was evidence. How does that show your argument?

S5:   Everyone was not equal.

Sal: So, that is what you could say. Everyone was not equal.

When Sally was confronted with a student who did not know what to do, she determined first whether they could summarize the content and, second, whether they could explain how the content was relevant to the argument. Regardless of which aspect the student had yet to accomplish, Sally responded by doing the reasoning for them.

Often, students could describe the content of the document but not the explanation. In most of these instances, Sally responded in a way that I initially coded as High I. That is, the response focused on a student idea but put teacher reasoning about the student idea on display. After the first pass of the data, however, I created a code for Sally that represented a pattern I noticed, a pattern I suspected was more akin to a Medium response. I named the code "how does this help?" because Sally used this phrase so frequently in the New Deal lesson.

This pattern of responding took numerous forms but each was designed to lead students to a right answer, or completion of reasoning that Sally had in mind. For example, Sally had this exchange with a student one-on-one:

S15:   In Document D people are adding electricity to rural America.

Sal: Ok, how is that helpful that they were adding electricity to rural areas and building highways? How is that helpful?

I decided that form of the question "how does that help?" show the argument, was not actually a probe of student reasoning designed to promote discourse but instead was designed to lead students to a right answer that could be written on the graphic organizer.

This appearance of High I responses which turned out to be Medium responses was not limited to "how does that help" questions. For example, Sally sometimes asked students to cite a document, as in the following:

Sal: What did you write S18?

S18:   From 1929 to 1940 unemployment was very big and during 1933, even after Roosevelt's election was the biggest unemployment percentage.

Sal: What document is that?

S18:   Document F. Even after that year it never got better.

Sal: (to another student) what did you write?

Although a request to cite a document can be a means to promote discourse about historical reasoning, Sally's use of the move was not an attempt to generate student reasoning about evidence.

Because Sally's task in the New Deal lesson was designed only to have students use the content of the documents to identify and explain support for an argument, she did not intend for students to disagree on how the quotes should be interpreted. Once a description of the content was matched and explained, Sally saw no further need for discourse. Sally approached almost every exchange as if explaining how document content supported an argument was an end in itself. The implication was that there were right answers to the only two questions that appeared throughout the lesson: 1) which argument does that description go with and 2) how does that description support the argument?

Because Sally used primary source documents in early lesson plans, there were frequent opportunities for her to respond in ways that could have promoted student reasoning about evidence in history. Her responses, however, consistently suggested that student reasoning was complete if it addressed Sally's two-part analysis structure. In light of this, Sally's responses in the early data tended to affirm student answers or lead student thinking to a desired end.

Change in eliciting. Sally's eliciting remained remarkably consistent over the course of the program. She used tasks that, at times, were potentially generative but were undermined by her launch of those tasks and her invitations for student reasoning.

Change in generativity. Although Sally's lesson design and materials varied in generative design, Sally consistently did not launch generative tasks. In several lessons, Sally drew on materials and activity structures, similar to those in the New Deal lesson, that could have facilitated generative tasks. For example, the third lesson I observed followed a similar pattern to the New Deal lesson. Sally assigned students to one of two teams, which would use primary source documents to argue that Brown v. Board of Education was or was not successful in ending school segregation.

Rather than have students examine all the documents, Sally again assigned four documents to the 'yes' team and different documents to the 'no' team. Again, the graphic organizer prompted students to make observations about the sources and explain how each supported their assigned side. As in the first lesson, Sally had predetermined which sources best fit each argument and assigned only those documents to the respective team. Thus, amongst their team members, they only had to contemplate one argument and were only faced with the documents that Sally provided to that side, all of which seemed to support their assigned claim.

When it came time to present the documents to the other team, students just summarized each document and explained how it demonstrated the assigned claim. Similar to the first lesson I observed, Sally had students "take a stand" at the end of the final discussion. Again, this resulted in a mass of students clustered in the center of two poles, shrugging their shoulders about any way to answer the historical question.

As in the first observation, the Brown v. Board task was adapted from materials on *Historical Thinking Matters* and the activity structure was an adapted form of a methods course IA. In observations during which Sally did not draw on these two resources, the opportunity for reasoning about evidence was sometimes entirely absent.

For example, in the second lesson I observed, Sally conducted a lecture on the Holocaust with the support of a moving PowerPoint presentation. After the presentation, students participated in a "gallery walk" of pictures of the holocaust and readings from interviews with holocaust survivors. While the task was emotionally stirring, Sally explained that her use of primary source documents was to promote student interest. Few of the photographs were sourced and captions gave only scant contextual information, if any at all. Even if students had been invited to reason about the evidence (e.g., Identify the sources to determine which provide the strongest evidence that the U.S. knew what was happening and did not intervene), the resources available in the classroom did not allow for such reasoning.

Although the opportunity for historical reasoning was present by way of the documents and essential questions in some of the lessons I observed, Sally rarely launched any task that positioned students to reason historically about evidence and argue diverse interpretations. In fact, Sally's use of graphic organizers and adaptations of activity structures actually undermined interpretive discourse because the tasks only invited a set interpretation of the meaning of each primary sources.

Changes in thinking elicited. Throughout the time I observed Sally teach, she elicited justification of arguments, but did not elicit sourcing, contextualization, or corroboration. Sally continued to invite students to engage in a two-part "analysis" that included citing textual data and explaining how those data supported a claim.

As in the outset data, the first half of the analysis was a straightforward description of "this is happening," rather than an interpretive interrogation of the source or context, as compared to the other available documents. Evidence, as Sally consistently approached it, was straightforward information from the content of a source that could be used to back up an

116

argument. For example, the following exchange from the Brown v. Board lesson illustrates the thinking that Sally targeted:

> Sal: So you are doing Document D? So, what is Document D? A photo of…?
>
> S: Kids walk to school … soldiers (partly inaudible)
>
> Sal: So, how does this show that Brown v. Board was successful in ending school segregation?

That was it. Once a student could answer those two questions, he was ready to make the argument.

Rather than eliciting thinking about evidence in history by asking students to consider the author, date, type of document (e.g., What is important about one being a federal report and the other being an op-ed), or historical context (What was happening at the time that helps us understand that?), Sally asked the students to build an argument by describing the content of the source and explaining how that content demonstrated an assigned claim. Because this two-part analysis did not typically depend on details important to historical reasoning such as source or context, Sally had little trouble eliciting similar thinking in the holocaust lesson and in other lessons in which content was accessible but details key to a historical analysis were not.

Because Sally's vision for analyzing historical sources was limited to a straightforward reading of the source and application of that content to an argument, students were not invited to "turn the sources into evidence" by any historical reasoning. Instead, Sally elicited reasoning about what the source "said" and then explanations about how that supported an argument. I never saw Sally invite a student to challenge the straightforward "meaning" of a source by eliciting thinking about the source or context that might rebut an initial interpretation.

The reasoning that Sally wanted was akin to thinking necessary for general argumentative reasoning and writing. In general argumentation a student states an argument, supports that argument with details and explains how those details support the stated claim. In this regard, Sally actually demonstrated a stable upward trajectory, scaffolding students' ability to make an argument, support that argument with details, and frame the writing with a formal structure for argumentation. What counts as evidence for a historical argument, however, is not any information, but information gleaned from an historical analysis of primary sources

A number of possibilities exist to explain Sally's lack of progress in eliciting student thinking about evidence. Sally's consistent interest in eliciting general argumentation may be explained by her placement context. In debriefs, Sally regularly mentioned the structure of DBQs as a model for organizing student thinking about the use of texts as evidence. During the third observation, Sally even mentioned to the students, "You want to make sure that you use primary sources to come up with evidence to support what you say….Remember when you write DBQs and thesis statements?" Sally also described general writing and thinking skills needed for DBQs as a central goal of her co-planning colleagues and mentor.

I observed nothing in Sally's context, beyond the weekly methods course, that would have encouraged the eliciting of historical reasoning. Sally's mentor teacher did not approach history as an evidence-based, interpretive activity and the students rarely raised any such ideas on their own. But, if placement context was the only factor determining Sally's approach to eliciting student thinking about evidence in history, I expect that she would have explained this to me in the debriefs because we frequently discussed the advantages and limitations of her placement. I suspect that another primary factor was Sally's relatively weak disciplinary

understanding of history, which likely influenced a limited vision for what it meant for her students to reason about evidence in history.

Both the observations and debriefs suggested that Sally had a limited understanding of how history was constructed, a factor that certainly played into her deference to DBQ-style analyses of primary source documents. Sally believed that primary sources could increase student engagement but her use of these sources was not significantly different than one might use any text, historical or otherwise.

Another potentially influential aspect of Sally's context was an attitude among some of the teachers on her team that "these kids can't." Although Sally loved her students and worked hard to support their learning, I noticed many times in our interviews when Sally expressed doubt that many of her students could do anything other than the most basic tasks. It is difficult to know whether a more ambitious perspective on her students' abilities might have pushed Sally to overcome limitations in her own understanding of history and limitations in the way she understood her responsibility to support student reasoning and writing.

While at first glance, it seemed that the availability of materials and tools provided by the methods course had little impact on Sally's eliciting reasoning about evidence, it is worth another look. A significant difference existed in the generative possibility of Sally's lessons when she used resources from *Historical Thinking Matters* and an IA practiced first in the methods course. When Sally designed tasks independent from these materials and tools, as in the Holocaust Lesson, the designs actually prevented historical reasoning entirely.

Despite Sally's limited ability to actually elicit student thinking about evidence in the classroom, her use of essential questions, sources, and IAs at least created an environment in

which a teacher could have elicited such thinking. Without these tools, Sally's use of primary sources simply supplemented pre-packed historical narratives. Although I conclude that the availability of certain tools did not result in attempts to elicit student reasoning about evidence, these tools helped create an environment in which eliciting student thinking about evidence was possible, if Sally had the capability and support to pursue it.

Interpretation changes. Throughout the remaining observations, Sally's pattern of interpretation changed only slightly. The themes I identified in her early interpretations largely held throughout the course of the study.

Changes in noticing. For the most part, Sally continued to notice the form of student thinking over the substance of student thinking. For example, after the Brown v. Board lesson, Sally described the student thinking that she noticed:

> I mean a lot of them did really well where they looked at the photograph and (were) like, … there is a photograph of African-American and a white a girl sitting like they are about to start talking to each other…And … they are in the same classroom and so …obviously that shows that they are desegregated. So they used those really well…

When she noticed the correct form, it seemed to override any other aspects of the student thinking. For example, one of her students provided a particularly eloquent explanation that included a description of the document's content, a quote from the document, and an explanation for how the information supported the claim. When I asked Sally about it, she said:

Sal: S1 is really good…what he writes down is always above and beyond what everyone else writes down. I loved that he actually used a quote …He says it and he backs it up, like this is your quote and this is why I am using this because it supports what I am saying. So, he basically did what I was trying to get everyone else to do.

I: What was it about what S1 said, like using the quote, that made it better than just summarizing?

Sal: Just because they don't quote things… Not, only is he able to like pull a quote from it. But, then he explained the quote. So, like he gave the quote, explained what it was saying, and then tied it in with "this supports my evidence because."

Because Sally only noticed the form of S1's statement, she did not recognize that he was arguing that Brown v. Board was successful in ending segregation because the Supreme Court ruled that segregation was illegal, an assertion that was unclear at best.

In some situations, Sally seemed to notice student answers that did not fit within her preconceived framework. In the Brown v. Board debrief, I asked Sally about a time when she appeared to stumble in her response to a student statement about one of the newspaper sources. The student observation went well beyond Sally's preconceived answer:

So he was pointing out that it's like a nationwide thing, it's not just like one state that segregation is ending…So that was where he was going with that, or that's how I took where he was going with that and he didn't correct me. I'm assuming that's what he – because it threw me off too because I just wanted them to look at it – like the headline, it says 'school segregations banned' … that was all I was hoping to get out of it. So it threw me off … But that sounded like what he was going for. So that was how I re-voiced it and he kind of nodded his head.

Although Sally did not stop to explore the student's thinking, it was clear that something about the interaction was disruptive for her. The student thinking disrupted her pattern of interpretation that I observed in many other lessons.

Even though Sally did not articulate it, I suspected that her inability to rebroadcast the student's thinking and her doubt, even during the debrief, was because she recognized that the student thinking was disruptive to her interpretive pattern. Sally also continued to notice student engagement in her instructional interactions and judged lessons as effective chiefly by the level of participation. She often just seemed glad if students were following instructions, much less engaging in particular types of thinking.

Changes in factors considered. The factors that Sally initially reported considering held for the remaining lessons. Sally weighed student comments against an expected answer and mostly evaluated the student thinking as complete or incomplete. When students did not answer with a completed analysis, Sally considered how to "get them to see" what she intended. For example, when I asked her about a particularly leading exchange she had with one student, she explained:

> Yes, so I was just trying to get him to look at it and see, like, you know…"Hitler's whole goal was to get rid of them so that was kind of the point….So like, how does that show how Jewish people were affected?"

As in the earlier lessons, there was a particular observation and a particular inference that Sally had in mind. She wanted the student to say that he saw Jewish people in prison camps and that showed that they were affected badly. Although such an assertion is tragically true, it fails to point to the means by which historians know that it is true. Simply drawing conclusions based on source-less photographs leaves students open to the misconception that history is simply a matter of one persons evidence against another's.

This pattern continued in a lesson that Sally launched by having students examine the high school's yearbooks from a period covering 1951-1975. When Sally asked the students to

explain what they noticed about the kids in the yearbooks up to 1965, she got many answers including that they looked happy, were dressed properly, and all had short hair. Finally, she told them, "The most important thing I wanted you to notice was that they were all white, ok?" When several students agreed, Sally led the class with another question:

Sal: In what year did you start to see African Americans?

SS: 1967, 1978

Sal: So you don't actually start seeing African Americans until 1964 or so. So, again Brown v. Board ended segregation in 1954. But in our high school you don't start seeing African American students until 1964 or so. That should indicate to you that even though Supreme Court decided to end segregate, it did not immediately end.

When I asked Sally about this interaction, she said, "Once they finally figured out where I was going with it, they got it, but at first it was kind of a struggle." Again, Sally noticed whether students did or did not make the observations and inferences she expected.

When Sally noticed that a student did not respond as she expected, it did not appear to change her pattern of considering a response. Although Sally reported noticing thinking that disrupted her strategy for analysis, this student thinking did not appear to change her consideration. When confronted by student thinking that did not fit her formula, she reformulated that student thinking into something that she wanted to "get the students to see."

One change in the factors Sally considered was anxiety about student participation. Over time, Sally developed more confidence in her role as teacher and was reportedly less fearful of asking students to engage in the work she envisioned. However, her evaluation of the success of lessons was gauged by level of student participation and not by the quality of student thinking during the lesson.

Changes in Response. During the remainder of the study, Sally's patterns of response remained primarily low evaluative responses and Medium leading responses. What first manifested as "how does that help" in the New Deal lesson continued to be a primary response that Sally used to get students to do the second step of analysis. That is, she wanted them to explain how their observation supported the claim they were arguing. Sally stuck with a pattern of response that supported students in organizing an argument based on a straightforward reading of source content. The following exchange from the school desegregation lesson was indicative of this pattern:

S3:   I did Document B. The US army escorts black kids into the school.

Sal: Ok, so that doc was a photograph of the army escorting blacks into the school. How does that show that Brown v. Board was successful in ending segregation?

S3:   I did not do that.

Sal: Well, why? So, why? Like very simply, like…

S4:   They are protecting students.

Sal: Good. So they are protecting students that are going to a desegregated school. What does that show?

        (Silence)

Sal: That it worked. Ok? That was all. You were thinking too much. So, if you have the army protecting students going to a desegregated school, that shows that they are actually going to that school. That shows it was kind of successful then.

Although Sally responded with the question "Why?" this responding move stood in place of "How does that observation demonstrate the argument?" That is, she wanted the student to explain how the initial observation of the document supported the argument that Brown v. Board

was successful. As was usually the case, the response was not intended to promote historical reasoning about the evidence but rather to arrive at a completed answer.

In some cases, students went off script and raised questions or topics that were not immediately relevant to the two-part "analysis" of the document that Sally envisioned. Her responses brought them quickly back to the analysis, as demonstrated in an exchange about a students' observation of a photograph of students at Central High School:

Sal: Good. So you looked at the picture of The Little Rock 9. So, if you were African Americans and were trying to go to school and there are troops blocking you, does that show you are successful in ending segregation?

S8: If you are light skinned can you go to school?

Sal: Some areas had different rules. So, you said no it is not effective. Why does that show it was not effective in ending segregation?

S8: Because they could not go in there.

Sal: Right, good.

In this exchange, Sally affirmed the observation and then asked the student to do the next part of the graphic organizer (explain how the observation supports the argument). Instead of answering Sally's question, however, the student raised one of his own. Sally's response was to brush it off and get back to the analysis. Throughout most of the observations I conducted, Sally continued patterns of response that were aimed at getting students to engage in her two-part analysis. Once students got to a completed answer, the conversation was almost always over and Sally either walked away or moved on to another topic.

Because Sally elicited little historical reasoning, she had few opportunities to respond to thinking about evidence in history beyond thinking about justification, or use of documents as

evidence as defined by Nokes, Dole, and Hacker (2007). Occasionally, Sally's students engaged in sourcing, contextualization, or corroboration despite few supports to do so. Rather than respond by facilitating further discourse, however, Sally tended to marginalize or highjack this thinking. In the following exchange, a student goes off Sally's script:

S5:   The newspaper shows non-segregated schools. The headline shows that segregation has
        been banned.

Sal: So, Document C is a newspaper article with a headline that says 'school segregation
        banned.' So, that shows that Brown v. Board was successful because this paper says there
        is no more segregation. And, S15 would like to add to that. (hand up).

S15:   If you look in the bottom corner it shows that it is effective in a completely different
        county as well. That just shows you how widespread and influential the verdict of the
        Supreme Court was in desegregating schools.

Sal: Ok good. So, S15 is pointing out how it was occurring in different counties. And, that
        shows how it was kind of all over the US. Anyone else from Team A have anything? Any
        documents we have not talked about?

In this exchange, Sally was presented with two opportunities to engage student reasoning about the evidence. She allowed the first student's observation to go unquestioned and then reformulated the second student's analysis of the document before moving on. A helpful prompt might have been something like, "Can you show us what you are looking at in the corner of the newspaper and explain what you mean by that?"

At the time, I noted that Sally's reformulation of S15's thinking was not true to the student thinking or the historical source. Rather than explore the thinking, Sally reformulated it in order to support the two-part analysis she wanted the students to complete.

In my final visit, I observed Sally teach a lesson on distinctions between the philosophies of Malcolm X and Martin Luther King. I observed a teacher who had gained significant confidence and poise and who could handle the many distractions of a busy classroom. But, I also observed a teacher who had perfected low and Medium responses in the face of student thinking about primary sources. With almost every student comment, Sally rebroadcasted the statement and moved on (low response) or coopting the statement and led students to a complete answer (Medium response).

Sally improved dramatically across the course of the year in terms of managing class dialogue according to a particular purpose. Unfortunately, however, Sally's responses tended to constrain student thinking about evidence rather than propel it. As a consequence, Sally's responses constrained her students' ability to understand evidence in history.

Summary of Key Findings

In this chapter, I have considered how Sally elicited, interpreted, and responded to student thinking about historical evidence. My first research question asked how Sally engaged in these practices at the outset of the program. Initially, I did not see evidence of an ability to launch generative tasks, despite the potential of the documents and task structures Sally chose. During the outset tasks, she elicited thinking almost entirely about justification. Her interpretation of student thinking appeared largely governed by an assessment of whether the student contribution was aligned with the two-part vision of analysis. In light of this focus, Sally's responses in the early data tended to affirm student answers or lead student thinking to a desired end with Low and Medium responses.

My second research question asked how Sally's capabilities to engage in these practices changed during the study. Sally consistently did not launch generative tasks because she adapted

materials and activity structures in ways hampered that generativity. Sally continued to invite students to engage in a two-part "analysis" based on justification that included citing textual data and explaining how those data supported a claim. Throughout the study, Sally's instructional responses to student thinking about evidence in history were either evaluative Low responses or leading Medium responses. Her responses tended to constrain student thinking rather than further it.

Chapter 6: Gabby

Gabby was a 25 year old white woman with some non-traditional teacher experience. After graduating with double major in government and political science from our university, she was accepted to enter a genocide studies Ph.D. program. Instead of entering the program straight out of undergraduate, she decided that she wanted to do something totally different.

At 22, Gabby moved to southern Thailand and taught English in a small, public elementary school. She described the biggest shock of the experience as the slower pace of life: "Growing up on the East Coast, I had a hard time figuring out how to relax." Gabby described herself as Type A, highly organized, and a person who likes "to get things done well in advance." Nonetheless, her time in Thailand must have had some impact because I found her to be funny, laid-back, and extremely creative.

Gabby was a member of the summer methods course I taught prior to Social Studies Methods II. Although the names and descriptions of the historical thinking heuristics I introduced in that class were new for her, Gabby was comfortable with the concepts from an early stage. By the end of the summer, Gabby scored highest in the 13-person cohort on historical understanding in the SS Methods II pretest. Although not a history major, Gabby had a strong understanding and interest in U.S. history post-Civil War, knowledge that would serve her well in the 9th grade U.S. history placement she eventually assumed.

Despite her teaching experience, Gabby initially described classroom routines and discipline as areas in which she felt least confident at the beginning of the internship. When she imagined a social studies classroom prior to beginning the program, she emphasized the importance of group work and student engagement through video clips, PowerPoints and posters. By the end of the summer, Gabby was envisioning a classroom grounded not only in student

engagement but also in "creat[ing] written arguments using primary source documents, advancing reading comprehension, and contextualize[ing] different periods of history."

From early in the year, Gabby worried that the rigorous classroom she envisioned for her students might outpace their willingness or sour them on her as a teacher. Periodically, she wondered aloud whether she was "burning her students out" with the expectations that she placed on them. In my observations, however, Gabby was well-loved by her students and highly esteemed by her mentor teachers.

Context

Gabby was enrolled in a track of the program that required placement in both middle and high school internship experiences. Her initial placement was in a seventh grade human history and geography class with a mentor whom she adored. After three months at the middle school, she moved to a high school in the same district, with similar student demographics. Of the roughly 1,300 students, over half were white and another quarter were Asian. About 7% of the students were Hispanic/Latino and the remaining students were African American (6%).

Because the district was populated by relatively affluent families, only 6.3% of students qualified for Free/Reduced Lunch Program. The school had a small number of students who were Limited English Proficiency (6%), none of whom were in Gabby's classes. She taught only honors and gifted/talented classes, which also meant that few of the students enrolled in special education were in her classes. The class I regularly observed was a 9th grade gifted/talented U.S. History class that used the same materials and lesson plans as Gabby's other classes.

Despite the fact that Gabby's classes were "upper-level," she noted that the evidence suggested that students had never worked on argumentative writing and had certainly never been

explicitly taught historical thinking skills. She described her task in regard to writing and historical reasoning as, "working totally from the baseline."

Gabby described her new mentor teacher as "very laid back," a fact that had both positive and negative repercussions. On the one hand, the mentor gave Gabby the freedom to experiment and operate the classroom however she chose. On the other hand, Gabby described the classroom before she took over the teaching as "lots of fun stuff like movies." The expectations and norms that the mentor had established with the students initially made it difficult for Gabby to implement the ambitious approach to history that she envisioned.

According to Gabby, the mentor did not ask the students to write regularly, did not often use primary sources, and never taught historical thinking skills, all goals that Gabby mentioned in early course reflections. The district curriculum had recently been revised and included many links to the common core and explicit directives to teach historical thinking skills. Gabby mentioned "feel[ing} sorry for" her mentor because the mentor's materials and thinking about social studies were not well aligned to the reading and writing-focused requirements of the new curriculum.

In the times I observed Gabby, I noted that she had a great rapport with her students and I was often shocked by the level of rigor in which they were willing to engage. Gabby described students in all the classes she taught and observed as having an "attitude of 'We are here for a reason'." Although she was under no illusion that the school was perfect, Gabby was initially amazed at the way students responded to the teachers' high expectations. She frequently described her students as "incredible" and described their work as "fantastic."

Eliciting, Interpreting, and Responding

In the following section, I describe Gabby's ability to EIR over the course of the study. The data that informed these findings include information drawn from the pretest she took prior to coursework, course participation in the following months, six classroom observations and seven interviews. My first observation of Gabby's classroom teaching was mid-October 2014 and my final observation was in March 2015.

Eliciting at the outset. As described in Chapter 3, I analyzed eliciting according to two features: 1) generativity of the task the TC launched and 2) the specific type of historical thinking the TC elicited. At the outset, Gabby launched generative tasks but struggled to communicate clear expectations to her students. She elicited thinking about justification and irregularly elicited thinking about sourcing.

Outset generativity. In the outset data, Gabby demonstrated an ability to envision tasks that positioned students to engage in diverse interpretations of historical evidence. However, when it came to launching such a task, she initially struggled to articulate her vision to the students.

The first lesson I observed Gabby teach in her new placement was the course-required discussion lesson. She launched a generative historical task around an essential question, required students to explore multiple primary sources, and facilitated student discourse in an structured academic controversy (SAC). Student desks were arranged in pods of four and each counter-facing pair was assigned to one side of the question, "Were 19th Century industrialists robber barons or captains of industry?"

Before they began working, Gabby established a common definition for both descriptors in the question saying:

G: You have one definition there. Captains of Industry and Robber Barons. Go ahead and write down those two definitions. So, those people you read about who made billions in oil, railroads and steel: these are two ways that you could use to describe 19th century industrialists.

S2:   (S2 defined captains of industry)

S3:   (Defined RB and explained) they exploited workers...

G: So, do they mean the same thing or completely different things?

S1:   (Silence)

G: Do they mean the same thing or are they telling two different stories?

From the beginning of the lesson, then, Gabby established a generative problem space by demonstrating that arguments could be grounded in multiple interpretations of the same phenomenon.

To drive the point home, Gabby introduced students to several political cartoons that depicted robber barons, and conversely, captains of industry. Because this was the first time students had engaged in a SAC, Gabby emphasized that the task was not a debate. Rather, she explained that students were working together to better understand both sides of the argument and eventually come to a consensus as a foursome.

Gabby left it up to the student to read the discussion guide and follow the directions. The first step was to:

Read through your documents to find evidence that supports your position. Feel free to highlight or underline text that helps support your point. Using your Document Analysis Chart, list 3 pieces of evidence you will use to present your position.

The guidelines then prompted students to follow a specific pattern of self-guided discourse.

Step 2: Position Presentation
- Side A presents their position using supporting evidence from the texts.

- Side B <u>restates</u> to Side A's satisfaction.

- Side B presents their position using supporting evidence from the texts.

- Side A <u>restates</u> to Side B's satisfaction.

Step 3: Consensus Building
- Abandon roles.

- Build consensus (come to an agreement) on the issue using evidence from the documents.

- If you can't come to a consensus, explain why you are having group gridlock.

The design of the task, based on a SS methods course instructional activity (IA), positioned the students as historians, tasked to use interpretation and evidence to argue their way to an evidence-based answer for the essential question.

The design of the task offered significant generative potential. However, Gabby initially struggled to launch the task in a way that prompted the student thinking she envisioned. During the debrief, Gabby criticized the way she gave directions at the start of the task, suggesting that her directions were vague and the process should have been modeled. This point was especially problematic because it was the first time students had engaged in an SAC and the first time they had been asked to analyze primary source documents in this way.

As she directed students to get started reading and working in pairs, she had the following exchange with a group in the front of the room:

G: So, Position A, read through your documents and then start talking with your partner.
       Then, Position B do the same thing.

S1:   So, what do we do?

G: So, there you are writing quotes from the documents. Things that support your side. As

       work together to find evidence and Bs work together to find evidence. Then, you will

       share with people across from you.

As Gabby made her first lap around the classroom, it was clear that many students did not

understand how the SAC was supposed to work. Gabby spent about 10 minutes during the

evidence-gathering period re-explaining the procedure to each group.

       During her second lap around the classroom, Gabby spent most of her time correcting

students' idea that simply writing quotes on the chart would be enough. She told most of the

groups, "You have to be able to explain how the evidence proves your side," a piece of

information that she breezed past in the instructions and left off the document analysis chart.

That chart provided space for evidence from each document but did not provide a specific space

or a specific prompt for an explanation that would connect a quote to a claim. Once students

realized what was required, Gabby was able to spend her time checking in and determining

whether they were only collecting quotes or actually explaining those quotes in ways that

bolstered their assigned arguments.

       The structure of the task required students to do something that they did not seem

particularly comfortable with: take a position that might not align with their initial opinion. In

several groups, students wanted to take the opposite side than they were assigned. Gabby

exhorted the students to engage in interpretation, saying, "The point is to be able to support any

position. You don't have to say you believe it. Just find evidence."

       During the debrief, Gabby explained that because she was relatively new in the internship

placement and because she appreciated that students had little experience with primary source

documents and self-guided discussion, part of her intention was to gauge what they could do.

While the task did not demonstrate a particularly strong link to learners' prior knowledge, it is perhaps understandable given that Gabby only began the internship a few weeks before. On the other hand, the task connected to content the students were studying and pushed them to develop an ability to defend an argument using evidence while introducing them to a discourse structure that promoted perspective taking.

Outset thinking elicited. Although the essential question and structure of the task was generative and held possibility for historical reasoning, Gabby elicited thinking about justification almost exclusively. During the debrief, she told me that she intended for students to 1) recognize a side of the argument, 2) identify quotes and/or textual information that supported that argument, and 3) explain how those quotes supported the argument.

Despite Gabby's justification-based vision for student thinking, neither the graphic organizer, discussion guide, or even the instructions made explanation an explicit part of the task. For example, she told the class, "Once you feel comfortable with the documents, start talking with partners to try to find evidence." As Gabby walked around the room, she frequently asked students, "You guys finding evidence?"

When one group was confused, Gabby pointed to the document analysis chart and said, "So, there you are writing quotes from the documents. Things that support your side. As work together to find evidence and Bs work together to find evidence. Then, you will share." In these and other interactions, Gabby clearly articulated an expectation that students would identify evidence that supported their position. But, she did not explicitly clarify what counted as evidence, what it meant to "explain," or even that they needed to explain how evidence linked to claim.

In a few small group interactions, however, I overheard Gabby prompt students to explain how the quotes they noted supported the argument. For example, she asked one group to present their arguments and said, "I want to hear quote and explanation." But, most of her elicitations were vague, as when she got the whole class' attention and prompted students to "find the evidence, find the evidence to support your position." Although some students simply read quotes when it was their time to "present their side," I was astonished to observe many students actually explain quotes and even argue for the side they were assigned because I expected Gabby's ambiguous elicitations to result only in confusion.

Throughout the lesson, Gabby's elicitations remained focused on justification, a process that required identifying evidence and explaining how the evidence was related to the assigned claim. Although she included a note on the graphic organizer that said, "Record questions about the sources … below," she did not focus explicitly on the source of the documents. Gabby seemed to make room for sourcing and other historical reasoning but neither directly prompted it nor explicitly taught it.

It is worth noting that if a historian engaged in this task, it would require a significant amount of corroboration. For example, an historian engaged in this task would certainly ask, "Am I finding the same information in each of these sources? Am I finding different versions of the story? Why would there be different versions of the same story about these industrialists?" But, Gabby did not explicitly teach the students how to cross-check the available sources in order to make sense of the conflicting accounts. Instead, she left the notion of argument and evidence somewhat vague in her explanation and simply asked the students to find evidence to support their assigned argument.

While observing the lesson, I assumed this lack of attention to historical reasoning was because Gabby did not value teaching historical thinking. However, she made clear in the debrief that anything beyond justifying an argument by supporting it with quotes and explanations was outside the intended scope of the lesson. She explained that she had expected students to "start to source ... So, 'Who are the kind of people who are saying these things'?'"

When Gabby described some of the arguments that students were making with the evidence, I asked her whether anything could have made the arguments stronger. She answered:

> …I think if they … had really looked at the sources. So a lot of them did pick up Andrew
> Carnegie. So the position he is coming from is going to have a certain bias…with
> Position A, the bias was more clear (and) with Position B the bias was less clear.

Even though it was beyond her stated intention, at a few points in the lesson, Gabby actually did elicit thinking about sourcing. For example, when faced with a group of boys in the back of the room who rushed through the activity in only a few minutes and claimed to have finished. During the debrief, Gabby told me that she was unsure how to proceed with these boys because they were so confident that they had sufficient answers. She hoped that by prompting them to attend to the sources, they might recognize perspectives that could challenge their conclusions.

During the debrief, I was surprised that Gabby understood sourcing, contextualizing, and corroborating much better than I expected because these heuristics were mostly absent from the task I observed. She approached the task as an exercise that made space for historical reasoning but did not explicitly require it. Every time she discussed historical reasoning skills, she spoke of "beginning to" or that she "was glad to see it." At the outset, Gabby rarely described actively promoting student thinking beyond the identification of textual evidence, explaining the

connection between quotes and arguments, and facilitation of SAC discourse. Based on Gabby's debrief explanation of the purpose of the task, the omission of historical thinking was intentional.

Interpreting at the outset. I analyzed interpreting according to two features: 1) what the TC reported noticing about student thinking and 2) the factors the TC reported considering for an instructional response. Across the outset data, Gabby noticed whether the students were able to accomplish the target thinking for the task and she noticed that some students engaged in historical thinking. She considered factors related to assessment of student understanding, extending student thinking, and managing tensions between student opinions and student ability to defend an argument with evidence.

Noticing at the outset. In the outset data Gabby tended to notice a number of aspects of classroom life including students being off task and some groups working faster than others. She also noticed that students who initially did not engage in the task as directed were able to accomplish it when prompted. A third aspect Gabby noticed was student historical thinking, although this thinking was not the target goal of the lesson.

As Gabby walked around the room during the SAC, she described noticing that some students were initially "off task" but that all students seemed able to do the reading and competently engage in the evidence gathering activity. She described noticing "on the second lap" around that room that most students were gathering evidence as she expected but "some kids were just writing down quotes" rather than using explanation to connect the quotes to the claim.

Gabby described noticing this incomplete thinking in some of her students but also recognized that when she prompted them, they "knew how to do it." That is, students "knew exactly when …I said write it down in your box as evidence…you need to be able to explain …

how does that show that they are Robber Barons?" Thus, even for students who initially were not executing the task, Gabby recognized that they were able to accomplish it when prompted with a simple question that asked them to connect argument with quotes.

Gabby noticed an unexpected classroom management circumstance that would not be relevant to noticing evidentiary thinking, except for the way that she chose to address the situation. Gabby described noticing a group of boys in the back who rushed through all three steps of the SAC and claimed to have even moved on to the "consensus," that was scheduled for the following day.

During the debrief Gabby admitted that she was surprised that the group was able to get through the activity and did not really know what to do. She described thinking that they "just knew what they were doing…didn't need to be prompted" and that "the documents were too easy for them." Gabby recognized that the boys would need extra attention, attention that she decided to provide by concentrating on their thinking.

In the pretest, Gabby had exhibited an ability to notice student historical thinking when she pointed out the way that Matt and Larissa used evidence to argue, recognized distinctions between accounts, and contextualized the sources that they read. In both cases, however, Gabby also failed to detect important misconceptions present in both examples of student thinking. Given the equivocal pretest results, I was unsure of what to expect when she faced real students' thinking.

Gabby described noticing a good deal of attention to the sources, and particularly to the credibility of the authors of the documents that suggested that industrialists were captains of industry, and that many students recognized that the "rich guys" had a "certain perspective." Conversely, Gabby reported noticing that even for those who sourced the documents for Side A,

few students were able to source the documents for Side B because they were not as familiar with characteristics that could suggest something about the authors' perspectives, what Gabby described as "where the sources are coming from."

That is, Gabby noticed that students recognized authors such as Carnegie and considered his perspective as a way to question author motivation and credibility. But, when it came to authors such as Ida Tarbell, a muckraker journalist opposed to big business interests or data from the U.S. Census Bureau, she recognized that students were less likely to attempt sourcing. Even with the attempts at sourcing she noticed, Gabby reported that there was not a lot of sourcing going on and that student arguments would have been stronger if they could have attended to the source more effectively.

Gabby also noticed that many of the students "brought in" background knowledge from earlier lessons as they discussed the documents. For example, she described how she heard students discuss concepts from the previous day's lessons such as "monopolies" as a way to "relate the documents to the economy at the time." She heard some students also discussing the conditions in factories, from a unit the class had just completed, and how these conditions showed that robber barons "just want to make a ton of money." In the debrief, Gabby correctly described this thinking she detected as contextualization.

When Gabby did not bring up corroboration in the debrief, I specifically asked her whether she noticed any corroboration. She described noticing that students were "comparing evidence from the two positions…looking at different sources and seeing which ones are the best, which ones may have different biases, and which ones may agree." However, Gabby's description of corroboration was more in the general terms of what was required to successfully

complete the task rather than specific examples of student thinking that she observed during the lesson.

At the outset then, Gabby noticed a number of aspects common to group work including students being off task and some groups working faster than others. She also noticed that even students who initially only wrote quotes in the evidence box were able to engage in the target thinking when prompted. Finally, Gabby noticed student thinking about sourcing, contextualization, and corroboration, although this thinking was not the target goal of the lesson.

Factors considered at the outset. Based on her explanations in the early data, Gabby considered a number of key factors that influenced her responses. First, she tended to assign competence to all students when she observed target performance in one student. Second, she saw historical thinking as a way to make the task more complex and, thus, provide enrichment. Third, she calculated how to manage a tension between student opinions and the task's requirement to defend the assigned argument.

When Gabby responded to student thinking she expected that if she could just point them in the right direction, they would be able to accomplish the target task. For example, when she noticed students who had only written quotes on the graphic organizer, she expected that a simple directive would support their ability to explain the link between evidence and argument.

Gabby's expectation that students would be able to resolve problems when they were presented meant that she rarely reported calculating complex instructional responses. During the debrief, she did not describe considering scaffolds to enable student thinking during the lesson or being stumped by student thinking stalemates. The expectation that students could "figure it out" seemed to promote autonomy among the class of highly motivated and interactive students that I observed.

On the other hand, Gabby's high expectations may have been associated with a tendency I noticed to assign competence when evidence for student learning was not necessarily clear. For example, she often spoke in general terms, making such statements as, "They seemed pretty comfortable with the documents [and] knew how to read the documents."

Likewise when Gabby discussed the student historical thinking she observed, she spoke in similar terms saying, "Most of them understood" and "Some of them noticed." She said "They were really thinking about how these people lived," when she discussed contextualization. With the exception of one particular student and the group of boys, Gabby tended to speak of the thinking she observed of the class in generalized terms and not about the learning of particular students or particular SAC groups. As a result of this inclination toward generalizing thinking, Gabby rarely reported the need to explore student thinking more deeply or address misconceptions.

There was one clear exception to Gabby's tendency to avoid deeply engaging student reasoning. With the group of boys who finished quickly, she described not knowing what to do with them. Gabby said at first she "just wanted to make sure that they weren't just like BS-ing it." As she listened to these boys talk through both sides of the evidence, Gabby considered how to make the task more complex for them. She described keeping those boys engaged as an important factor that impacted her response.

A consequence of Gabby's need to engage the boys was that she considered how to make the task more challenging, by inviting them into more complex thinking about the texts and the arguments. Although she noted that students in other groups were engaged in more complex thinking than she expected, Gabby did not describe considering how to push other groups into more complex ways of thinking. In fact, her debrief suggested that historical thinking was not a

goal of her lesson and was, thus, not a primary factor she considered in her responses. Instead, the factors she described were associated with helping students to engage in the pattern of discourse the SAC facilitated and ensuring students could use documents to identify and explain evidence for one side of the argument.

Another factor that Gabby considered in her responses–related to this group of boys but also to other students in the class–was how to manage a tension between the students' opinions and the task's requirement to defend the assigned argument. She described how numerous students were unhappy with the argument that was assigned to them and wanted to change sides. Gabby did not consider allowing students to change sides because the task was based on an ability to "support any position regardless of whether you agree with it."

Gabby did not try to change that fact that students were emotional and their opinions were getting involved. She laughed as she described many students who were "hell-bent on saying that industrialists were bad people." Gabby then recounted how she calculated ways to get students to focus on defending their side with evidence. She said, "I'm like, I know your opinion and that you don't agree with this….but, put that aside … you need to be able to support any position given regardless of whether you agree with it."

Gabby's commitment to ensuring students argued their assigned position demonstrated a factor that she considered in her instructional responses: that students must learn to draw conclusions from the evidence first, rather than their opinions. She suggested that the notion of having a position grounded in evidence was new for most of the students in the class. She described the students as "very opinionated" and having no trouble articulating their viewpoint. But, because Gabby wanted them to learn to examine evidence and base their opinions on evidence, she considered how to encourage them to suspend their initial opinions.

To summarize, Gabby considered a number of factors that contributed to her instructional responses at the outset. She frequently assigned competence to all students when she observed target performance anywhere in the classroom. Second, she approached historical thinking as an enrichment activity. Third she managed a tension students' desire to express their opinions and her goal that they defend an argument with evidence.

Responding at the outset. In analyzing response patterns, I considered how the TC responded to student thinking when students vocalized their thinking about evidence in history. At the outset, Gabby's responses to student thinking about evidence in history tended to focus on her own thinking rather than the students.

In the pretest and early coursework, Gabby envisioned instructional responses to sample student thinking about historical evidence that prioritized inquiry over didactic moves. For example, in the pretest, Gabby envisioned questioning Larissa as a means to help her identify and address misconceptions present in the essay, namely that Larissa ignored contradictions between the two sources. Gabby's inquiry approach, which she said was asking Larissa "to play devil's advocate to her own essay," seemed a realistic way to explore student thinking while promoting perspective taking.

Gabby then planned to ask Larissa a series of questions that seemed more overwhelming than helpful:

What were the costs of the Spanish-American War? What did the U.S. have to gain? What other imperialist activities had the United States done leading up to the Spanish-American War that may have been good for the United States, but bad for the territories they conquered? Were there instances when conquering territory and people wasn't a

good thing in U.S. History, specifically with Manifest Destiny, Native Americans, slavery, and the Mexican-American War?

These questions were intended to help Larissa focus on the side of the argument she ignored in the essay, reasons besides U.S. grandiosity and the sinking of the U.S.S. Maine.

In the decontextualized settings of the pretest and early coursework, Gabby often envisioned responses that explored student thinking. When engaged in teaching enactments at the outset, both real and rehearsed, Gabby tended to take over the reasoning herself rather than explore the students' reasoning. On one occasion in the methods class when Gabby rehearsed discussion moves designed to practicing revoicing ideas, proposing counterarguments, and uptaking ideas, Gabby made it through several minutes of the discussion without using any of the targeted moves that focused on student thinking. At one point, she proposed a counterargument herself, but did not prompt a "mock" student to do it. In the first observation, I often observed Gabby engage in similar Medium and High I responses that put her own reasoning on display. On a few occasions in the lesson, she responded to student thinking with High II responses, or moves that made student reasoning the focal point.

Gabby's most common pattern was Medium responses, or follow-ups that took the form of a response but were not actually responsive to student thinking. For example, when students interpreted cartoons of robber barons and captains of industry, Gabby responded in ways that highlighted clear definitions of 'robber barons' and 'captains of industry.' In the following exchange, a student began by describing the political cartoon that was displayed on the screen:

S3:   It just shows that the Captain of Industry is just trying to give back to the nation to create a utopia of economic revitalization.

G: Ok, so these guys are just handing out money.

146

S2:   The second one is more philanthropic than the first. They are actually giving their own

        money back....

G: What is that word you used?

S2:   Philanthropic.

G: Say it nice and loud.

S2:   Philanthropic.

G: OK, so philanthropy is a nice word. What does that mean?

S2:   It is like how people with lots of money use their money and power to give back to

        society and use it for the greater good.

G: So, what you are saying is that that is what this guy is doing. Ok, anyone else?

Gabby's response interrupted the student thinking in order to focus on the term "philanthropy"

rather than pursue the student's thinking, of which the term seemed only a small part.

        Instead of allowing the student to develop the reasoning further, Gabby finished the

thought and moved on. The exchange had the form of responsiveness, in that Gabby engaged the

student and asked questions but the response actually aborted the student's thinking.

        Similarly, when they got to the robber baron cartoons, Gabby's responses followed what

seemed like a scripted response to a known answer. When she displayed two political cartoons

depicting robber barons, a student said:

S5:   They consider themselves bigger?

G: Ok, bigger than everyone else.

S3:   In the first picture, he is like playing with a building. So, maybe he thinks he can control

        everyone.

G: What city does this look like? What building?

147

SS: The Capitol.

G: He is just playing with the buildings like they are toys. (Pause). So, what we will do next is look at Handout 1.

Although Gabby had several opportunities to explore student thinking in greater depth, her responses did not closely attend to that thinking. Despite the form of responsiveness, the exchanges appeared intended to coopt student thinking to make Gabby's points about the cartoons. At the time I noted that it was difficult to know whether she abandoned student thinking because she needed to move on to the day's central task or because she did not notice the opportunity to develop the thinking further. During the interview, Gabby answered this question. She explained that her intention was simply to deepen student understanding of the two poles of the essential question (robber barons and captains of industry) prior to the SAC, and not to engage in an analysis of a primary source.

Even when the SAC was underway, however, many of Gabby's exchanges were characterized by Medium responses as she modeled the discourse and as she walked around the room checking in with different groups. She talked several groups through the reasoning and although she asked questions, the questions, for the most part, did not push the thinking further than identification of quotes and explanations of those quotes. In such instances, her responses were more formulaic than exploratory.

The second primary pattern of responses that Gabby demonstrated at the outset was High I responses, or responses that displayed her own reasoning on a foundation of student reasoning. For example, during the final discussion, Gabby extended a student's thinking by affirming his interpretation and building on it with her own additions:

S10:   In Document 2, the whole quote just describes how the corporates really thought they were higher than everyone, even the law.

G: Ok, yeah, so they had all the money. They could buy out the government. They gave money to different government organizations as well. So, they could do whatever they wanted. The government needed them. S5 yeah?

Rather than pressing in to the student thinking, Gabby extended the student idea by reasoning about it herself. This extension, however, was not particularly helpful because it was not explicitly substantiated by evidence and was almost a hyperbolization of the second document.

I only noticed Gabby use High II responses, or follow-ups that made student reasoning the focal point, a few times during the lesson. In one of the exchanges she had with the group of boys, Gabbie pushed the students to provide evidence and revoice the sourcing that the student had demonstrated.

G: What is your evidence?

S1:   Carnegie is saying that the rich are good for society. But, really this shows that he is just protecting himself by… (inaudible).

G: So, you are turning that evidence on its head and using it against him?

When I later asked Gabby about the exchange, she noted that the student "actually managed to turn…Position A around," in that he delegitimized Carnegie's reliability as a source. The students in the group did not seem to notice Gabby's attempt to highlight S1's thinking about how the perspective of Carnegie influenced interpretation of the document. When students ignored her comment, she did not press it further.

Immediately after, another student in the group explained that they had already completed the entire SAC process and had arrived at a consensus: "We need a bigger middle

class." Gabby pushed back by asking them to return to the documents, and then she returned to a High I response:

G: Lets talk about captains of industry.

S1:   But, we all agree.

G: Who do you think that guy is?

S: A rich guy.

S3:   (inaudible)

G: So, in evidence box write support...and maybe talk about the sources too. What do you think the sources are? Where are their biases?

Gabby's response was an attempt to build on the student thinking and direct the boys past their own certainty and into new questions based on the sourcing information. But, she had to wrestle the line of reasoning away from the students in order to make the idea public. The response could have initiated student reasoning that went beyond the straightforward interpretation of the documents that the group had settled for, but the idea was Gabby's and the reasoning seemed to flounder without Gabby's presence.

Admittedly, analysis of Gabby's responses in the outset lesson was difficult because of the structure of the activity. Most of the discourse about evidence was limited to small group conversations and when Gabby brought the discourse back to the whole group, the class time was running short. However, of the responses I was able to observe, there were only a few in which Gabby responded to student reasoning by making that thinking the focal point of the discourse. Most of the discourse I observed was characterized by Medium responses and High I responses.

Change in eliciting. Gabby continued to design tasks that had generative potential and she increasingly launched these tasks in ways that strategically built on the content and skills she observed in her students. Gabby intentionally elicited historical thinking over the course of the study and rarely explicitly taught these as historical thinking skills.

Change in generativity. Similar to the outset data, Gabby's later lessons were filled with generative possibility. These tasks positioned students to reason about evidence with essential questions that created historical tensions resolvable by interpretation of documents and argument building.

For example, I had the chance to observe students participate in another SAC later in the year. Gabby used the essential question "Was the prosperity of the 1920s real or artificial?" She chose documents that included primary sources from W. E. B. Dubois, Stuart Chase, President Hoover, and others. Also included were graphs and data that demonstrated both GDP and real family income from the 1920s.

Gabby's clear instructions and clarity about what counted as evidence made the launch of the task almost seamless, as compared to the previous SAC I observed. She started with a nod to the purpose of the SAC, to promote discourse that results in argumentation and eventual consensus.

G: Ok, what is a consensus guys?

SS: Agreement …discussion

S2:   An agreement between two or more parties.

G: What is it?

SS: Agreement …

G: Ok, it is an agreement. So, you will be discussing two points of an argument and you will

         be coming to an agreement.

The SAC, as Gabby implemented it, offered a generative opportunity for students to engage in

the type of discourse that historians use. Rather than debate or argue about unsupported opinions,

her design intended that students would take both sides of the argument seriously and use

evidence to argue their way to consensus. In her instructions, Gabby referred several times to the

SAC on robber barons and captains of industry and the discourse patterns that she introduced in

that lesson.

Gabby's instructions directly addressed many of the problems that I observed in the first

SAC. When she gave instructions for the graphic organizer, her clarity on the nature of evidence

and the necessity of explanation that would turn that quote into evidence was a clear change from

the first SAC on industrialists: This time, she said, "Using your document analysis chart, list

three pieces of evidence you will use to present your position….So, evidence will be, in this

case, a quote and an explanation of that quote."

When it was time for the second step of the SAC, listening to the counterargument and

repeating back the argument, Gabby modeled the process by having one student present her

argument and having a second student try to restate these arguments. When the second student

was unable to restate the argument, Gabby smiled and asked the first student, "S14, would you

like to restate what you just shared?" She continued this modeling until a student from Team B

was able to restate the argument to S14's satisfaction.

Gabby's launch of the 1920s SAC positioned students to reason about evidence and

develop arguments grounded in the interpretation of evidence. Furthermore, the task required

students to use content and skills that Gabby had been developing in previous weeks, as she explained in the debrief.

On another day, I observed students preparing for a document-based discussion on US imperialism that would occur the following day. Gabby felt that for students to understand the primary sources she would provide for the discussion on imperialism, they needed stronger foundational knowledge about context and specific vocabulary. While many teachers might have envisioned this necessity as a lecture, Gabby told me in the debrief:

> I could have said, okay, "Well, these are the five reasons [for US imperialism], like copy them down." I wanted them to find it by looking at a document from that time from a man who's very famous for his ideas during that time...So rather than handing it to them I wanted them to hand it to me by looking through and figuring out what this guy is saying about why the United States should become an imperial power

The task I observed required students to read a speech by Albert Beverage, a 19th century advocate of U.S. expansion, and "identify reasons for US imperialism." After modeling interpretation of a Beveridge quote that Gabby displayed on the power point, she gave students time to read through the selected portions of the speech on their own. Then, she led the students in a discussion that began:

G: … the reasons that Beveridge says the US should imperialize. Ok, so S1 can you start us off?

S1:   Right off the bat he is talking about how America is a God-given land (quotes from document). God decided that America needs to be better than everyone else.

Students offered multiple reasons from the Beveridge document that demonstrated why the US should imperialize.

153

For each reason, Gabby articulated attitudes of the time. After discussing Beveridge's reasons for imperialism, she introduced students to generalized motivations for imperialism that she called the "Izes:" civilize, colonize, Christianize, democratize, and 'economic-ize.' With each of the "*Izes*," Gabby stopped and asked students if they could identify a selection from Beveridge's speech that exemplified that motivation for imperialism. In this way, she took an objective that could have been presented as information for student consumption, and she turned it into an opportunity for students to reason about evidence and develop arguments based on evidence.

I observed another example of Gabby's eliciting when I visited on the first day of the 5-day unit on the prosperity of the 1920s. (The SAC mentioned previously was a part of this unit). The culminating activity of this unit built around the question "Was the prosperity of the 1920s real or artificial?" was an essay in which students defended an answer to the essential question by using evidence from the dozen or so primary sources that Gabby provided over the course of the unit.

On this first day of the unit Gabby launched a lesson built around a reading from *The Great Gatsby* and pulled on popular images that idealized the "Roaring Twenties." The lesson was another example of an opportunity that she took to ground students into the complexities of historical interpretation. Rather than simply provide students with an authoritative narrative about the 1920s, Gabby used fiction to conjure popular images of the decade and repeatedly returned to the unit's essential question.

Although students did not engage with primary source evidence in the lesson (except for a quote from President Coolidge) or build arguments using historical evidence, Gabby prepared them for the following day in which they would examine primary sources that depicted

sharecropping, the lives of migrant workers, the convict lease system, and the farm crisis of the 1920s. The lesson, when viewed on its own did not meet my standards for generativity because the students were not analyzing *The Great Gatsby* as a primary or secondary source. However, when viewed in the scope of the broader unit, I recognized that Gabby was positioning the students for a highly generative task.

Changes thinking eliciting. Throughout the time I observed Gabby teach, she continued to focus on justification but became more intentional about historical reasoning. Over the course of the study I observed her intentionally elicit thinking about sourcing, contextualization, and corroboration. However, I rarely observed her approach these thinking skills in an explicit way with her students.

During the 1920s prosperity SAC, most of Gabby's eliciting focused on justification. She emphasized identification of evidence, as a part of this, explanation of how selected quotes related to the claim students made in their arguments. When Gabby launched the SAC, she reminded students how the process worked and pointed them to the graphic organizer, which instructed students to "Put both evidence and document number in your position's boxes - Use at least one quote [in each box] and explain."

By reminding students that evidence would be "a quote and explanation," she emphasized her definition of what counted as evidence. Evidence could not simply be a quote. A quote became evidence when it was linked to a claim by way of explanation. As Gabby walked around the room she asked students "What evidence have you found?" By this, she meant, "What evidence have you found to support your side?" This focus on general argument building skills was common across all of the lessons I observed Gabby teach. But, her eliciting was not limited to justification alone.

Gabby also elicited student thinking about the sources, but not always sourcing, exactly. For example, when I observed her introduce the 1920s prosperity unit, she established expectations for a student essay that suggested attention to the sources. These required students to "use three documents and cite the name of the document, the author, and the date it was written." In one interview, Gabby told me that she told the students to "introduce quotes." She related how she provided students with sentence starters designed to help them avoid, "hanging quotes" and joked that "if you showed up to a party with a new friend, the first thing you would do is introduce him. It is the same thing with quotes, we have to introduce them before we can use them."

Of course, citing source information is not necessarily sourcing because it does not require students to consider the source as a means to make sense of the document. Had Gabby extended the introduction analogy and said, "Before we would go introducing this person around as our friend, we would probably like to find out what kind of person he is," that could have been a more apt means to connect with reliability, the central issue in sourcing.

I did not see Gabby explicitly teach sourcing. Instead she seemed to just provide students a problem space that invited sourcing and some direction as to why sourcing was important by providing feedback such as, "Why should I care what this person has to say?" Despite the lack of explicit attention, some students actually began sourcing documents anyway.

Gabby later explained that she did not focus on sourcing until she graded the 1920s unit essay and noticed how rarely students were attending to the reliability of sources. She told me that once she provided specific written feedback about sourcing and required rewrites of some of the essays, she began to see students attending to the source information to make their arguments.

Based on her concern, Gabby told me that she decided to make sourcing a central component of the final two units of the year, an investigation of Japanese internment experiences and an examination of accounts about the dropping of the atomic bomb on Japan. Gabby told students that she was "going to trick them with some documents that might not be as good as others." For the atomic bomb lesson, she explained that she would provide students with documents of various levels of reliability and expect them to attend to the sourcing information in order to decide what to use for their arguments:

> I had one that was Secretary of War Henry Stimpson…an op-ed newspaper
>
> editor…versus a soldier, or a student who survived Hiroshima… And, I had a section on
>
> their graphic organizers that they had to fill out about why they thought that this was a
>
> reliable source and where was the bias.

Gabby explained that the students "got super into it because they were like, 'which one is she tricking us on'?" And, she further described how it was a nice addition for a debate she facilitated at the culmination of the atomic bomb unit because students "had a chance to question each other … like, 'You are assuming that this newspaper guy can be taken at his word. Why does he have the last say on what is going on'?"

By the last debrief, Gabby spoke openly of neglecting sourcing for most of the year. But, even when she described prioritizing sourcing, she apparently did not teach it as a distinct type of disciplinary thinking. Rather, she emphasized its importance, created a problem space, and encouraged the students to figure it out.

Gabby elicited contextualization more frequently than any other historical thinking during the lessons I observed. Indeed, she acknowledged in the debrief after the 1920s SAC that she prioritized contextualization because she believed contextualization was the hardest of the

157

historical thinking heuristics. As one example of this focus, Gabby wanted students "to be able to contextualize the Imperialist Age of the late 19th century/early 20th century." She selected the Beveridge speech and included contextual notes below the text of the speech:

> Beveridge was US Senator from Indiana (1899-1911), and, as is evident here, a fervent supporter of American imperialism. He gave this speech as a campaign speech on September 16, 1898. He was also a prominent historian of his time, and …and was highly influential in social reform in the early 20th century.

Even though Gabby selected the speech herself and included the information above as part of the design, she did not explicitly point the students to this information during the launch of the task.

When Gabby launched the 1920s prosperity unit with *The Great Gatsby* and popular images that idealized the decade, she established a strong contextual basis for the unit. She introduced students to the context of the time by discussing aspects of apparent prosperity. As she pointed out new products and technology, Gabby said "You have hair dryers, you have toasters, you have electric stoves and refrigerators…things that you have to have that you don't know how you lived without before...all being sold with radio advertising….Before it was just in catalogues but now it is on the radio." Gabby referenced the expansion of the economy, lowering of regulations, new products and technology, buying on credit, and the increasing wealth gap that made the 1920s so revolutionary. Where I saw all of this as grooming the students for contextualization they would do in the coming days (contextualization I later observed), Gabby did not use the term "contextualization" or refer to a specific type of thinking required to frame historical context during the lesson.

Many of the ideas and attitudes Gabby raised in the lesson would be keys to understanding the context of the sources that the students examined in the following day's SAC

because the background information helped them recognize what was going on at the time the documents were written. But, despite all the contextual thinking Gabby was eliciting, I never heard her use the term "contextualization" or refer to it as a specific skill in any lesson. I never observed her explicitly model contextualization by saying something like, "Notice the way I am thinking right now as I use the background information about the farm crisis to interpret the author's point."

Gabby approached corroboration in the same veiled manner. During my visits, I observed students regularly consider multiple pieces of evidence at once, question different versions of the story, and sometimes even consider which sources were more believable than others. One of Gabby's primary objectives for the SAC on prosperity in the 1920s was "Students will corroborate between documents to compare and contrast different arguments on the same topic." During the lesson I observed students considering each piece of evidence and beginning to identify distinctions and similarities between the documents.

During the debrief, Gabby casually showed me a graphic organizer from a previous day's lesson. The graphic organizer walked students through historical thinking for each of the documents they used to examine poverty in the 1920s:

- Who is the author? What is their perspective?

- When, where, and why was it written?

- Is this source reliable? Why or why not?

- What does this document say about what life was like in the 1920s?

When I saw this graphic organizer, I recognized that Gabby indeed was eliciting thinking about historical reasoning. In what I observed, however, this thinking was rarely scaffolded and rarely

even explicit. Gabby seemed to assume that the students would be able to think historically if she asked them to and gave them the opportunity.

In the beginning, Gabby had not expected to see any historical thinking beyond justifying arguments with quotes. But, by the end of the internship, she expected students to exhibit historical thinking regularly in their discourse and writing. Yet, Gabby seemed to avoid explicit instruction with these skills throughout the course of the study.

Changes in Interpreting. Throughout the remaining observations, Gabby's patterns of interpretation changed in some ways and held in other ways. She continued to notice aspects of student thinking and her attention increasingly focused on more complex thinking. However, she maintained a tendency to assign competence when the evidence for student learning was not necessarily clear. She was influenced by a desire to promote contextualization and particularly considered the tension between student judgments as historical understandings.

Changes in noticing. As the study proceeded, some of the aspects of student thinking that Gabby noticed remained the same and some things appeared to change. In the later data, she described noticing students' ability to engage in the assigned tasks, noticing contextual thinking, and noticing student ethical ideas that made contextual thinking more complex.

In the later data, Gabby continued to notice student historical thinking, especially contextualization. For example, after the lesson on imperialism, she said, "They nailed contextualization." She went on to describe examples of comments she heard that connected Beveridge's ideas to content the class had previously studied. In one of these descriptions she noted how students "tied it back to the Native Americans, which [in terms of the curriculum] is finished but is still happening during this time."

160

Gabby described noticing that students were making the connection that "industrialization and imperialism were happening simultaneously" and that industrialization, imperialism, and westernization were all intertwined. For Gabby, this indicated that students saw Beveridge's imperialistic attitudes as contiguous with time periods that she had previously taught. She noticed that they were taking new ideas and facts about imperialism and framing these within their understanding of the time period. In some cases, Gabby noticed that the students understanding of the context surpassed what she introduced in class, a fact she attributed to their being "well-informed."

Similarly, in the 1920s SAC, Gabby noticed students bringing in earlier contextual information and outside content knowledge to understand the documents and the time. In one of these references, she described how she noticed that "S4 was bringing back what he remembered about industrialism," to make sense of a passage in the Dubois text. In other student conversations, she noticed consideration of widespread poverty and the depression looming, aspects of the context not mentioned in any of the task's texts.

Related to Gabby's noticing of contextualization was that she continued to notice that students had emotional reactions and strong opinions. In the imperialism lesson in particular, she conveyed negative reactions she heard as they engaged with the ideas and attitudes of the time. For example, in the Beveridge lesson, she heard students say things such as "that is crazy" and "that's kind of scary." Gabby also related that other recent lessons, such as Indian removal, evoked feelings that many of these ideas were "wrong."

Where in the robber barons and captains of industry lesson, Gabby seemed to just notice students' turning to their opinions to bolster arguments, in the later lessons she noticed what she called "the elephant in the room." By this she meant that students believed the ideas of

161

Beveridge and others "were unacceptable" and they were simply dismissing the prominent ideas of the time as evil or stupid. She struggled with how to approach contextualization when a historical action or actor was clearly wrong, by modern standards. Gabby said, "it is not just that it's wrong but that it happened for all of these reasons."

The later data indicates that Gabby noticed key aspects in student thinking but she also failed to notice other aspects. For example, where she noticed some aspects of sourcing in the outset data, she neglected sourcing for an extended period during the later data. Granted, she always provided sourcing information on the documents. But, when I asked her about sourcing in both the Imperialism lesson and 1920s SAC, she said that she did not notice students sourcing during the class period. Gabby's did not notice her students' penchant for describing documents as class resources rather than a particular type of historical source with a particular author. For example, in several exchanges during the 1920s lesson, Gabby failed to notice that students were referring to sources as "Document A" and "Document B," rather than calling it "the Carnegie article" or "the Tarbell quote."

Gabby actually critiqued her failure to notice sourcing weeks later when she admitted that she did not actually start paying attention to student sourcing until she noticed a lack of attention to the sources in student writing, well into the 3rd quarter of the school year. Given these points, it seemed that she was attending to many aspects of student thinking but that attention was often constrained to certain aspects, especially to contextualization.

Gabby also continued to notice students' ability to successfully engage in the tasks she provided them. In particular, she noticed their high levels of reading comprehension, their abilities to accurately interpret difficult documents, and their capacity to identify an author's argument. For example, she noted after the Imperialism lesson that the students "did a fantastic

job"…looking through [the document] and figuring out what this guy is saying about why the U.S. should be an imperial power."

In some cases, Gabby noticed that students were successfully engaging in the activity, but in the debrief, could only vaguely explain what it was that she noticed. For example, in the 1920s SAC, she described students "pulling good quotes." But, when I pressed her on what a "good quote" was, she replied that it was "a quote that I would have pulled." Similarly, Gabby struggled to articulate exactly what she noticed when she said that students were "accurately interpreting the documents."

Altogether, it seems that Gabby's noticing remained constant in some ways but changed in others. She continued to notice student thinking but appeared to notice that thinking at a more complex level. She continued to notice student opinions as a factor in their arguments but recognized a new and difficult element in those opinions. Finally, she noticed whether students were successfully accomplishing the target thinking, but only in nebulous terms.

Changes in factors considered. The factors that Gabby initially reported considering changed in some ways and held in other ways. In the remaining lessons, she described a clear framework for assessing student thinking but continued to assign competence to the entire class based on the thinking of a few students. She was influenced by a desire to promote contextualization and particularly considered the tension between student judgments and historical understanding.

Gabby reported noticing many things typical for a new novice teacher, such as time management and student participation. But, most of her descriptions in the debriefs suggested that many factors impacting her instructional responses were related to the substance of student thinking. In the 1920s SAC debrief, Gabby described the way she evaluated student arguments.

She determined that an argument was "good" if the student was pulling, what Gabby called, "good quotes," or "quotes I would have pulled." Then she considered if they were "interpreting the sources accurately and … building off of it by incorporating things like sourcing and contextualization." This description suggests that she was engaged in a significant amount of formative assessment that could impact instructional responses.

However, based on what I observed in the debriefs, Gabby continued to attribute the understanding of a few students to the entire class. For example, in the imperialism lesson debrief, she frequently referenced noticing that students understood something that was based on the comments of a single student. Even when Gabby described her supervisor's observation that the closure of the imperialism lesson only assessed the thinking of two students, she did not appear to view this as a tendency in other parts of her teaching.

This trend continued in the 1920s SAC as Gabby walked around the room and checked in with different groups. She described groups as "being on track" and "accurately interpreting" but regularly attributed these understandings to the group as a whole rather than notice the specific thinking of individual students. Nonetheless, Gabby did pay attention to student thinking.

Where historical thinking beyond justification was rarely a factor at the outset, it was a factor that Gabby considered for instructional responses in the later data. In particular, she described focusing on helping recognize the continuity of ideas and the overlapping time periods in the curriculum. For example, in the lesson on imperialism, Gabby said that she was aware that she sounded like a "broken record" as she hammered the continuity link of imperialism with other periods of history that the class previously studied, especially the industrial revolution. Gabby mentioned one motivating factor was that she hated the way her high school history classes had lacked links from one unit to another. She remarked, "I did not want them to view

imperialism and industrial revolution as separate." This is how she described the motivation of many of her instructional responses in both the imperialism lesson and the 1920s SAC. She said, "I wanted them to see that all of this was happening simultaneously."

Gabby described her questions and comments throughout the imperialism lesson as intended to link student prior knowledge to the task's content in order to demonstrate the continuity of ideas in Beveridge's speech. Likewise, in the 1920s SAC debrief, Gabby repeatedly described her responses as intending to remind them to place the content and sources into the historical frame the class had already built. Although historical thinking was an important factor in her calculations, Gabby rarely raised sourcing or corroboration as explicit factors that she considered in instructional responses, until the later part of the internship.

Even as Gabby promoted contextual thinking, she wrestled with how to address "the elephant in the room." She weighed how to support students in contextualizing the attitudes of the time while allowing their judgments of unethical ideas and behavior to remain in place. In the debriefs, she described a tension between wanting students to have an emotional response to the study of history and getting students to understand historical ideas and actors in their own time period. She wanted them to understand the "reasons that these things happened," while also allowing them to recognize that many of these things were ethically "wrong."

This tension did not emerge publically during the classes I observed for Gabby but it was clearly something that was on her mind, as she described after the imperialism lesson. I did not observe her resolve this tension but did discuss it with her in subsequent interviews. She did, however, describe working patiently to support students in contextualizing historical actors and their actions while also encouraging students' sense of justice.

Changes in Response. Over the course of the remainder of the study, Gabby increasingly

utilized High I responses to put her own reasoning on display in response to student ideas and, as

in the outset, rarely engaged in High II responses.

During the SAC on the 1920s I was again faced with the challenge of trying to observe

Gabby's responding when much of the discourse was happening in small groups. As Gabby

made her way around the room, she listened to student discourse and joined the conversation at

certain points, with mostly High I responses that built her own reasoning on student reasoning.

For example, when one pair was working through the Dubois document:

S7:   "Gambling," they did actually mention it in the document. But I was not really sure what

        part of that quote to pull.

G: So what does he mean by gambling? Does he mean like playing poker and rolling dice?

        What gambling is he referring to?

S7:   (Pause. S7 read the Dubois quote again.) Like monopolizing things to increase its value?

G: Yeah, so he is talking about industrialists and bankers gambling. So, how are they

        gambling? They are gambling with stocks. They are speculating. They're making bad

        investments. So, that's what he means by gambling.

In the exchange above, Gabby began with the student idea and tried to help him develop it.

When his answer was unclear, Gabby took over the thinking and brought in contextual

knowledge for him (High I).

Likewise, in the lesson on imperialism, most of Gabby's responses were either High 1 or

Medium. She frequently repeated students' ideas and expanded on that thinking. When Gabby

modeled reading Beveridge and thinking with the whole class, she started by eliciting students'

interpretation of a section from the speech that included, "the trade of the world must and shall

be ours." When a student pointed out that Beveridge believed that the US was fated to control world trade, Gabby responded by repeating the student statement and asking:

G: What does that sound like, what vocab term does that sound like that we just learned about?

S10:   Like sphere of control.

S16:   Sphere of influence.

G: Sphere of influence? Yeah. Why would he say we need to control trade?

S10:   So we can consume the resources.

At this point, Gabby shifted the conversation to reason about the context of the time period in which Beveridge was writing:

G: Ok, so, American soil is producing more than Americans can consume. Remember, we have the industrial revolution going on right now. We are producing at unprecedented rates. We are making steel, we are making…What are we going to do with all the stuff because we don't have enough Americans to buy it?

S11:   Ship it away.

G: …We need to control trade so that people will buy our products that our country is producing. So, what I want you to do is…Take a look at why Beveridge thinks that the US should imperialize.

This is an example of one of Gabby's High I responses. Her initial response was to identify "sphere of influence" as a point way to frame the student's idea. Once a student answered her question about beverage's motivations, she completely departed from the student thinking and contextualized herself, connecting Beveridge's ideas to contextual ideas of the industrial revolution and vocabulary that students had recently studied. Had she been explicit that

she was contextualizing and then returned to the original student statement after stabilizing the context, it could have prompted students to use these contextual details to reconsider Beveridge's speech. But, instead of returning to the student thinking, Gabby moved directly to the independent reading portion of the activity.

The exchange above also demonstrated a High I response pattern that Gabby used in several lessons to put her own contextual reasoning on display. A common way that Gabby built on student thinking in the later data was by personifying that thinking in the form of a present tense mock quotation.

When I first saw her use present tense and personal pronouns to speak in the language of an earlier time, I was worried that Gabby had a major misconception about history. However, in the debriefs, Gabby made clear that these were intentional attempts to bring students into the attitudes of the time, attempts that she believed the students recognized as contextual personification and not as Gabby's ideas or actual contemporary attitudes. The following High I response from the discussion on Beveridge and imperialism was indicative of this pattern:

S5:   He thinks only certain people are able to be self-governed. They are not strong enough to govern themselves.

G: … He says that we have a responsibility to govern those who can't govern themselves.

S6:   He says that we are already governing people without their consent like Indians. So, it is not that immoral.

G: Yeah, we are already doing it. How did we get America and gain the territory we did? We have land from the Native Americans, from the Spanish… Why is it a big deal?

Gabby initially responded to S5's thinking by simply highlighting the observation that the student made. But, after S6 extended that thinking, Gabby moved into personifying the attitude of the time.

The lesson on imperialism, in particular was marked by a pattern of discourse in which a student would present one of Beveridge's reasons for imperialism and Gabby would follow it up with several impromptu statements personifying the attitude of the age. In each case, after making the personified statement, Gabby immediately moved on to elicit another of Beveridge's reasons rather than return to the student's original idea. Again, had Gabby explicitly stated that she was stabilizing the context and then returned to the student thinking that prompted the response, this move could have kept student thinking as the focal point of the exchange.

These High I responses that extended student thinking by personifying contextual notions were not limited to the imperialism lesson. In the lesson that introduced the unit on prosperity in the 1920s, Gabby again responded to student statements by taking student ideas and personifying those ideas in a first person voice of the time.

For example, after several students described their interpretations of a quote from Calvin Coolidge, a final student said, "We love business." And, Gabby responded, "We love business. We love money. This is what we are concerned about, prospering in the world." In this way, Gabby used first person personification responses as a tool to develop contextual thinking that students would use throughout the unit on the 1920s. I never observed Gabby use these responses as a means to stabilize the context and immediately return to the original student idea that prompted the response, as would be required of a High II response.

To sum up Gabby's response patterns in the later data, she increasingly utilized High I responses to put her own reasoning on display in response to a student ideas and, as in the outset,

rarely engaged in High II responses. She also sometimes used non-response when student thinking was made public but her rationale for doing so was not clear to me.

Summary of Findings

In this chapter, I have considered how Gabby elicited, interpreted, and responded to student thinking about historical evidence. My first research question asked how Gabby engaged in these practices at the outset of the program. Gabby initially struggled to launch a generative lesson, elicited thinking about justification, and only elicited thinking about sourcing when she determined a need for enrichment. Although it seemed that Gabby was not focused on historical thinking in the lesson, in the debrief she actually described aspects of historical reasoning that she noticed. She juggled student enrichment and student opinions as she weighed how to respond but also tended to assign competence to the entire class based on the thinking of one student. Her outset responses were characterized mostly by Medium and High I responses.

My second research question asked how Gabby's capabilities to engage in these practices changed during the study. In the later lessons, Gabby regularly launched generative tasks and became more intentional about eliciting historical thinking. She continued to notice student historical thinking and weighed new challenges of contextualization. Gabby continued to attribute the thinking of a few students to the understanding of the entire class. In the later lessons, Gabby increasingly used High I responses to put her own reasoning on display.

Chapter 7: Kendra

Kendra was a 25-year-old white female who graduated from a private high school prior to finishing undergraduate as an honors government/politics and philosophy major at our university. She spent two years in corporate recruiting before deciding to enter the Master's program. Her fiancé was a local high school history teacher, a graduate of the same master's program in which she was enrolled. Kendra became pregnant early in the internship and managed the roller coaster of coursework demands, internship expectations, and significant life change with a wry smile and positive outlook.

Kendra described herself as an introvert, a trait I noticed when she was in unfamiliar settings with people she did not know. From the first day of my summer methods course, Kendra exhibited tremendous intellectual confidence as well as a curious combativeness. Despite sometimes feeling annoyed, I tried to encourage her questioning and, given the small class size, I was able to entertain many of the challenges she posed. Most of her frustration involved historical reasoning and historical concepts that Kendra did not initially understand.

At various points, Kendra clearly stated her frustration that the course was focused on historical thinking at the exclusion of other social studies subjects. During one class that summer, Kendra suggested that historical heuristics might not actually be important. Annoyed that we were plowing ground already established, I tersely responded that it was fine if she wanted to focus on facts but she should not call it 'history.'

This exchange resulted in a meeting after class. Through her tears, Kendra explained that the concepts were all new to her and, although she was working hard, she felt confused by the content and had no confidence that she would ever be able to teach it. This conversation, though

unsettling for me at the time, proved a turning point both in my relationship with Kendra and in her participation as a TC in the course.

Over a four-month period (June-September), I observed Kendra become increasingly comfortable describing historical thinking and the value of attending to student thinking when teaching. Kendra's pretest demonstrated these developing abilities in some aspects of historical reasoning. She built arguments primarily around sourcing information and attended little to matters of contextualization and corroboration. Her overall score for historical thinking heuristics, however, was in the middle range of the class.

Kendra seemed to have a strong sense of what good teaching looked like and her reflections suggested that she admired teachers' classrooms when students engaged in challenging tasks that the students valued. She seemed to know good teaching when she saw it but was not usually able to articulate what the teacher was doing that made the teaching particularly good.

Kendra's critical eye combined with a lack of experience meant that she tended to be hard on her own attempts at teaching. Rehearsals and early enactments left her frustrated that she could not accomplish what she wanted in the classroom. For example, Kendra's rehearsals in the methods courses frequently demonstrated a commitment to include historical reasoning and a developing, but often directionless, capacity to focus on student reasoning when teaching. She explicitly sought critical feedback and expressed her frustration at what she perceived as missing the mark in her own teaching.

Context

Almost from the first day of her internship in a twelfth grade government classroom, Kendra was disappointed with her mentor's teaching. Kendra's concern was partly the teacher's

172

single-minded focus on test preparation. Kendra worried about the "lack of norms for independent primary source analysis" and expected it to be "very challenging to help students analyze sources" because the mentor never engaged in such activities. What Kendra wanted most in a mentor was a great role model, guidance, and frequent feedback. What she got was a person whom she admitted not respecting as a person or a teacher.

In one of her first classroom enactments, Kendra tried an approach that was different from her mentor's instruction. She reported that the students were unwilling to participate in the lesson and the mentor's only feedback was "I told you so." Kendra struggled in this placement week after week and eventually, Meredith, the supervisor, and Kendra mutually decided to place her to another school.

In mid-October, Kendra moved to an 8th grade history classroom and found a mentor she respected and admired. Her new placement was a large middle school with a high rate of students who qualified for free and reduced meals (89%). The student population was mostly Hispanic (67%) and black (28%). Almost all of the students received Title I funds and 30% of the school's students qualified as Limited English proficiency.

Despite similar student demographics, Kendra's new placement was very different than her original one. She stepped into a situation in which the students were already normed into active participation, were working with primary sources, and were using tools that promoted historical reasoning. After the change of placement, Kendra had nothing but positive things to say about her mentor, her school, and her students.

Kendra's new mentor had participated in a research study several years prior that provided professional development for teaching historical reasoning. The new mentor emphasized historical thinking skills, reading primary source documents, and argumentative

historical writing. In fact, some of the mentor's materials were the very resources we introduced in the methods course and his preferred activity structures were some of the methods course IAs. Thus, Kendra had more support in the development of historical thinking than only what was provided in the summer and fall methods courses. In addition to on-level and honors course US History I classes, Kendra taught inclusion and ESOL sections of the course. The class I regularly observed was the inclusion section.

Eliciting, Interpreting, and Responding

In the following section, I describe Kendra's ability to EIR over the course of the study. The data that informed these findings include information drawn from the pretest Kendra took prior to coursework, course participation in the following months, four classroom observations and seven interviews. My first observation of Kendra's classroom teaching occurred in late October 2014. My final observation was in mid-March 2015.

Eliciting at the outset. As described in Chapter 3, I analyzed eliciting according to two features: 1) generativity of the task the TC launched and 2) the specific type of historical thinking the TC elicited. At the outset, Kendra elicited an array of student thinking about evidence in history but was unable to launch and maintain tasks that resulted in cogent lines of student thinking.

Outset generativity. Initially, Kendra had difficulty launching tasks that elicited generative student thinking. The first time I observed Kendra was only a few weeks into her new placement. Her mentor decided to "throw her in" and let her teach a 3-day investigation about "Who shot first at Lexington Green?" Even though Kendra was excited to teach a lesson that focused on historical thinking, her pretest, early coursework, and prior internship experience suggested that Kendra might not have been ready to manage the student thinking on her own.

In her pretest, Kendra envisioned a task in which students read and discussed primary sources that represented "different biases" on a particular concept. At the conclusion of the lesson, students would write a short exit card with the perspective they found most convincing. Although Kendra called this cognitive work "thinking historically," she did not describe any tools for evaluation beyond the students' opinions.

Although Kendra designed several tasks in her early coursework that focused on historical thinking, the problems of her first placement prevented her from enacting these in a classroom with students. Accordingly, the Lexington Green lesson was one of the first times she had ever taught a history class to actual students.

In the Lexington Green lesson, Kendra's resources and a task structure were potentially generative. Materials positioned students as "historical detective(s) trying to uncover who fired the first shot." On the prior day, Kendra used a reading scaffold designed to help students focus on the author's argument, the reliability of the sources, and the timeline of events, as described in each source.

On the day I observed, Kendra used a structured academic controversy IA to promote argument and interpretation of the essential question. Because the design of the task seemed generative, I was excited to see Kendra's students engage in the investigation. Within the first few minutes, however, it was clear that both Kendra and the students would struggle.

The problem was not the task's generativity. Throughout the lesson, student thinking repeatedly materialized around the room, sometimes in rapid succession. For instance, a student seated in the front of the class pointed to a quote as evidence of the author's argument. A student in the back wondered aloud about how afraid the colonists must have been and another student explained why the diary of the British soldier was reliable. Kendra's task required that students

175

select quotes, evaluate evidence, and explain arguments. But, Kendra launched this task in a way that created confusion rather than cogent lines of student reasoning. Consequently, student reasoning fizzled out almost as quickly as it emerged.

Confusion began with the introduction of a two-sided graphic organizer intended to scaffold the structured academic controversy discussion. When Kendra explained the graphic organizer–which had side-by-side columns labeled "evidence" and "explanation"– she told the students that in the first column they should "identify quotes that supported the argument." In the second column, students were to "try to explain how this quote explains why the colonists might have shot first."

Although it seemed that quotes counted as evidence, I did not understand what counted as explanation. In the original modeling of the graphic organizer, however, explanation appeared to be a way to connect a quote to one argument or the other. As the lesson went on, explanation seemed to morph into something else. When Kendra tried to demonstrate the explanation portion of the graphic organizer for a second time, she asked "Why is this the strongest evidence?"

After one student provided a quote as evidence that the British shot first, Kendra then asked him, "Why does him saying that make it a strong piece of evidence? He could have said anything. What specifically about what he said or where he said it makes it reliable?" Such questions made me wonder whether Kendra actually knew what student thinking she wanted to elicit because these questions seemed to be exploring reliability rather than inviting a warrant to connect a quote to a claim

Each time a student articulated what was written on their graphic organizer, they provided evidence but were unable to meet Kendra's expectations for what went in the baffling

second box of the chart. Not understanding what Kendra meant by explanation, students tended to address the second box with paraphrases of the documents' content.

Although Kendra entered the lesson with a task that could have created a generative problem space, her instructions and eliciting of student thinking began a spiral of confusion. Kendra elicited a lot of student thinking about evidence in history but that thinking was so disorganized that the reasoning materialized and faltered over and over again.

Outset thinking elicited. During the Lexington Green lesson, Kendra elicited thinking about justification and sourcing. She elicited thinking about justification by asking students to identify a piece of text that supported a side of the argument. For example, she frequently asked, "Who can provide a piece of evidence that the British shot first?" More commonly, she asked "What is the best evidence?" By this Kendra intended that someone would offer a quote from one of the documents that suggested that one side or the other shot first.

Second, Kendra elicited thinking about source reliability. After student contributed a quote that justified one of the claims, Kendra often asked, "Why is this piece of evidence the strongest piece of evidence?" At that point, inevitably, the discourse stalled. Even though Kendra asked students to provide the "strongest" evidence, it was difficult to understand what she was trying to elicit because she had not established a criteria for strength of evidence.

Students were able to point to quotes that supported one side or the other but did not appear to understand how to reason that some evidence was stronger than other evidence. So, despite repetitive prompts, student disregarded questions about "best evidence."

Kendra's emphasis on "the strongest" evidence signaled to me that she was trying to elicit sourcing and was particularly interested in focusing student attention on the reliability of the sources. For example, with one small group, Kendra engaged in the following discourse:

K: So, what do you all think is the strongest piece of evidence?

S: (provides evidence - inaudible)

K: …Lets look at the sourcing information.

S: It is like a testimony.

K: Why would that change whether we should believe them?

Kendra followed a similar pattern in the large group where she began with a question about the

most compelling piece of evidence and then moved into eliciting thinking about reliability of the

source:

K: What is most compelling?

S: (reads quote)

K: Why would someone believe the minuteman's statement? Why is he believable?

At various times while groups were working, Kendra's eliciting thinking about the reliability of

the sources appeared forced. For example, with one small group she had the following exchange:

K: Ok, what else. What about the fact that it was a diary?

S: (No response.)

K: Are people honest in a diary? Why would they lie? Where would that go in your chart?

SS: (Silence)

K: Is that evidence or explanation?

SS: (Silence)

K: So, this is a diary of a British guy. What helps us understand it if it reliable, believable?
         Do we think he will lie?"

S5:   No.

178

She later returned to this topic in the whole group discussion, prompting students to think about sourcing and calling on a student she had previously discussed the matter with:

K: Who can use the sourcing info from Barker's diary to create evidence?

S: (Silence)

K: So we know that the evidence we are taking is from Barker's diary. Is everyone looking at the first source? Yeah? So, what does the fact that it is a diary tell us about whether the author is believable? We are trying to think about the author's reliability and how that makes his account believable.

S5:   Why would he lie in a diary. Lie to himself?

Kendra explained during the debrief that she had expected the students to naturally consider the reliability of the source because the two authors' accounts were "so blatantly contradicting …(and) match the two sides of the essential question." She went on to say that the essential question "set up dominos" in which "you end up trying to figure out who to trust more, which pushes you to the sourcing information."

In the debrief, Kendra described the way that she hoped asking about the strongest evidence would get students thinking about the reliability of the author:

> … in hindsight, that question without having been explained on the front end was just over their heads…I was going to have them hold pieces of evidence next to each other and I really wanted them to rank them (and)… justify why something was number one or number two… I feel like the issue was that I was asking them to evaluate two pieces of evidence at the same time and I didn't make that explicit enough and I didn't give them a way to think about (author reliability) that could work for them.

As Kendra suggested above, the students did not turn to author reliability in the domino-like fashion that she had expected. Rather, they focused on the content of the texts, which was not necessarily problematic, but just something that Kendra had not anticipated. In turn, Kendra saw no other way to elicit thinking about reliability and scrambled for some means to promote historical thinking.

It is worth noting that Kendra did not initiate thinking about contextualization during the lesson. Given that the investigation packet included a timeline to support student thinking about each account's depiction of the battle, I expected to hear talk about what happened when and direct comparisons of each account's depictions of the battle. But, Kendra elicited very little contextualization or corroboration on the day I observed.

Because Kendra was not sure about her goals for student thinking in the lesson, she elicited that thinking in an erratic manner and sent it in many different directions. Kendra was unsure of the role that reasoning about evidence, beyond the identification of specific quotes, should play in the investigation. In particular, she was confused about whether reliability of the source was evidence or explanation. In part due to all of this confusion, the student thinking she elicited misfired, and never built any momentum toward cogent lines of argument.

During one of the debriefs, Kendra told me that many of her students did not even understand that the British account and colonists' accounts differed. In other words, even though the fundamental purpose of the lesson was to resolve a tension between two conflicting accounts, some of her students never identified the accounts as conflicting. During her debrief, she did not suggest that the students were not capable of doing the task, saying:

I thought that they could handle both [accounts] at the same time. Perhaps they could

have but I couldn't. That is what I noticed. Maybe with the right scaffolding they could

have done it. But, I was not ready to provide the right scaffolding.

With this, Kendra suggested that at the outset, she could not handle instructionally what the

students might have been able to handle cognitively.

Interpreting. Based on the literature, I defined interpreting in this study as 1) what the

TC reported noticing and 2) factors the TC reported considering for an instructional response.

Based on Kendra's explanations of classroom interactions about historical evidence, I identified

several themes related to noticing and factors considered before responding. Below, I explain

these themes starting first with the outset data and then move on to how interpreting changed

over the course of the study.

Outset noticing. Kendra tended to notice many features in her classroom interactions, so

many that she was overwhelmed. Despite an overwhelming influx of data, Kendra still noticed

some student thinking about evidence. At various points in the debrief, she described particular

incidences of students thinking about sourcing, contextualization, and even corroboration,

although she called corroboration "matching." Recognizing these aspects of student thinking

demonstrated Kendra's ability to notice aspects of student thinking about evidence in history,

something I saw her do often in decontextualized settings in the early data.

Kendra most often noticed that students were not thinking in ways she wanted and

expected them to think. When she expected the students to focus on the documents' authors and

dates, they focused only on details in the accounts. Where she expected the students to make

inferences based on sourcing information, they merely paraphrased portions of the documents.

Even when Kendra noticed student thinking that she wanted, she realized that she was unable to leverage that thinking for an instructional goal. For example, Kendra described noticing that students saw the details of each author's account and the author's reliability as entirely unrelated. She shook her head and said, "there was just not any student who saw it" and she described feeling like ideas were just "floating in free space and not attached to anything."

Kendra also noticed many things that distracted her from attention to student thinking. For example, she noticed that she was confused about the graphic organizer. She noticed classroom management problems and worried that she was "losing them" at multiple points in the lesson.

Factors Considered. While the factors Kendra considered in the pretest and coursework were often based in thoughtful consideration of student reasoning, the factors she considered in the Lexington Green lesson were confused by an apparent flood of data. What was going through Kendra's head as students engaged in the task was best summed up with a thought she remembered during the debrief: " I don't know what to do."

For example, when one student cited a quote about the colonists' backs being turned at the first shot, it set off a string of student ideas about soldiers turning around during a battle. When I asked Kendra about the exchange, she said:

> Like you have to…at least get close enough so that I can get you there because right now you're so far off Google maps that Google can't even find you.…I guess I should've been like "Why is that weird that they turned their backs"…but when I'm standing up there, I'm like "How do I get this back?"…I'm so afraid I am going to go where they are and then end up talking about baseball.

As in the example above, Kendra often noticed particular student thinking, recognized it was different, and potentially distracting, from what she expected, and then did not know what to do with that thinking.

Throughout the lesson, Kendra noticed that students exclusively pointed to quotes in the texts, no matter how many times she asked them to explain their evidence. She described how, "I was trying to get them to say more, to build…something around it…I just did not know how to grab onto them when they just said the sentence again." Kendra also noticed the way some students discussed the author's reliability and others discussed quotes from the documents but no students could "make these two things add up" to a cogent argument about who shot first at Lexington Green. She characterized this as students holding the quotes in one hand and the sourcing information in the other hand. Kendra confessed, "That is as far as I can get them and then I don't know what to do."

Kendra had a vision that students would understand each author's argument in response to the historical question and consider the reliability of each source in order to eventually argue their own response to the essential question. When students started discussing, however, "they had really interesting ideas about context but our goal was reliability and evidence."

Kendra and her mentor had decided that in addition to focusing on identifying quotes and explaining those quotes, the students should be pushed to consider the reliability of the sources. Given students' wandering ideas, Kendra thought, "Should I abandon the objectives and go where they are?" "Where they (were)" was sometimes sourcing, sometimes contextualization, and sometimes even corroboration. Kendra noticed this thinking and considered following many of the student ideas.

During the interview, Kendra demonstrated that in addition to trying to make sense of student thinking, many other aspects impacted her interpreting. For example, she described the influence of the mentor who wanted her to make a "last minute game change" to the graphic organizer. This change, to add the fated explanation column that caused so much trouble in the lesson was a change that she only realized she misunderstood once the lesson began.

Kendra also described the influence of classroom management challenges on her interpreting. When I asked about a particular response, she told me:

> It is not that I need them to sit down and be quiet because they are not going to learn if they don't. It is because I can't think if they don't [be quiet]. I don't know what that student just said. You are reading it to me now and I am like, I don't remember. All I remember is the [student misbehavior].

Kendra described a host of emotions that played a role in her responses including a fear that she was "losing them" and anxiety that "they weren't getting anything out of it." In unpacking what she was thinking during interactions with students, she described thinking about her supervisor's advice and ideas from the methods class.

Part of Kendra's interpreting difficulties were due to her failure to elicit student thinking in an organized way in the first place. Because student thinking emerged so quickly all over the room, Kendra was too overwhelmed to interpret that thinking. Because she did not understand her own graphic organizer, even the supports she had put in place to help manage student thinking became obstacles for interpreting. Had she been able to elicit student thinking in a more targeted and organized manner, she might have been able to make better sense of it.

Additionally, I suspect that Kendra was unable to handle the cognitive load required for the task she launched in the Lexington Green lesson. In the debrief, she suggested that she

underestimated the challenges of managing multiple layers of student thinking at once. For the task Kendra envisioned, she needed to be able to monitor students' reading comprehension, support their understanding of each author's argument, and assess multiple heuristics of student historical reasoning, all at the same time. Even for a veteran teacher, the ability to manage all of that at once would be daunting. For a student teacher, only a few weeks into a new placement classroom, the load was probably impossible. To put it briefly, Kendra jumped into student thinking with both feet but was not able to manage the required cognitive load.

Responding at the outset. In analyzing response patterns, I considered how the TC responded to student thinking when students vocalized their thinking about evidence in history. At the outset, Kendra sometimes responded in ways that made student reasoning the focal point (High II) and at other times responded in ways that only minimally attended to student thinking (Medium).

In the Lexington Green observation, Kendra made numerous attempts to follow student reasoning but she was unable to maintain a thread of reasoning to any meaningful end. For example, in one of the earliest exchanges in the lesson, Kendra entertained a student misconception that the British had fewer soldiers:

K: Did you want to explain why that is strong evidence?

S1:   When you read it you see it was less British and more colonists, so the colonist thought they would win.

K: Did anyone find how many British soldiers there were?

SS: (Several students responded with varying answers).

K: Do we know how many British soldiers there were?

SS: No.

K: So, how does that change your argument, that we don't know how many there were?

With this question, Kendra challenged the student statement by asking the class to address the

misconception and then returned to the student thinking after correcting the misconception. At

that point, however, a student raised his hand and took the conversation in a new direction:

S2:   One of the colonist saw the British marching down and they were afraid and started

        firing.

K: Does anyone hear the context that S2 is building? She said the colonist saw British troops

        marching toward them and they got scared. That might be a reason the colonist shot first.

        Where would the evidence he provided go on your chart?

Rather than answering Kendra's question about S1's thinking, S2 tried to explain why the

colonist might have fired first based on a completely new line of reasoning.

In response, Kendra abandoned S1's reasoning and responded to the second student with

a High II response, asking the class to notice the student reasoning and especially that S2 was

thinking about the context of the event described in the article. But, rather than unpack S2's

thinking further, Kendra quickly moved on to model the graphic organizer and left both S1's and

S2's thinking hanging.

Throughout the lesson, Kendra's attempts to respond in ways that allowed student

reasoning to be the focal point (High II) tended to leave student reasoning 'out there,' detached

from other student ideas, with no means to become a cogent argument. I saw this tendency when

Kendra engaged with small groups as well. Once, when she stopped to check in with a group that

was working their way through the British soldier's diary account, she engaged their thinking but

did not seem to know where to go next:

K: What does this table think?

S1:   The colonist shot first.

K: How can you prove it?

S1:   (Student described the story from the colonist perspective).

K: But what if the British saw people gathered and thought they would be shot?

S2:   But they kept on marching.

K: Look at your chart and circle the piece of evidence you think is strongest.

S3:   What if it was all just a misunderstanding?

S4:   We need a time machine.

K: Let's focus back and circle the piece of evidence we think is strongest.

Kendra's initial response to the group's decision to privilege one account over another for no reason was to offer a counterargument. When this High II response prompted further student thinking, she quickly reined it back in.

Sometimes Kendra's inability to maintain student reasoning with High II responses had an obvious cause, as in an exchange that was sidetracked by student misbehavior:

K: So, what does the fact that it is a diary tell us about whether the author is believable. We are trying to think about the author's reliability and how that makes his account believable.

S1:   Why would he lie in a diary. Lie to himself?

K: So, S1 thinks that she has reason to believe that he would not lie in a diary.

At that point, a disruption occurred in the back of the room and a group of boys laughed so hard that it disturbed the whole classroom. Kendra handled this by turning and asking one of the disruptive boys to answer her question:

K: S11 what do you think? (Students laughed and S11 did not answer) Can someone at this table tell me why the fact that it is a diary makes it more reliable?

S3:   First person point of view.

K: Can someone over here tell me why it is a first point of view is important? Why does that have value?

S1:   Because they want someone….(inaudible)

Although Kendra tried to recover the line of reasoning after the disruption, the conversation turned to another topic and Kendra lost the reasoning about the reliability of a diary.

Despite numerous exchanges that ended in failed attempts to reason together, Kendra showed her resilience by trying to respond to student thinking again toward the end of the lesson. She asked a student to explain why she should believe the colonists' statement?

S5:   Because the source says they are in court.

K: Ok, so we are looking at the source here - so why does that matter that they are in court?

S6:   They all swore on bibles and that means to tell nothing but the truth.

K: You want to add something (to student with raised hand).

S7:   I want to add to what S3 was saying. He said "to our knowledge no one was firing" How could the colonist fire when the colonist backs were turned?

K: So, you are looking at the story to see if the story makes sense. Very good…(Kendra paused for several seconds)…now we are going to talk about our individual positions. You can choose either side now. (Long pause) So, now lets look at our opinions.

Initially, Kendra responded with a high 2 response that kept the conversation focused on the reliability of the author. Kendra's question moved the thinking toward the type of document and initiated more student thinking. But, when the next student interjected, he returned to a piece of

188

evidence raised earlier and derailed the line of reasoning that S5 and S6 were building. Kendra reformulated the comment and then froze, seemingly unsure of what to do. In the end, she responded by leaving both lines of reasoning undeveloped.

Although Kendra attempted numerous times to respond to student thinking in ways that promoted and advanced that reasoning, she was unable to maintain a focus on particular lines of student thinking to a meaningful end. In the face of this frustration, Kendra sometimes responded in ways that appeared to marginalize or ignore student thinking and compel the students toward her own reasoning. In one exchange, Kendra took the observation that a student contributed and made the inference for her:

S8:   The colonist shot first. They were frightened.

K: So you think they were frightened and they wanted to scare them off, ok good.

K: So, who found, is everyone finished writing. Let me give you a minute.

Kendra offered no explanation for what made the student thinking "good." Kendra made the inference but did not model how she made the inference or why she did so. Instead, she took the student's idea, made the inference for the students, and moved on.

In one exchange defined by Medium responses, Kendra repeatedly asked about reliability despite every indication that students did not understand her questions. At the time, students were discussing evidence that the British shot first:

S1:   They said when their backs were turned a shot was fired.

K: Is that what the direct quote says?

S3:   (Student reads quote) "While our backs were turned on the British troops…

K: Why is this piece the strongest evidence that the British shot first?

S4:   (S4 started to read same quote as S3) "While our backs were turned…"

K: (Interrupted) Why does him saying that make it a strong piece of evidence? He could have said anything. What specifically about what he said or where he said it makes it reliable?

S3:   Why would they turn their backs in war?

K: It is a good question and it might go somewhere on the chart? But, why was that quote the strongest piece of evidence?

S5:   (addressing S3) They were turned around because they were leaving.

K: That is what it says but why is this sentence the best evidence that the British shot first?

S6:   I found it.

S5:   Someone had to shoot first to start the war.

K: True. So we are looking for reasons to believe it was the British.

S7:   (student reads another quote) "The number of colonists…"

In the exchange above, Kendra's attempts to get the students to think about strength of evidence seemed to just ricochet off of them. Eventually students just seemed to ignore her questions. Rather than changing course with a High I response such as modeling sourcing, Kendra just kept asking the same question over and over.

At some point, Kendra backed off this responding strategy and unsuccessfully tried to make clear what she wanted students to do. Eventually, she reformulated a student statement to make it sound as though he had addressed reliability (although he had not). By that point in the lesson, both Kendra and the students seemed exhausted.

In summary, Kendra's responding at the outset was characterized by 1) failed attempts to make student reasoning the focal point and 2) follow-ups that were minimally responsive to student thinking.

Change in eliciting. Across the remainder of the study, I observed Kendra inconsistently launch generative tasks in her classroom. I observed Kendra elicit several different types of thinking about evidence but her focus on eliciting such thinking was not consistent over the course of the study.

Change in generativity. During the remaining lessons I observed, Kendra inconsistently launched generative tasks. In some cases, I observed substantial improvement in Kendra's ability to launch tasks in ways that positioned students to reason about evidence and engage in argument building, based on diverse interpretation.

For example, the second observation was like the Lexington Green lesson in that students investigated a historical question over three days (i.e., Were Shay's men rebels or freedom fighters), used two primary sources to discuss, and eventually crafted a written response to the historical question. Most of the materials were the mentor's and Kendra told me that she and her mentor created a graphic organizer to support students in distinguishing between paraphrasing a quote and explaining how a quote related to the author's answer to the essential question.

In the debrief, Kendra pointed out that her students often struggled to articulate an author's position, despite an ability to point to quotes that represented the author's position. The graphic organizer was intended to help students recognize a clear difference between a key quote, a paraphrase of a quote, and identification of an author's argument, all aspects that the students tended to confuse.

Kendra walked the students through the thinking she expected to be most difficult. For example, after giving students time to review the first document, which they read the day before, Kendra focused them on the position of one of the authors in relation to the essential question. Kendra asked, "Is anyone brave enough to explain whether this piece of evidence supports Shays

Rebellion or says it is not justified?" With this question, Kendra launched the task by positioning students to reason about the document and share their interpretations of the document. The Shays Rebellion lesson was like the first observation in its generative design but in the second observation, Kendra was better able to launch the task in a clear and directed way.

The readings, graphic organizer, and spoken instructions all aligned around a three-part focus. First, students needed to read and comprehend the document and interpret the author's position on the essential question. Second, students needed to identify quotes that demonstrated on which side of the essential question the author was positioned. Although less of a focus on the day I observed, the packet for the three-day investigation also included explicit prompts to help students consider each authors' reliability and answer the essential question for themselves.

Although many students struggled with reading comprehension and the thinking the task required, Kendra managed to launch a task that positioned students to reason about evidence in history, engage in interpretation and argument building, and connect to targeted skills. The Shay's Rebellion lesson was the last time I observed Kendra use one of her mentor's "investigations" and it was the last time I observed her launch such a generative task. On two other occasions, I observed Kendra launch tasks that were, by her own admission, problematic because of changes she made in light of design miscalculations.

For example, in the 3rd observation, Kendra designed a task in which students were positioned to reason about Madison's arguments in Federalist 10. Kendra's original design was intended to have students recognize the way that the author was building his argument, claim by claim, in a similar way that her students built arguments when they wrote essays. She created a graphic organizer with space to note the author's main claim at the top and space under each selected quote to paraphrase and explain how the quote related to the author's claim.

With this design, Kendra wanted to continue developing students' ability to identify an author's claim and explain the link between specific textual evidence and the claim. She envisioned students reasoning about how portions of the text related to Madison's overall claim, in light of the philosophical controversy going on at the time Federalist 10 was written. Although the design of the task was potentially generative, Kendra reported that the first class of the day fell flat because most of the students were unable to comprehend the text she had chosen. By the time I observed the lesson 3rd period, Kendra had decided to focus on reading comprehension and identification of the author's claim alone.

Kendra's abandonment of her plan in the Federalist 10 lesson launched a task that was different from her original vision. The revised task was an exercise in reading comprehension and applying the meaning of particular vocabulary terms to students' prior knowledge. Thus, while the original task had generative possibility, I observed Kendra launch a task that focused more on reading than reasoning and more on comprehension than argumentation.

In another lesson I observed, Kendra launched a task noted for its lack of generativity. Again, Kendra designed the lesson without the support of her mentor's materials and again she described having to hurriedly change the task because it did not go as expected in the first class. On this occasion, however, the original task was not particularly generative and the adapted task was even less so.

The lesson occurred at a time that Kendra was laying the foundations for a unit on the progressive reforms of the early 20th century. Kendra wanted to use a primary source to get students to identify a particular content goal: factory life was terrible. Although the graphic organizer included a space to note the source, Kendra explained in the debrief that its only relevance to the lesson was that the author was a factory worker who could describe the hardship

of factory life with authority. The real focus of the graphic organizer was to identify the "main idea of the primary source, "key points and details," and draw a picture of the "events described in the source."

Although the graphic organizer called for the author of the source, the task did not position students to reason about the source beyond a straightforward comprehension of the text. Because the source was presented as an authority on factory life rather than a single perspective on factory life, the students were not invited to interpret the author's argument using the historical thinking tools of previous lessons.

In the debrief, Kendra explained that during the first class, she realized that the source did not clearly convey that factory life was terrible. In fact, the source actually portrayed a much more nuanced depiction of factory life than Kendra intended to convey.

Faced with the loss of her content goal, Kendra's fix was to emphasize certain portions of the primary source and then turn to the textbook to emphasize the difficulties of factory life. This adaptation to the task, which elevated the textbook as the ultimate authority and undermined student reasoning, was misaligned with Kendra's content objective. The result was that the task did not invite reasoning about evidence and actually undermined interpretation by promoting texts as authoritative narratives on factory life.

The full corpus of Kendra's data suggests that she launched generative tasks inconsistently throughout the period of study. When supported with her mentor's materials and activity structures (that mirrored methods course resources), Kendra learned to launch generative tasks. Without this support, and especially when pressed to make adaptations, I observed Kendra launch tasks that focused on structured reading comprehension to the exclusion of student interpretation and reasoning about historical evidence.

Change in thinking elicited. Throughout the remainder of the study, Kendra elicited sourcing and contextualization. But, when pressed with challenges in some observations, she did not elicit historical reasoning.

Kendra's questions and prompts in the Shay's Rebellion lesson were focused primarily on recognition of each authors' perspective by way of a three-part pattern. First, she wanted students to think about which side of the essential question the author was on. Second, she wanted the students to identify quotes that demonstrated that position. Third, she eventually wanted the students to be able to determine whether Shay's men were rebels or freedom fighters, based on the reliability of each author's account.

Where in the first observation Kendra's eliciting seemed intended to promote a free-for-all with historical reasoning surfacing over the room, elicitations in the second observation constrained student thinking to the three manageable categories noted above. For example, Kendra asked students:

K: What side of our essential question do you think he is on?

SS: Freedom fighters.

K: Who can give me a piece of evidence from Document A that supports that he thinks they are freedom fighters?

When one student provided a quote from the document that she believed represented Daniel Gray's position, Kendra revoiced the student quote and then tried to get students to explain how that quote demonstrated the pro side of the argument by saying, "S5 thinks this sentence about the Riot Acts is a strong piece of evidence. Can someone who agrees with her explain why that might be good evidence." When student answers were initially unclear, Kendra stuck with the explanation portion of the discourse and asked the question again in a different way: "Everyone

take a second and look at this sentence (reads fist part of sentence). How does this support or justify that Shays men are justified in doing what they did?"

In the example above, Kendra elicited thinking about the authors' perspectives on the essential question and quotes that demonstrated that position. She also asked students to provide a warrant that connected the quote to the claim. Kendra's ability to subdivide these particular student moves (identify author's position, identify representative quote, and make warrant) was a substantial change from the eliciting I observed in the first lesson.

When debriefing the lesson, Kendra spoke of "getting out" misconceptions. By this, she meant getting to the bottom of confusions that she anticipated students would have about the text or the author's argument. I observed Kendra take a number of risks in an effort to allow misconceptions to become public in the classroom. Her focus on eliciting misconceptions, however, was mostly targeted at comprehension-level understanding of the text and, rather than making the discourse more complex, it tended to devolve discourse to a search for 'correct' interpretations.

Even when Kendra elicited thinking about author reliability, she did so in a way that hinted at a correct answer. For example, she began the analysis of the Abigail Adams document by reading the headnote herself and asking the students, "What about Abigail Adams is important to note? I should see everyone looking at [the headnote]."

Kendra wanted students to notice the author's location (London) as far from the scene of the rebellion. Based on the way she elicited, I assumed that the attention to sourcing was just an add-on and that Kendra was giving up on reliability in order to focus on reading comprehension. However, her debrief explanations made clear that author reliability was central to her vision of student thinking in the three-day investigation and eventual student writing.

The eventual task Kendra envisioned was for students to write an essay answering whether "you think that Shays men were rebels or freedom fighters." To get students to a point where they could engage in that task, Kendra had to open multiple problem spaces. First, she needed the students to consider whether "this author thinks that Shays men are rebels or freedom fighters?" Then, she needed them to consider, "How reliable is this author?" Only then, were students ready to make an argument that Shay's men were rebels or freedom fighters.

Opening these overlapping problem spaces required that students comprehend the text, identify the author's perspective, point out text that supported the author's perspective, and evaluate the author's reliability. Kendra reported during the debrief that each of these overlapping and interconnected inquiries was not difficult when isolated but difficult when students were expected to reason with all of the information at once.

While Kendra identified the separate aspects of overlapping inquiries much better than in the first observation, she confessed that she did not know how to help students bring the thinking together in a way that helped them address the essential question. This confession perhaps explains why most of the eliciting in the Shay's Rebellion lesson was focused on relatively straightforward tasks of reading comprehension and identification of argument.

Kendra saw reasoning about author reliability as a second level of attention for which she wanted to make space but did not want to initially distract from comprehension of the author's argument. She explained in the debrief:

> I wanted them to give me a piece of evidence and then explain it because that was still my overarching goal … And I wanted to break it into two sections like, we are going to have a claim about what this author thinks and we are going to have a claim about what this author thinks and I want you to make an argument for your claims about what the

authors think…And then later I wanted to compare them…I wanted to find a piece of evidence and I wanted to see them explain it. And I thought that that was worth my crunch time.

Unsure of how to incorporate historical reasoning as an additional thread of reasoning during the lesson, most of Kendra's elicitations focused on students' comprehension of the documents and identification of the authors' arguments.

In the Federalist 10 lesson, Kendra intended to elicit student thinking about the author's perspective in light of a particular historical context. Kendra explained her vision for contextualization, although that is not what she called it, when I asked her why she showed a video clip about the Constitutional Convention:

… I wanted them to go back and remember [it] just happened; remember the states are trying to figure out where power is, people are no longer sure of themselves … The Constitution has been drafted but it's not been adopted yet… I wanted them to be able to put the conversation there and I wanted them to have a brief reminder…that this is one side of a two-sided debate.

Kendra wanted the students not only to comprehend the argument of the Federalist 10 author but to place that argument in its historical context.

As I noted earlier, however, when Kendra initially realized that her students could not handle the reading, she scaled back her goals for student thinking and abandoned the contextual goal entirely. Rather than linking the argument in Federalist 10 to the context, she focused on eliciting comprehension of the text, identification of the author's argument, and a definition of two key terms in the reading.

198

The revised version of the lesson was so focused on reading comprehension that any contextualization goal fell off the radar screen. For example, when Kendra tried to do a "think-aloud" model of the first selection of text, she demonstrated only defining words and paraphrasing. She connected the paraphrase to a claim but that claim was disembodied from the author and context. After demonstrating the reading, she asked:

K: Who can raise their hand and tell me what the claim is for Quote 1?

SS: (Silence)

K: No one? (Wait time)

K: What is the author trying to convince the reader of?

This last question would have invited students to contextualizae the document if the students had recognized the author, context, or type of document. As it was, the scaffolding and prompting was all in support of reading comprehension and paraphrasing.

As students struggled with reading comprehension, Kendra elicited thinking about the meaning of key words. For example, after the first segment of text, she asked, "What does the author think about factions?" But, some students did not know the meaning of 'factions' and they first had to address the vocabulary. At another point, Kendra asked, "What does self-interested mean?" She then tried to get students to recognize how the author felt about self-interested persons.

Outside the showing of the video, I did not hear Kendra elicit any thinking about the context of Federalist 10. Recognizing the author's claim in the absence of recognizing the author or the context of the author meant that Kendra only elicited thinking about justification. Given the miscalculation Kendra made in the reading ability of her students, it is understandable that she narrowed in on the fundamental task of reading comprehension. However, Kendra could

have retained a focus on contextualization by noting the author and connecting the document more explicitly to the video clip she showed at the beginning of the lesson.

During the lesson on factory conditions, elicitations focused only on comprehending the meaning of the texts and recognizing Kendra's predetermined conclusion about factory life. As the students read the primary source on factory life, Kendra walked around the room asking students "What did you notice about the source?" This question pointed to nothing in particular and did not appear substantively different than her earlier instructions to write down the "main points" and describe life for the author in the factory.

Kendra did not elicit thinking about controversies present in the source or between the source and the textbook. These controversies only existed because Kendra had chosen a source that did not align with her learning objective.

Despite the use of the two sources, Kendra did not ask students to corroborate between the textbook and primary source, and in fact, minimized differences when students raised them. She did not identify the textbook as a secondary source and did not ask students to interpret it differently from the primary source. Although in the debrief Kendra envisioned a lesson that might have promoted corroboration of the source and the textbook, she felt that there was not enough time to engage in that activity.

Overall, Kendra's eliciting was inconsistent. Her ability to launch lessons that elicited student thinking about evidence certainly improved from the Lexington Green lesson to the Shay's Rebellion lesson. In other cases, such as the Federalist 10 lesson and factory conditions lesson, Kendra adapted lessons in ways that deemphasized and even undermined historical reasoning such as corroboration and contextualization.

One possible explanation for this inconsistency was the availability of the mentor's materials and task structures in the "investigations" I observed. When Kendra designed tasks independently, she miscalculated in ways that required adaptation. The adaptations were not particularly generative and resulted in the limitation of student historical reasoning.

Changes in interpreting. Throughout the remaining observations, Kendra's patterns of interpretation changed. Although Kendra continued to notice many aspects of student thinking and classroom life, she developed a framework to help interpret the complexity and manage the factors she considered before responding. This framework was both helpful and constraining.

Changes in noticing. Where Kendra reported noticing a cacophony of inputs at the outset, her later descriptions suggested that she developed specific outcomes to direct her noticing in subsequent lessons. Kendra explained that after the Lexington Green Lesson, she began to organize her lessons more tightly around "outcomes," a term she used to describe short-range, assessable learning objectives. She used these outcomes to organize the flow of her lessons and she described her noticing in the lesson in accordance with these outcomes. She described this in detail in our final interview:

> I want to check for my outcomes. I don't want to skip them. I want to go in order. So, I have spent a lot of time ordering my outcomes when planning my activities. Because, I know if I have that checklist, I can move them forward in the lesson, even if they don't get as far forward in the lesson as I expected them to.

Although her outcomes for each lesson were different, she described them as a "checklist" that typically began with noticing whether students were able to comprehend the reading. Whenever she spoke of moving through the lesson according to the outcomes, I noticed that she moved her hands in the shape of a pyramid.

And, a pyramid of thinking was how Kendra seemed to organize the complexity of student thinking. It began with assessing reading comprehension, then moved toward whether students could summarize the argument, and, as long as students were able to proceed, it progressively moved toward higher levels of thinking. Toward the top of the pyramid Kendra envisioned for her students was historical reasoning, an elevation on the pyramid that they sometimes were unable to reach.

For example, in the Federalist 10 lesson, Kendra immediately knew that her vision for the lesson was in trouble when it became clear that most students could not comprehend the reading. She described in the debrief what she noticed after she model-read the first paragraph:

> There was definitely a group of students who must have checked out while I was reading the quote…they didn't want to engage in the text at all. Which tells me, overall, I completely overshot what they can do.

Kendra described noticing that some students refused to participate and, even "students who are very agreeable, really struggled." Kendra described thinking at the time, "Hmm, my lesson is really hard." She attributed this difficulty to the reading level of the text.

Second on Kendra's noticing checklist was whether students could select a quote that demonstrated a claim and explain how the quote demonstrated that claim. For example, in the Shay's Rebellion lesson Kendra reported:

> They're very good at picking out which quote is important…but getting them to that next step of explaining is what I really wanted them to be able to do, to explain why it supports it and not just read it back to me or paraphrase it.

Even in the Factory Conditions lesson, Kendra reported noticing whether students were able to select a quote and explain how the quote demonstrated evidence of a claim.

202

In many cases, what Kendra noticed that provided evidence about whether or not to move up the pyramid was whether students made a comment that reflected what she "wanted to hear" or expected to hear. For example, in the Shay's rebellion debrief lesson Kendra described the way that students engaged each of the author's perspectives without rejecting one or the other as "just an opinion." She noted,

> I was afraid …they would…be like well, he is on this side but that doesn't matter because it's just an opinion…but you know, I did not hear that once…that's improvement from last time when an opinion was something insurmountable and …should be dismissed.

Likewise, when students were struggling through Federalist 10, she noticed that one of her students was able to evaluate how the quote supported the author's claim, as she had envisioned.

More frequently, Kendra noticed when student statements reflected something she did not want to hear. For example, in the Shays' Rebellion lesson she noticed that many students picked quotes and "just read them to me." Kendra described how students "are able to pull out quotes that are important but they are not able to explain how that quote fits into their argument."

Similarly, in the Federalist 10 lesson, Kendra described how she wanted the students to see how each segment of the text was built into "one narrative that was going to support …the claim and I wanted them to dissect it, like an essay that I would have them write,…but that was too sophisticated."

What Kendra did not notice is perhaps more important than what she did notice. While in the outset data, she described noticing dozens of aspects of student thinking about evidence in history, this attention seemed to wane in later lessons. Although Kendra discussed sourcing, contextualization, and corroboration in regard to other lessons I did not observe, Shays Rebellion was the last debrief in which she mentioned noticing these heuristics.

I suspect this devolution in noticing historical thinking had to do with the priorities she established with the pyramid of outcomes. If students could not understand the document, could not identify the claim, or could not connect a quote to evidence, historical thinking may have been a priority that was out of reach in Kendra's mind.

Although this prioritizing had detrimental affects for the historical reasoning that she noticed, it had a positive impact on her teaching overall. Where at the outset, Kendra noticed many other things unrelated to student thinking, many things that became distractions for her attention to student thinking, she reported that prioritizing specific, ordered outcomes gave her the chance to decide "which balls to drop."

> … in the beginning … I did not have the priorities ordered….later it was not that I wasn't thinking about those things it is just that I knew they were kind of irrelevant. So, the kid that was off task, I am not going to let one off task kid throw my entire lesson this time. He can just be off task for 30 seconds and I can keep going. And, I think that that prioritizing took a while for me to figure out how to do while listening to kids.

Where in the Lexington Green lesson, Kendra noticed many things about classroom management and described them as distractions for her attention to student thinking, she only mentioned classroom management one other time in the rest of the data.

Kendra continued to notice other distractions, such as the expectations of her supervisor and other competing interests. Of these competing interests, she said "It was not that they ever went away. It was just like, ok, I am only going to juggle four balls and these are the four." Staying with the juggling theme, Kendra said:

> Now, if I have to juggle seven balls and I know I need to drop one, I know which one I can drop. And, just being able to pick that is really freeing with the mental space. So, I

can call on a student that really struggles when I know I can spend a lot of time on

whatever she says…

Kendra learned to hone her noticing to outcomes that she prioritized going into the lesson. These

priorities helped her recognize if she could "drop one."

Changes in factors considered. The factors that Kendra reported considering changed

over the course of the study. Central to the factors that Kendra considered when she formulated

responses to student thinking were the outcomes that she had for each lesson. She described

outcomes as "a checklist" that would guide her in whether "I can move them forward in the

lesson, even if they don't get as far forward in the lesson as I expected them to."

Therefore, when she noticed that students were unable to accomplish a task's reading

comprehension, she recognized that she needed to either support or scale back the task. For

example, in the lesson on Federalist 10, Kendra recognized that she "overshot" their reading

ability and determined to abandon her contextualization outcomes in order to address the

reading.

In the same way, if students were able to comprehend the reading but not identify the

claim, she determined to support or scale back the task in some way. For example, when she

noticed that, "they didn't catch Madison's claim as fast as I thought they would…that threw me

for a loop." Kendra explained that she had not intended to spend much time helping students

identify the claim but she recognized that many students were confused.

Where the outset data were characterized by a sense that "I just did not know what to do,"

Kendra increasingly felt that she did know what to do because she was moving through her

outcomes checklist. She still reported "floundering" from time to time but the outcomes served

to provide direction in guiding her responses. For example, she reflected about the Federalist 10 lesson:

> It was like "Oh shit, they can't do this but I can fix this somehow in real time." It was the first time I ever felt like that. Because I remember during the (Lexington Green) lesson thinking, "this is going really poorly, and I have no idea what to do." But, by Federalist 10 I was like, "I can bring this back." So, ok if this is where they are struggling, I need to scale it back, and I need to go back to vocabulary. Like I knew where to go if they were lost and I knew how to build it back up. That was the first time I really felt like I could do that.

A focus on the outcomes allowed Kendra to narrow the field of student thinking significantly. In so doing, she was able to interpret the complexity of student thinking and classroom life more effectively. The downside of this narrowing was that historical reasoning fell lower in the checklist, or higher in the pyramid, than other cognitive tasks that students struggled to accomplish.

Changes in Response. Across the remainder of the study, Kendra's responding was much more strategic than at the outset. These changes appeared to run parallel with decisions that she made to narrow the scope of student thinking that she elicited.

There were times during the later data that Kendra's follow-ups were High II. That is, she explored student thinking and allowed student reasoning to be the focal point. For example, at one point in the Shays Rebellion lesson, Kendra engaged with student thinking about which side of the essential question Daniel Gray might be on. When a number of students insisted that Gray thought Shay's men were "freedom fighters, Kendra responded:

K: Who can give me a piece of evidence from Document A that supports that he thinks they are freedom fighters? (Wait time).

S4: (S4 student reads her evidence quietly - inaudible).

K: So, S4 thinks this sentence about the riot acts is a strong piece of evidence. Can someone who agrees with her explain why that might be good evidence?

S5: Because he says he thinks the acts are good.

K: Whose acts, the police's acts? The police or Shays men?

S5: The police are Shays men?

S2: The police OR Shays men.

K: Wait a sec. You said their acts are good. Whose acts?

S5: Shay's men.

K: So, how does this sentence "The riot act…" (reads quote) suggest that Shay's and his men's actions are good?

SS: (No response).

K: Everyone take a second and look at this sentence (reads first part of sentence again). How does this support or justify that Shay's men are justified in doing what they did?

In this exchange, Kendra's responses maintained a focus on the original student claim (i.e., that Gray's speech portrays Shay's men as freedom fighters). She used High II responses to focus the discourse including a substantive probe, asking for textual support, and revoicing a student idea. When the discourse threatened to go off track, Kendra maintained the line of reasoning by asking the students to make sense of S4's initial assertion.

At another point in the Shay's Rebellion lesson, Kendra responded to a student assertion by probing student thinking and then tried to correct a misconception and return to the original

student idea. This exchange began when she responded to a student's statement that government officials were portrayed in Shay's account as above the law:

K: So, what does that tell us about whether Shays and his men were justified in doing what they did.

S1:   If you look at one point (Reads quote)…It is impossible to bring them to court…revenge, hatred…I think what he is trying to say is that their actions are depending on how they want to end this. Like, if they want to go to court they have to have a lot of evidence for what is going on.

T: You are right if the court system is working correctly. But, what he is saying is that it is not possible to bring them to court, even if they had evidence. So, is that an argument for or against the court system being just? (long pause)…Take a minute to turn and talk with your table. Discuss with your partners whether this supports the actions of Shay's Rebellion or against the actions of Shay's Rebellion.

Kendra first responded to student thinking by asking the class to make sense of the initial argument. When S1 responded with a miscomprehension of the text, Kendra responded by explaining the author's point and returning to the initial idea of whether the legal system was just. Although the response was not a stroke of mastery, it appeared to be a serious attempt to place student reasoning at the fore.

Although Kendra occasionally demonstrated High II responses, she more frequently took over the student thinking with Medium responses. For example, in one exchange from the Shay's Rebellion lesson, Kendra controlled the conversation by bouncing questions back and forth in a way that led students to identify the author's perspective on Shay's men:

S1:   It is going against Shay's Rebellion. The government …

K: Uh huh. And what is government doing that is unjust?

S6:   Increasing taxes.

K: Yes, but what specifically makes Shay's men attack courts and not state government buildings?

S7:   Putting farmers in jail.

K: …So, what side of our essential question is Daniel Grey on? Were Shay's men rebels or freedom fighters?

Rather than promote student reasoning, Kendra asked questions that increasingly guided the students in the direction she wanted them to go. In the end, they arrived at the destination but it was Kendra's reasoning that brought the students to that destination and that reasoning was not elucidated for the students.

When students moved on to Abigail Adams' perspective on Shay's rebellion, Kendra engaged in a different type of Medium response pattern when the students did not immediately notice Adam's perspective on Shay's men:

K: Can anyone find an important quote to help us understand whether Adams thinks they are rebels or freedom fighters?

S2:   (S2 tries but confused)

S8:   (inaudible)

K: S8 can you say it again? Can I read it? (Kendra reads from S8's graphic organizer)

K: What side does that make us think she is on?

S11:   I have another one. This mob makes us…weakens our country.

K: Very good so she thinks these are rebels.

Instead of supporting S2's or S8's reasoning and working to make their understandings visible and articulate how to assess the author's perspective from the identified quote, Kendra transformed the last student statement to provide the class with the answer that Kendra intended.

Even though Kendra used Medium responses to lead students, most students still seemed unsure of Adams' perspective on Shay's men. Kendra tried again with another portion of the Adams source but kept running into problems when students struggled with the reading comprehension. Inevitably, she felt the need to lead their thinking or just tell them Adams' position. In the following exchange, Kendra started off by rereading a text a student pointed out as demonstrating Adams' position:

K: Ok, (rereads and continued reading beyond where the student stopped)…What does that tell you?

S8:   (inaudible)

K: Why do you think that?

S8:   (inaudible)

K: Ok, by helping the state, what does she mean? … If they are not talking about physical borders, what else could it be?

SS: (Long Pause)

K: …Let's see if we can figure this out. Shay decided to attack courts and not state houses. What do those two things have in common?

S10:   You can't have one without the other. If someone breaks the law, then you go to court to deal with them.

At this point, I noted that Kendra seemed unsure of where to go next and hesitated before her next move. The students, however, remained attentive and patient.

K: …So, Adams is hoping what?

S12:   She is hoping they won't find any justification for what they are doing? (several students

         expressed confusion)

K: You don't understand? What do you not understand?

S2:   At first I thought she disagreed with Shay and now I think she agrees with the rebellion.

K: If Abigail Adams thinks an investigation will show that the rebellion does not have

         justification then what side is she on? Talk about it in your groups.

In the exchange above, Kendra consistently tried to turn the thinking back over to the students

but she continued to face challenges posed by reading comprehension and student confusion. In

the end, she simply told them Abigail Adams' perspective.

       Kendra infrequently responded to student thinking about evidence with Low responses or

follow-ups that are not responsive to student ideas. However, in the lesson on the conditions of

factory life, she had few opportunities to respond to student thinking about evidence because she

did not elicit such thinking. Many of her responses seemed intended to highlight the meaning of

the text, accomplish her content goal, and in some cases avoid student thinking about evidence.

       For example, at many points, Kendra responded to student contributions by asking for

justification. In the following exchange, Kendra prompted students to explain what is known

about factory life from the source:

S7:   It is hard working.

K: What did you see in the source that says it was hard working?

S7:   (said something about source in inaudible).

K: Is that what she was looking at? Show me where.

S7:   (Pointed to picture above the text in the source)

K: In the picture (above the source) we see people working hard. What do we see in the
         source S8?

These requests for justification did not facilitate student reasoning as much as they demonstrated
that students comprehended the meaning of the text. In fact, Kendra worried that allowing the
students to reason too much might result in the disruption of her lesson objective.

Kendra actually avoided some students' comments that the textbook description of the
factory did not sound very bad. During the debrief, Kendra described wanting to rush students
through portions of the text that provided a nuanced picture of factory life because that put her
content objective at risk.

Although not equivalent to her ability to respond to students' historical thinking,
Kendra's responding to student reading comprehension challenges was relevant to my analysis of
her responding patterns. Reading comprehension was a factor that regularly impacted Kendra's
responding to students' historical thinking.

For example, in the Federalist 10 lesson, the discourse was constantly mired in confusion
about the meaning of the text. Kendra spent much of her energy responding to student thinking
about the reading. In one characteristic exchange, Kendra started by asking the students to notice
the author's claim:

K: What is the author trying to convince the reader of?

S14:   Make everyone equal.

K: Did he saying anything about equal?

S14:   (inaudible)…changing sides.

K: Has he said anything about changing their sides? S14, you are on the right track. He did

        say something about whether they were good or bad. S3, did he think factions were good

        or bad?

S11:   Bad?

K: S3?

S3:   Good.

K: Raise your hand if you agree with S3.

S10:   What did he say?

K: That the author believes that factions are good. (Some students raised hands)

Raise your hand if you disagree. S4, can you share why you disagree? (S4 laughs) Who can

share with us why they disagree? (S1 raised hand)

K: S1, go ahead.

S1:   They disagree because they know that some people will not agree with it. So they think

        that it might be a wrong thing.

The discourse proceeded in this way for several minutes. Students answered question-by-

question and Kendra plodded through, trying to help the students comprehend the text, and

sometimes forgetting the original question that began the exchange.

Kendra's tendency to respond to these reading comprehension challenges with a initiate-

response-evaluate pattern sometimes meant that fundamental misconceptions did not emerge

until late in a conversation. For example, an amusing exchange occurred after several minutes of

a large class back-and-forth about the meaning of the term "factions" in the text. Thinking that

students finally understood the term, Kendra asked them to "take note of what the author thinks

of factions." After a few moments of silence, a student up front said:

S12:   You mean clothes?

K: What?

SS: No!

S12:   Fashions?

K: Factions. Ok, (long pause), I know…this word here, 'factions,' based on this quote, does anyone want to guess what this meant? (S11 raises hand)

S11:   Groups.

K: Yes, very good. Groups of different sizes that have goals. Great job S11.

Because Kendra tended to control the discourse during these reading forays, it was difficult to determine what the students actually comprehended. Even in her response above, Kendra did not return to S12 but instead grasped S11's answer to emphasize the meaning of the term. The challenges that Kendra faced with reading comprehension made responses to historical thinking all the more difficult. Her attempts to explore student historical thinking seemed always checked by concern for whether students had understood the text.

Summary of Findings

In this chapter, I have considered how Kendra elicited, interpreted, and responded to student thinking about historical evidence. My first research question asked how Kendra engaged in these practices at the outset of the program. Initially, Kendra launched a task that elicited student thinking including sourcing but much of the resulting discourse was clouded with confusion. Despite the confusion, Kendra still noticed some student thinking about evidence but just did not know what to do with it. Even though Kendra tried to build on student reasoning with her responses, she was unable to maintain a thread of reasoning and sometimes turned to Medium responses in order to salvage a particular point.

214

My second research question asked how Kendra's capabilities to engage in these practices changed during the study. In the later lessons, Kendra inconsistently launched generative tasks and inconsistently targeted historical reasoning. The outcomes framework she envisioned helped her manage the complexity of classroom life but also may have constrained the reasoning she prompted and noticed. While Kendra sometimes responded to student thinking in the later lessons with High II responses, she frequently led student thinking with Medium responses.

Chapter 8: Craig

Craig was a 23 year-old white male from Long Island, New York. His gregarious personality and outgoing nature served him well in the relationships I observed him build with students and staff at his placement school. Even though Craig was new to classroom teaching, he had extensive experience with kids because of previous experiences as a camp counselor, tutor, and coach. At the start of the internship, Craig saw these experiences, along with his ability to effectively relate to students, as his greatest strengths.

Craig had an undergraduate history degree and was enrolled in the university's five-year Master's and teacher certification program. A part of Craig's degree plan also included an ESOL certification. Of the social studies, Craig was reportedly most confident teaching world history, a specialty that would serve him well in his world history internship placement.

Craig explained that Social Studies Methods II was the first time he heard historical thinking skills described explicitly. He remarked that as a history student he had engaged in sourcing, contextualization, and corroboration, but did not have language to describe those heuristics. He performed particularly well on the historical thinking portion of the pretest and his early coursework demonstrated interest and proficiency in using and discussing the historical thinking skills.

In Craig's reflections, he often described and analyzed lessons he observed in other social studies teachers' classrooms at his internship school. Craig was specific in both his critiques and compliments of the instruction that he observed, and demonstrated an early interest in questioning and discourse practices. In several reflections, he considered the effectiveness of questions that teachers were asking in their classes in promoting student thinking. He saw several teachers who used questions very skillfully to promote participation, scaffold student thinking,

and assess student understanding of objectives. In other cases, he critiqued questioning that was not effective. Craig's lesson plans and coursework demonstrated that he thought deeply about the phrasing of his own questions to students.

Craig's interest in questioning and the discourse that proceeded from questioning was born out of a commitment to what he called "constructivism." Once when discussing his undergraduate courses he explained,

If there is one thing I walked away with it is that idea of enduring understandings. When we looked at all those different philosophies of learning, I definitely identified with constructivism. Like people learn by constructing understandings and they are not going to remember little details. They will take away big understandings that they construct.

Part of Craig's effort to teach with "enduring understandings" was to relate his history classroom to his students' lives. This particular value was at the heart of Craig's philosophy of teaching and became the rationale for his action research project and an emphasis of many of his reflection papers. He explained:

I strongly believe that one of the most important aspects of social studies education is connecting historical topics to the contemporary world. In addition, I believe that this makes the content more interesting and improves student participation.

In early coursework Craig referred to disciplinary thinking skills in general terms. For example, he planned to use primary sources but what these sources could provide did not seem like a central focus of what he was trying to do. As early as November, Craig was thinking more strategically about disciplinary skills of history and was convinced that he knew how to teach these skills better than his mentor.

Part of Craig's teaching style was to ask questions to promote engagement and participation. Craig reflected on this after one of his teaching attempts in the fall:

My host teacher commented that he could not believe how many questions I asked. He found that almost everything I said, outside of explicit instructions for the activity, was a question… I was trying to focus students toward certain connections and pieces of information [but] I did it through questioning such as, "What do you see in this picture?" rather than saying, "Notice the child worker is not wearing shoes." Although this may take longer and lead to off-topic answers, I believe it builds a better, constructivist, understanding. Also those off-topic answers are sometimes great connections or perspectives that I never considered.

As this segment reflects, Craig was already thinking about eliciting student thinking and interpreting that thinking.

Context

Craig was placed in large suburban high school that had a reputation for student success because of its high graduation rate, relative to its 38% FARMS population. Although ESOL students made up only 12% of the 3000 students at the high school, Craig's classes had a high population of these students.

Craig's classes included eleventh and twelfth graders with widely varying reading and writing abilities, and many students with specific IEP and 504 accommodation requirements. In the class I typically observed, for example, seven students had IEPs and five others were former English Language Learners or spoke another language at home.

Craig described his mentor as kind, helpful, and honest. His mentor openly assessed his own teaching with Craig and pointed out weaknesses and invited Craig's support and participation. The mentor acknowledged that he did not do much to support his ESOL students' in target language objectives and suggested that, as a general education teacher, he saw his responsibility as relating to students' content literacy rather general literacy.

Given Craig's background in ESOL, Craig was surprisingly generous about his mentor's omission in these regards. Once the mentor provided Craig an opportunity to lead an activity and recognized that Craig could manage the behavior of the class, he turned over a tremendous amount of responsibility.

As Craig was transitioning into the lead role, he closely attended to the district curriculum and how his mentor and others in the building implemented that curriculum. Craig told me that his mentor used primary sources and even worked some with historical thinking skills, although these activities were distinct from the normal activities of the class. But, given the autonomy that his mentor trusted him with, Craig felt confident that he could pursue instruction in the ways that he envisioned.

Eliciting, Interpreting, and Responding

In the following section, I describe Craig's ability to EIR over the course of the study. The data that informed these findings include information drawn from the pretest Craig took prior to coursework, course participation in the following months, five classroom observations and six interviews. My first observation of Craig's classroom teaching occurred in early October, 2014. My final observation was in late March 2015.

Eliciting at the outset. As described in Chapter 3, I analyzed eliciting according to two features: 1) generativity of the task the TC launched and 2) the specific type of historical

thinking the TC elicited. Across outset data, Craig sometimes demonstrated an ability to envision and launch generative tasks. During these tasks, Craig elicited thinking about justification, contextualization, and elementary thinking about corroboration.

Outset generativity. Across the outset data, Craig increasingly envisioned and launched generative tasks. The pretest lesson plan that Craig designed was the least generative of the outset tasks. He sketched a lesson in which students would use several primary source documents to consider the question, "Is it within States' rights to secede from the union?"

Craig's pretest demonstrated that he valued the use of primary source documents and rich questions that could promote interpretation and evidence-based argumentation. However, his description of the lesson left unclear how the students would reason during the task. For example, Craig noted that students would use the Declaration of Independence and 1869 Supreme Court case to debate the essential question. But, he did not explain how these documents would provide evidence to address the essential question or whether multiple interpretations of either of the documents were possible.

Instead, Craig described wanting students to arrive at an "enduring understanding" about the push-pull between state and federal power in American history, a perfectly good objective but not necessarily one that positions students to think like historians. In the pretest, then, Craig demonstrated some beginning aspects of generative design but it did not appear that the lesson would have promoted rich discourse about evidence in history if it had been launched.

By the second month of the course, Craig sketched a very different type of lesson when asked to plan a task that would help his students understand the difference between history and the past. Craig envisioned a "Who Done It?" murder mystery activity in which groups of

students would examine multiple documents, some more trustworthy than others, in an effort to determine who committed a fictional murder.

In this lesson, Craig imagined positioning the students as detectives in order to bridge to a conversation about history as theories, theories based on evidence and not on authoritative narratives. He expected this task to introduce students to the importance of multiple forms of evidence in history, how some evidence may be more or less reliable, and what happens when some evidence contradicts other evidence. Although this activity did not include actual historical sources, it showed that Craig could envision a generative task anchored to students' everyday knowledge.

While, to my knowledge, neither of the above tasks were used in Craig's classroom, both demonstrate aspects of the potential and limitations I observed in the generativity of the tasks Craig launched at the outset of the study. For example, two early tasks that Craig launched demonstrated dissimilar generativity. In his first course-required enactment, Craig facilitated a task similar to the pretest, in that it included multiple primary source documents and a compelling essential question. Like the pretest, opportunities for diverse interpretation and reasoning about evidence in this lesson were unintentional, if present at all.

The task required students to read adapted writings of three Enlightenment political philosophers and work, first individually and then in groups, to answer three guiding questions. For example, as the Rousseau group read, they answered the following questions:

1. What is a social contract?

2. According to the Social Contract, where does the power to govern come from?

3. How could these ideas transform European society?

The questions positioned students to reason, not only about the content of the documents, but also about how the specific Enlightenment philosophies could challenge and transform 18th century European society. Engagement in this second part of the task would have required students to examine the philosophies and compare those philosophies to the concept of absolutism that they previously studied.

The third question offered opportunity for diverse interpretations and argument building grounded in available evidence. As I listened to their discourse, I noticed that the students were able to answer the first two questions and dialogue about the ideas of the philosophers but few could even begin to answer the third question. Thus, although the task pushed students toward new ways of thinking, it offered little engagement with students' prior understandings, except for their understanding of absolutism based on previous content. In this lesson, Craig treated contextual thinking (as demonstrated in the third question) as something that would be natural for students, as if the three questions would act to 'connect the dots' and precipitate a contextual inference (i.e., these ideas became revolutionary in the context of absolutism). As a consequence, rather than positioning students as historical detectives, the only two questions the students could answer positioned them as consumers of the political ideas articulated in the readings.

In the course-required discussion lesson, Craig launched a task that had students develop, critique, and "evolve" a hypothesis about the Industrial Revolution. Students first examined primary source evidence that suggested the Industrial Revolution had all positive societal impacts. In the second round, the primary source evidence pointed to negative societal impacts. In the final round, students examined two secondary sources that articulated each side of the argument. After each round, students "revised hypotheses" based on the evidence and, in the end, took a position and cited specific evidence from the documents.

Similar to Craig's early "Who Done It?" murder mystery, this task was highly generative in promoting student thinking about history as method rather than an authoritative narrative. Craig invited diverse interpretations and argument building grounded in evidence. He asked students to consider various perspectives and would not allow students to simply choose one side or the other as an easy way out of the question "How positive was the Industrial Revolution?"

As in the lesson on the Enlightenment, however, Craig's seemed to assume that if he created a task that required complex historical thinking, the students would be able to engage in those new ways of thinking with little explicit instruction in the thinking skill. The structure of the activity supported students in comparing different versions of the same story but many seemed confounded by what to do with conflicting evidence. Thus, at the outset, Craig demonstrated an ability to design and launch generative lessons but he struggled to successfully bring students into the type of discourse he hoped to see.

Outset thinking elicited. In the outset lessons, Craig elicited thinking about contextualization, and corroboration, and justification. Over the course of the outset lessons, Craig was increasingly intentional about how he taught historical thinking.

Although not launched, the murder mystery lesson was designed to specifically elicit conceptual thinking about evidence in history and serve as a foundation for future student reasoning about evidence. Craig envisioned the activity as "perfect for the first day of class," an anchor for future historical thinking. Later Craig explained that at the time he created the murder mystery lesson, he was becoming aware of the importance of sourcing, contextualization, and corroboration in teaching history. He recognized that unless he could elicit student thinking about the way that history is constructed, students would not ever engage with complex historical thinking.

Even though Craig's lesson on Enlightenment political philosophies did not make contextualization clear, his third question on the graphic organizer, and the discourse he tried to promote in groups, asked students to explain how a political idea could transform European societies that were based on absolutism. When I observed the lesson, I thought that Craig was trying to elicit contextual thinking about the cultural values or common attitudes of the time period in which these philosophies were being introduced, and why the philosophies were revolutionary at the time.

Only months later did Craig recognize it as contextualization saying, "if it was contextualization, it was accidental…I was still thinking mostly about this notion of enduring understandings and had not figured out how I wanted to deal with disciplinary thinking." Not surprisingly, students struggled with the third question of the graphic organizer because they had little sense of contextualization.

In the lesson on the Industrial Revolution, Craig repeatedly instructed the students to "corroborate the documents" as they worked through each round of the "evolving hypotheses." This lesson marked the first time I observed Craig intentionally elicit historical reasoning and clearly articulate to the students the thinking he wanted them to do. During each round, Craig elicited thinking about how the documents answered the essential question "Was the Industrial Revolution positive or negative?" This question required students to comprehend and compare the documents from each round and account for evidence to answer the question, round by round.

After each round of evidence, Craig would say something like, "Make sure you write your revised hypothesis...Update this based on whether those documents corroborate what you already wrote. Do they change your answer?" As students discussed their revised hypotheses and

filled out the graphic organizer, Craig elicited justification by requiring students to defend the argument of each round with evidence from the documents. As the rounds proceeded, Craig regularly reminded the students of what they were doing with prompts by saying, "as you are reading, you are constantly thinking, do these corroborate?"

Admittedly, much of the corroboration in the Industrial Revolution lesson was significantly scaled back from the corroboration that historians do. Craig did not have the students first attend to the documents' sources or contexts. Instead, they were invited to comprehend conflicting accounts and use the contents of the documents to defend an answer to the essential question. In some respects, Craig was inviting students to compare two narratives rather than compare the documents. However, during the final round of discussion, he pushed the students to begin looking across all the documents to corroborate, not only round against round but document against document. He said, "What if we connect it back to some of the earlier ones. What if we look at Document A (from the first round) versus Document D (from the second round)?"

After the final round, Craig again explicitly highlighted the skill of corroboration by saying, "Now, lets fill out the big box. This is your real answer. I want you to touch on multiple sources. Corroborate all these sources into one answer here." With this task, Craig's elicited thinking that challenged students' ideas about history as an authoritative narrative and established an elementary understanding of corroboration as comparing differing pieces of evidence to arrive at an answer to the essential question. In the debrief, Craig explained that his primary intention in the lesson was to introduce the "thinking skill of corroboration." Although Craig required student to note the sourcing information and even led a brief discussion about reliability, he later described sourcing as an "add on...just for practice."

Given this contextualization goal, Craig's lack of attention to important aspects of the sources was not necessarily problematic. He did not intend for students to engage in all types of historical reasoning at once. Instead, he wanted to isolate corroboration as a skill so that students could recognize the thinking they were doing, and hopefully, do it again in later lessons.

While Craig elicited justification of arguments by asking students to point to evidence that supported a claim, this work was almost always done in the context of some other historical thinking work. Across the outset data, then, Craig elicited thinking that challenged students' ideas about history as an authoritative narrative, unwittingly elicited contextualization, and elicited an elementary understanding of corroboration. Both the murder mystery and Industrial Revolution lessons demonstrated an increasing intentionality in terms of eliciting historical thinking.

Interpreting at the outset. In order to analyze interpreting, I limited its definition to 1) what the TC reported noticing and 2) the factors the TC reported considering prior to an instructional response. Based on Craig's explanations of classroom interactions about evidence in history and primary sources, I identified several themes related to Craig's noticing and the factors he considered before responding.

Outset noticing. Craig's noticing was irregular across the outset data. He noticed some student thinking at some points but failed to notice salient student thinking at other points. Of the thinking he noticed, two themes emerged: 1) whether thinking was within the problem space he envisioned for the task and 2) specific instances in which students were struggling in the problem space he created.

As early as the pretest, Craig demonstrated an ability to notice details of student thinking that went beyond general argumentation. For example, Craig noticed that Matt treated the

sources as if they were authoritative accounts about "what happened" rather than sources with particular perspectives. When Craig examined Larissa's example essay, however, he failed to notice key aspects of student thinking including Larissa's back-and-forth movement between present and past tense, complete neglect of one of the documents, and her use of quotes without attribution. Craig's reflection on the Enlightenment lesson demonstrated similar inconsistencies. He described a number of comprehension challenges that students encountered as they read the documents and answered the questions but he was unsure of whether students actually accomplished the primary objective for the lesson.

Craig reflected on the value of the assessment he used: "I am now questioning if this assessment displayed proof of the objective - how these three thinkers transformed European society. This objective was better judged in Question 3 of the packet. As noted previously, Question 3 was the one part of the lesson that few of the students were able to accomplish. Yet, Craig seemed unsure about what the student thinking suggested about his goal.

By the time Craig taught the Industrial Revolution lesson, he demonstrated closer attention to student thinking about evidence in history than in any of the previous lesson. Craig seemed to come into the lesson with a clear vision in his mind of the problem space he wanted to create for the students, and in turn, the type of thinking he wanted to see. This vision for problem space and student thinking formed a backdrop for Craig's noticing in the lesson.

In the debrief, Craig's explanations for what he noticed were clearly aligned with a problem space he envisioned for the task. For example, I asked Craig what he noticed during an exchange in which several students built an argument that the Industrial Revolution was positive by pointing to various portions of the sources:

So, I thought…it was almost piecing together, almost thinking out-loud, right? I think this

is how I would want a student to think through this. And, they … did it together.

At other points in the debrief, Craig described noticing whether students comprehended the

documents, identified whether the source depicted the Industrial Revolution as positive or

negative, and whether the student could explain how the source was evidence for the claim.

When I asked Craig about a time I heard him tell a student, "That is a good hypothesis,"

Craig explained, "I was particularly looking for 'this is positive' or 'this is negative'…and then

could they pull out evidence …he answered the question and gave a reason that I could tell he

got from the documents." Craig's explanation indicates the path he saw toward the problem

space he envisioned for student thinking.

Craig further described noticing whether students were making sense of the tension he

wanted to create between each of the rounds of evidence. For example, Craig noticed during the

second round that several students tried to resolve the problem of conflicting evidence by saying

that the Industrial Revolution was both positive and negative. The entire lesson was an attempt to

bring students into a tension between conflicting accounts of the same phenomenon. So, when

Craig noticed that students were saying, "both" he knew he had them engaged in the tension.

Craig also noticed points where student thinking was not operating in the problem space

he envisioned. For example, in the debrief of the Industrial Revolution lesson, Craig described

noticing that one of his first questions after reading the first round of documents prompted a

student answer that was not what he envisioned. He said, "I remember…[thinking] there's going

to be an issue because of the way I set it up. Everyone's going to agree with it…there is not

going to be a [discussion]." Here Craig noticed that he had a problem in the first round because

the first round of evidence suggested what the students already thought: that the Industrial

Revolution was uniformly positive. Given that his vision was to bring them into a tension that was unresolvable by straightforward thinking, he was alarmed.

Craig also noticed specific instances in which students raised confusions or misconceptions when confronted with the conflicting accounts. For example, he described noticing how some students sought to resolve the problem by "weighing" the quantity of documents that suggested one side of the argument and the other. That is, students wanted to count the documents for and the documents against in order to determine which side to choose. Thus, Craig's outset data told a mixed story on his ability to notice student thinking. By the Industrial Revolution lesson, he tended to notice whether thinking was within the problem space he envisioned for the task, and he noticed misconceptions or confusions related to the problem space.

Factors considered at the outset. Based on Craig's explanations in the early data, the factors he considered before responding were primarily driven by an intention to connect his goals to the everyday thinking of his students and a commitment to maintain the problem space he envisioned for the task.

The first factor that Craig considered in calculating his instructional responses at the outset, and especially in the Industrial Revolution lesson, was how to connect the problem space to an aspect of students' everyday lives. Craig believed that historical thinking was similar to the type of thinking that students did every day as they sought to resolve conflicting stories and tensions within various accounts.

Craig explained that he wanted students to recognize that "You're already doing this. You have the ability to do this…you just don't call it [corroboration]…You don't have to change your mindset because now it's school." Craig did not believe that corroboration was unnatural

229

for students. As such, his instructional responses often leveraged connections to students' everyday thinking.

The second factor that figured in Craig's instructional responses in the Industrial Revolution lesson was his commitment to maintain the problem space he envisioned for the task. Craig had a vision that students would examine evidence for the positive side and rush to that conclusion. Then, they would examine evidence for the negative side and "forget everything they just read," and rush to the negative.

Throughout the lesson, Craig worked toward a tension that would force students to try to make sense of all of the evidence, not just the evidence for one side of the argument. The factor he considered for his responses, then, was how to maintain that tension in a way that would invite student corroboration rather than dismissal of one side or the other.

Craig's decision to frame the task as an "evolving hypothesis" rather than simply a discussion lesson was indicative of the problem space he sought to create and captures the vision that he had for student thinking in the lesson. He explained that he did not want his students to approach the task as, "'This [document] proves me right. This [document] doesn't prove me right, so I'm not going to use it.'…I wanted them to approach it as…'It's an evolving answer, using evidence'."

Craig's commitment to the problem space played out differently as the lesson on the Industrial Revolution proceeded. For example, when students were first confronted with the negative evidence, some students suggested that the Industrial Revolution was both positive and negative. Craig initially refused to accept this answer. As some students began to explain reasons that it could be both positive and negative, however, Craig began to allow the apparent contradiction to remain. When I asked him why he did not initially accept this answer, he

explained, "Because, that's what it comes down to…I want them to get the idea that it's both. But, at a certain depth of understanding…[because] at a shallow understanding …something can't be positive and negative…because if they're saying 'both'…just trying to get them to think, 'What does that mean?'

Craig envisioned bringing students into that tension in order to introduce them to the reality that accounts of the past are just that, accounts and not retellings of reality. Craig's description of what he was up to at one point in the lesson captures the vision that he held of the problem space he wanted to create:

What I was trying to do was…make it an explicit… like yes, it's negative and it's completely different from what we just read and everyone in the class agreed with…I wanted to be like, 'wait…remember 10 minutes ago when you all said it was positive?

Although he gave a show of resisting the students who said, "both," he told me later that when students started to argue both it indicated that they were working in the tension he envisioned for the task.

Despite some irregularity across the outset data, by the Industrial Revolution lesson, factors that impacted Craig's instructional responses included an intention to connect target thinking with the everyday thinking of his students and a commitment to maintain the problem space that he envisioned for the task.

Responding at the outset. In analyzing response patterns, I considered how the TC responded to student thinking when students vocalized thinking related to evidence in history. At the outset, Craig's instructional responses to student thinking about evidence were not consistent. Sometimes he responded with follow-ups that explored student thinking and sometimes he did

the thinking himself, either by building on student reasoning (High I) or by vaguely reformulating a student idea (Medium).

As Craig walked around the room in the Enlightenment lesson, he asked students to explain their answers to him and appeared to listen closely when students answered. As it became clear that students floundered on the third question, which asked them to explain how the political idea could transform European society, Craig responded with Medium responses.

These responses either guided students to a conclusion in a question-and-answer pattern or vaguely reformulated a student comment that allowed Craig to reason it himself. For example, when the Rousseau group told Craig that they did not understand Question 3, he responded:

C: Ok, so it comes from #2. So, in absolutism, where does power come from? Where did kings get their power?

S2:   Divine right.

C: Ok, a divine right. So, in a social contract they say that power comes from the people. That is a huge difference. Because with divine right, who was the only person the king had to answer to?

S2:   God.

C: Now, they are saying that the king has to answer to the people. How would that be different?

SS: (Silence).

C: Are you following me? Let's think of an example. I really like your jacket and I am the king. I have decided I will kill you and take the jacket for myself. In absolutism, I only have to answer to God. So, people come to me and are like why did you do this, I can just say, I don't have to answer to you. Now, flip it, where power comes from the people in

232

the social contract. If the people agree to give me power and I kill you, then they come and say, "Why did you do that?" I have to answer to the people. I can't just do whatever I want. Does that make more sense?

S1:   I think so.

C: We will talk more in the next class.

After presenting the analogy in the exchange above, Craig might have leveraged it to return to student thinking with a prompt like, "So, let's return to the original question. If the understanding of the time about where a ruler's power came from was changing, how could that transform European society?" Instead, he left his own analogous reasoning as the last word. In the lesson on the Industrial Revolution, in contrast, Craig was more responsive. After each round of student examination of evidence, Craig attended to student thinking and followed up with prompts and questions that sometimes pushed the students to articulate their reasoning (High II responses).

Craig's responses aimed to move students through the thinking he envisioned for the lesson: declare an argument and point out evidence from the documents that demonstrated that argument. For example, in the following exchange, Craig pushes students to answer with evidence:

C: Who wants to start us off … this round of evidence? Was it positive or negative?

S: Negative.

C: Give me a reason why.

S: (reads quote from the document)

C: Ok, so how does that prove it is negative?

S: Working conditions were bad.

C: So, the working conditions were really bad. Who can add to that?

S: Some of the living conditions were bad.

C: Living conditions. Not just where they worked but where they lived. What were they?

S: Windows broken stuffed with rags, floors wet… (citing from doc)

At that point, Craig made another High II move by having the students dig into one of the documents and consider the context before returning to the original student thinking:

C: Good. So, picture that. Which document was that again?

S: Document D.

C: Imagine that. Look at Document D for a second. Close your eyes and think about that. Look at the last sentence – kids rolling around "in the filthy moisture of the street oozing up." That is not positive. (Pause) I thought the whole point about factories was that it made life better...

Craig could have simply ended the conversation by asking students to envision the context depicted in the document. Instead, he followed the visualization with a High II response by accentuating the contradiction between the vision of life depicted in the document with the point that the students had already established: the Industrial Revolution was positive.

Craig responded in High II ways at other points in the Industrial Revolution lesson. In the following exchange, Craig pushed student thinking and resisted several opportunities to take over:

C: So, who can read me their revised hypothesis considering all of this? Does this corroborate with first round of evidence?

SS: No.

C: So, give me revised hypothesis.

234

S4:   Overwork, child labor…(inaudible).

C: So, do we agree with S4 that we should forget about the good things in the first round? It is completely negative? Who wants to argue against that?

Instead of taking over the reasoning by reconciling the problematic interpretations himself, in the exchange above Craig invited the students to articulate their hypotheses and their evidence.

As students struggled to explain how multiple documents could suggest different conclusions about the Industrial Revolution, Craig pushed the students to explain their answers. In one of these exchanges, he responded by asking a student to reason through the day's dilemma. After Craig revoiced a student's idea that "it does not have to be perfect to be positive," he noticed the wry smile of one student:

C: S20, you smirked. What is your response to that? Prove me wrong.

S20:   Well the definition of positive is "to be good for everybody." So, in a way, positive does have to be perfect. So, if outcome is perfect then…(inaudible)

C: So, how do we...almost to what you were saying earlier... "if the outcome is positive but the method is negative," what do we do with that?

Here Craig took a risk by revoicing an earlier student statement and asking S20 to explain a challenge to that student reasoning. After S20's explanation, Craig framed the contradiction again, by returning to an earlier student statement. These High II follow-ups maintained a focus on student reasoning.

At various times in the Industrial Revolution lesson, student thinking seemed to go in a different direction than Craig had expected. In these situations, his response patterns varied. For example, in the following exchange, he paused a conversation about negative aspects of the

Industrial Revolution in order to address a contextual misconception that children were not allowed to work:

C: Like are you saying today?

S5:   No like ...(inaudible).

C: …this is a good moment. Let's take a step out here. So, at this time, during the Industrial Revolution, were kids allowed to work in factories?

SS: No.

SS: Yes.

C: Yes they were. And, they did. They did work and the pay went to the family.

S8:   Children were paid… (Cited from document).

C: Good. Yes, they worked for little pay and they worked just as long of hours.

C: You are right, S5. This eventually changes. After they start seeing these problems they start making reforms but at this time, no. Children are working in these factories

C: Why is that negative?

In this apparently unexpected detour, Craig "stepped out" to build context. After clarifying the circumstances of the time period, he linked the contextual diversion to the original discussion about the negative aspects of the Industrial Revolution.

Even in the Industrial Revolution lesson, however, Craig did not always respond to student thinking with High II follow-ups. In some cases, Craig's responded by building his own reasoning on a student idea and even occasionally reformulated a student answer to make a particular point.

As students talked through their arguments after the final round of investigation, Craig responded to one student's idea by cinching the conclusion himself rather than asking the student to further articulate his reasoning:

S3:   Originally, I said that even though products that came from [the Industrial Revolution] was beneficial the process was a problem …

C: So, S3, what you are saying it was good but it is almost like, 'at what cost.' S3, what you got? What do you want to say? (To student with raised hand).

When students struggled during the Enlightenment lesson, Craig especially tended to do the thinking for them.

Craig sometimes responded to student ideas with a High I pattern that affirmed and elaborated, and did the hardest thinking for the students. For example, when one student cited a quote during the first round discussion as evidence that the Industrial Revolution was positive, Craig responded by interpreting the quote for the students rather than bouncing it back to them.

S10:   The immediate effects of this phenomenon… (Quoting from document).

C: So, the last paragraph and sentence. Do you agree that these are positive things? Wealth, industry, population, political influence–Not bad for Great Britain, right? So, do we all still have that it is pretty positive?

By pointing out the location of the quote and then making the inference that connected the quote to a positive claim, Craig reduced the complexity of the thinking the students had to do.

In another exchange in the Industrial Revolution lesson, Craig skipped the chance to ask students to explore thinking about a quote that he identified:

S6:   [The Industrial Revolution was negative because] working conditions were bad and they were working with dangerous machines.

C: Yeah…Look at this picture. He is just sticking his hands in there… can you turn to Document C for a second? In the last paragraph and the last line, it talks about unguarded machinery … Follow along with me. (Read from doc). Now, flip back to the picture. Is that a picture of unguarded machinery? …What happens if his finger gets cut off?

SS: Gets fired.

C: …He gets fired. Why does he get fired?

S: Cause he does not have finger.

C: Cause he can't work anymore…

Although Craig's response focused on S6's idea, Craig moved away from student reasoning. He then identified a quote that supported S6's statement and made a connection to another piece of evidence. The questions that followed did not invite complex reasoning.

At the outset, then, Craig responded to student thinking in an inconsistent manner. Especially in the Enlightenment lesson, but at points in the Industrial Revolution lesson, Craig responded by doing the hard thinking himself. At other points at the outset, Craig responded by exploring student thinking.

Change in eliciting. Throughout the remainder of the study, Craig continued to launch generative tasks. In these later observations he explicitly elicited particular student thinking about evidence including sourcing, contextualization, and corroboration.

Change in generativity. Throughout the remainder of the study, Craig launched generative tasks that positioned students to reason about evidence in history and, to a limited extent, he invited argument building based on interpretation of primary sources. Craig increasingly connected these tasks with students' prior knowledge.

For example, in one lesson I observed, Craig began by asking whether any students remembered what sourcing was from the prior semester. A few students gave some ideas and then Craig said:

> I am going to prove that you already know how to do this .... Let's pretend there was a fight in the lunchroom and you were not there. Of course, you need to know what happened…Or maybe you are more mature than me.

Craig showed four possible sources that students could ask about the fight: a kid in the fight, the security guard, the best friend of a kid in the fight, and a student in another grade. Argument immediately arose among the students about which of four people they would talk to in order to determine what happened in the fight. Craig shouted over them:

C: So this is what is interesting. There are pros and cons for each one. Right? Sourcing, it is not just bad or good. Let's break this down 1 by 1. So, who would go to the kid in this fight?

SS: Yes! No!

C: Who says yes? S1 why would you go to the kid in the fight?

S1:   Because he was in the fight and he knows what happened.

C: …Who says no? S2 why did you say no?

Craig pursued the same student reasoning with similar questions for each of the four possible sources of the fight saying, "Who would ask this person?...Why yes?...Why no?" Each of Craig's elicitations focused on voicing student thinking about the credibility of the source. After having students discuss each of the four potential sources he transitioned to the actual historical work saying: "The point of this is that you all did sourcing in your head like THAT…You know how to do this already."

As Craig transitioned to historical work, he grounded the skill of sourcing in something the students had just demonstrated that they were able to do. Craig then introduced the historical documents he wanted the students to source, saying:

> There are two documents here….[after the reading] there are three questions. Take a look at the questions. One, who does this document blame for WWI and why? Two, do you find the argument credible? … So, you will have to do sourcing to answer that. And, three, what about this source made you say 'yes' or 'no?'

Craig emphasized that the questions about the historical documents were only an academic version of the questions the students had just answered about the fight.

By the end of the lesson, many of Craig's students were discussing the inherent problems with biases in the two sources. As the end of the period was drawing to a close, Craig summarized the problems of credibility that the students identified saying:

C: Does this mean we can't use these? We just throw these documents out?

S10:   No

C: Garbage? Can't use them? (pause)

SS: No!

C: So how are they useful S11?

S11:   If I could look at enough other documents these could be good.

With this question, Craig invited thinking about a new historical reasoning skill, corroboration.

During the debrief, Craig explained that once students were able to identify issues of credibility by examining source information, he believed that he could then have them envision other sources that might help them build a stronger understanding of what happened. Even though the majority of the lesson time focused on an analogy that prepared students to engage in

sourcing, the task that Craig launched was highly generative. By anchoring into the lunchroom fight, Craig positioned students to reason about evidence with actual historical documents and established a pattern of argumentation that transitioned smoothly into the interpretation of primary sources.

On another occasion that I visited Craig's classroom, students were engaged in a three-day unit on totalitarianism in the 1930s. The unit culminated with an essay that required students to pick one of three ideologies (Nazism, Fascism, or Communism) and use evidence from multiple primary source documents to argue whether the ideology counted as totalitarianism.

As I observed the second day of the unit, Craig separated the students into groups that were tasked with studying one of the three ideologies: German Nazism, Italian Fascism, and Russian-Stalinist Communism. Each group examined 2-3 primary sources and then reported out to the class about 1) why or why not the sources were reliable, 2) what the sources said about life under Nazi (fascist, communist) rule in Germany (Italy, Russia), and 3) How that ideology was an example of totalitarianism.

After reviewing some of the terms and ideas that students explored in previous days, Craig launched the central task by saying:

> We are going to interpret what it was like to live under a totalitarian state. The way we are going to do that is to look at documents from each state. … What is it like to actually live in that state, how are they acting in a totalitarian state?

Craig had students think about the context of each regime by using the documents to envision what life was like for people at that time.

The second question on the graphic organizer was intended to get students thinking about life under the regime that each group was assigned. As Craig walked students through the graphic organizer, he said:

> Second question is "What does this source say about life under whatever the ideology is?" That is our objective today. The source will talk about a specific thing that happened in this place…Whatever it is, you are trying to say, what does it mean to live in this system.

During the debrief, Craig explained why he designed the second question as he did:

> …That was more like do you comprehend this document through this lens of what it was like to live that way? So it's not just like a basic question – like for the Nuremburg Laws it's not like, "Oh, what laws did they have?" Because that's just like searching, that's not good thinking… So they're actually interpreting, doing stuff on their own.

By the end of the lesson, students had begun to design an outline for their essay that would use evidence from the day's task to argue that one of the three ideologies was an example of totalitarianism. In these and other tasks Craig launched, he tried to position students to reason about historical evidence and engage in evidence-based argumentation. Based on the data I examined, Craig's facility to launch generative tasks began during the outset of the study and consistently developed throughout the remainder of the study.

Change in thinking elicited. Where Craig's outset data demonstrated an inconsistent focus on historical thinking, in the later observations I saw him elicit sourcing, contextualizing, and corroborating. In some cases, Craig made historical thinking explicit for students and intentionally connected it to students' prior knowledge.

Craig's course-required reflections throughout the fall shed light on the development of his approach to eliciting historical thinking. In a December reflection, Craig described his plans for the coming months, in which he would take over all the responsibilities for teaching the class:

> I am planning to explicitly re-introduce and practice each skill; however, I am planning to do so by isolating each skill…will make it more manageable and possible to build student understanding one skill at a time… Once students gain a better understanding and ability to perform each task, we can then integrate many skills in the same lesson.

These reflections helped provide a running commentary of the development I observed in Craig's approach to eliciting historical thinking. In mid-fall, he was observing teachers in his building working on historical thinking skills and he determined it was "something I will continue to play with." By the time he was taking over teaching, he had determined an intentional direction for historical thinking skills in his classroom.

Craig approached historical thinking skills in isolation, meaning that he taught them as a specific skill first rather than trying to embed them in content. He then used the isolated examples as anchors to return to throughout the semester. For example, at the conclusion of the lunchroom fight discussion he told the students, " So, this is what we will do for each one (of the skills)….When you think sourcing, you will think of this example."

Later in the year, I observed Craig use the lunchroom fight analogy to elicit student thinking about sourcing during other tasks. For example, during the totalitarianism lesson when some students wanted to simply copy and paste the source in the space labeled 'sourcing,' Craig brought the lunchroom fight up to prompt thinking about how the source information figured into an argument.

Despite the fact that sourcing was not directly relevant to the essay students would eventually write on totalitarianism, Craig maintained an emphasis on the importance of recognizing the source and considering its credibility. Craig made this point when discussing sourcing in the debrief: "Sourcing itself didn't really fit into my writing objective, which is using evidence to build arguments…But I was just trying to remind them – where it's like, 'What are you reading?'"

Although Craig decided to explicitly introduce historical thinking in relative isolation from historical content, he did not continue to keep content and historical thinking skills separate. As the sourcing example above demonstrates, Craig used the inclusion of primary source documents as a means to practice historical thinking skills.

Similarly, Craig explicitly taught contextualization with an analogy that amused his students. He asked them to imagine a scenario. In this scenario, the student observed Craig walking down the hallway holding a baby doll. What, he asked them, might someone conclude from observing this phenomenon. Students laughed when someone suggested that perhaps Craig liked to play with dolls, but that was exactly what Craig wanted to hear.

Craig wondered aloud what they would need to know in order to understand what he was doing with that baby doll. In this way, Craig described contextualization as "not seeing any event in a vacuum." Like Craig's corroboration in the Industrial Revolution lesson, contextualization remained elementary in Craig's classroom. When Craig looked back on the year, he acknowledged, "In my mind contextualization was always an explicit priority, but I was not ambitious in my effort to help develop this skill."

Craig's explanation made a distinction between eliciting thinking about context and actually eliciting contextualization:

I always kept considering context as a priority in my class, especially in cases that most

people struggle with, such as the Nazis and other darker moments in the 20th century.

But, on the other hand, I never truly asked students to use contextualization on their

own...I find this skill so difficult. Contextualization requires students to fully understand

a …document, have the background knowledge of what is occurring …, and then the

ability to make a connection between the two….Overall, my approach was to keep it

more teacher centered and … address it when I heard a student make a comment that

ignored context or I felt the topic lent itself to it.

As Craig suggested above, he did more demonstrating contextualization than enabling students

to contextualize. However, a few examples of contextualization are worth noting. Craig related a

task his students completed as they examined the "steps to World War II". Craig selected a

secondary source that depicted the hardships of post World War I Germany including the

widespread suffering, terrible living conditions, humiliation of Versailles, international

repudiation, and domestic political turmoil.

He described prompting students to think contextually, "If you were living in this

scenario and you had two choices for your leader, one is a nice guy in a suit saying, 'Let's follow

Versailles and the other is a man in a military uniform, banging on a table, and swearing that he

will make Germany strong again, which would be most appealing?" This activity was designed

to elicit contextualization, and return to the lesson of Craig and the baby doll, by helping students

recognize that the rise of the Nazism "did not happen in a vacuum."

I observed him return to this contextual anchor in a later lesson. When Craig showed the

clip of Hitler's speech in the lesson on totalitarianism, students debriefed the elements of

totalitarianism. At one point, a student said:

S8:   But, why would people support him if it was not about them and it was just the state?

S10:   Because they were stupid.

C: No, no, no, not that they were stupid. An entire population of people cannot be stupid. Why were people into this? Why was this appealing? Think about what Germany was like before Hitler came along?

S10:   They were going through a tough time.

C: A tough time is an understatement. What was happening?

S3:   They had a war or something.

C: Remember those pictures with the wheelbarrows, hyperinflation, also the war blame, all that. These people are wanting something more, something he promised them.

SS: (Murmuring)

C: Yes, the Nazis took over but there were reasons for people buying into this.

In the midst of a lesson in which students were working to understand what it was like to live in a totalitarian regime, Craig saw an opportunity to anchor back into a task that had demonstrated contextualization.

After the lesson on the Industrial Revolution, I did not observe Craig teach corroboration again. However, Craig described the lunchroom fight as "a good jumping off point to get into corroboration." Although I was not present to observe, Craig related the way he explicitly taught corroboration at another point in the second semester. He described having students imagine that they were "picked up by the cops and taken to an interrogation room to be questioned about a crime." All the students initially said that their first reactions would be to say, "I didn't do it." But, Craig explained that the cops did not believe their story because "they did not think you were credible…so, now what do you do?"

Craig then walked the students through a process of how the police might "try to corroborate" their story with other evidence. Craig then examined the idea of corroboration explicitly in this situation before turning it to a historical investigation by asking, " so when examining historical documents, what would it mean to use corroboration."

I did not observe Craig explicitly elicit corroboration during the later portion of the study because in the lesson on totalitarianism, Craig did not emphasize cross checking the documents. During other interviews, however, Craig cited multiple lessons in which students engaged in cross checking documents to successfully complete a historical task.

During the final interview, Craig described how he began the master's program with a dual commitment to student tasks that promoted enduring understandings and student engagement. Based on his experience with the methods course, Craig said that historical thinking is not the third pillar in his conception of what it means to teach history.

Interpretation changes. Throughout the remaining observations, Craig's patterns of interpretation largely held constant.

Noticing changes. In the later observations, Craig continued to notice student historical thinking in light of his learning targets. Although he recognized that many of his students were unable to engage in this thinking on their own, they could do so when he prompted them in certain ways.

In the later lessons I observed, Craig continued to notice student thinking about evidence in history and compare that to the vision for student thinking that he envisioned for the task he launched. During each of the debriefs, Craig articulated a clear vision for a problem space that he had hoped to create for the students and he cited examples of students who were working within that problem space.

First, Craig continued to notice student thinking about evidence in history and compare that to the student thinking that he envisioned for the task. For example Craig explained that in the lunchroom fight analogy, he had expected students to prefer one source over another, "but many of them were able to see the good part and the bad part [of each source]." Although the lunchroom fight was not even an actual historical example, Craig recognized that the student thinking was operating in the exact space he envisioned for sourcing and, thus, made way for the transition to actual historical sources.

As Craig moved from the fight to the World War I documents, he continued to notice when student thinking was operating within the problem space he envisioned. What he noticed was that even though many students were unable to "sift through the documents…[they] used the source and were able to answer the questions, which is the point." Thus, even though Craig noticed that some students were struggling with the reading, he also noticed that they were able to operate in the problem space he envisioned, which required students to simply recognize that every source had value and every source was problematic.

Similarly, in the lesson on totalitarianism, Craig described noticing when student thinking was in line with what he envisioned for the task. He noticed but was not overly concerned with confusions or disruptions unrelated to students' ability to recognize what it was like to live under a totalitarian regime, based on the documents.

Just as Craig noticed when student thinking was in a good space, he noticed when student thinking was not operating in the problem space he envisioned. In some cases, this included students not accomplishing a learning objective. For example, in the lesson on totalitarianism, Craig noticed that some students did not explain how information from the documents they cited related to claims. He described this as a central matter of concern because without explanations,

students would be unable to engage in the target thinking and unable to write the essay on the third day.

Likewise, Craig noticed threats to his problem space that involved confusions or misconceptions. For example, when Craig recognized during the lunchroom fight lesson that some students thought that the document written by the British was blaming Russia for starting the war, he described his thoughts: "Some students called me over and were confused…some people did not get that [authorship] part and if that's the case, it blows up the whole sourcing thing because it's almost the opposite."

Craig also noticed misconceptions that threatened the problem space. For example, as Craig transitioned from the lunchroom fight to the World War I documents, he said, "You all know how to do it. You just need to know how to do it…(paused as if looking for the word)." When Craig paused and seemed to search for the word that he wanted, a student in the back jokingly finished his sentence, "know how to do it, right?" Craig later told me that he noticed this because "I'm like, no!...They all just did it 'right.'" Students regularly made jokes in Craig's class that Craig either laughed at or ignored but at the point described above, Craig noticed the student's suggestion that their sourcing of the lunchroom fight had not been "right," and saw it as a threat to the thinking he was trying to develop.

The second aspect of student thinking that Craig noticed was that students were frequently unable to engage in historical thinking independently. In both the lunchroom fight and the totalitarianism lesson, Craig described noticing that many students did not engage in sourcing without significant support. Even in his final interview, Craig described recognizing that most students still did not engage in historical thinking independently, especially contextualization.

Even though Craig noticed student inability in this regard, he also noticed that when he prompted them, they were often able to engage in the target thinking. For example, when Craig described the sourcing he observed during the totalitarianism lesson, he explained that most of the students still could not do it without prompting but when he reminded them of the lunchroom fight, many remembered sourcing and were able to attend to the sourcing information in the documents.

Likewise in the totalitarianism lesson, Craig noticed that when he prompted students to explain the information that they pulled from the documents, they frequently were able to do it. When he prompted students to remember the "storm clouds leading up the to World War II," some were able to explain how Nazism might have seemed a legitimate option to many Germans.

Changes in factors considered. The factors that Craig initially reported considering largely held for the remaining lessons. He continued to consider how to connect with student everyday thinking and how to maintain the problem space he envisioned. In the later lessons, he also described viewing target student thinking as if on a trajectory, rather than right or wrong.

The first factor that impacted Craig's instructional responses in later observations was a commitment to accessibility. Like the Industrial Revolution lesson, Craig continued to approach historical thinking, and other challenges in his classroom, as "something you already know how to do." In most examples I observed, this came by way of analogy at the launch of a task.

In the lunchroom fight lesson, for example, Craig described how pleased he was that the students disagreed on who the best source would be for the fight because he then knew, "it's going to be easier to bring up the fact that history is a matter of perspective….so, I'll be able to make that connection." When he made that transition to the World War I documents, he

explicitly identified the thinking that the students had just done and then described how they could transfer that thinking to the documents. Craig told students that they were just going to apply that same thinking to history, as he passed out the documents.

During the debrief after the totalitarianism lesson, Craig continued to talk about "good anchors" to student thinking. "It's a lot how teaching is, right? You're just trying to make stupid or funny or some sort of tangible connection so that when I go back to it, it's something they remember…it's how people think."

Craig described the way his students were "very much… still in the mode where it's like 'I'm in school and this was given to me by a teacher'." By this he meant that even when his students were capable of engaging in thinking such as sourcing, his responses typically had to anchor to something like the fight in order to get them to do it.

The second factor that figured in Craig's responses in the later lesson was a trajectory-view of student historical thinking. Craig accepted that he was unlikely to see ideal historical thinking, so he seemed to frame it on a trajectory, with small steps that indicated what type of thinking that he was aiming to facilitate. For example, in the lesson on the lunchroom fight, Craig calculated that the students did not have to be doing sourcing perfectly or completely to be sourcing.

When Craig launched the lunchroom fight lesson by asking students if anyone knew what sourcing was, one student immediately said "after you paste it, you tell what website you got it from." Craig described his own thought at the time, saying, "I was good with it because at least he saw that sourcing has to do with sources. But, that alone isn't sourcing, the point of sourcing is to use that to make an argument…So, I was happy that he at least saw that…but the word source isn't sourcing."

Craig assessed student thinking toward the end of the lunchroom fight lesson as "a good start." He described the thinking as "kind of what I was going for…but it did not get deeper…than 'this is pure bias'." Even as he acknowledged the limitations of the thinking, he said, "They definitely got sourcing" and went on to describe how he needed to reinforce the skill with future tasks that go beyond simply identifying bias. Similarly in the lesson on totalitarianism, Craig wanted students to contextualize but saw contextualization on a broad spectrum that included simply recognizing a contextual element to using historical knowledge to frame what it was like to live under a totalitarian regime.

Both the first and second factors described above may be partly attributable to the fact that Craig recognized that many of his students had been unsuccessful in academic settings and one of his primary intentions was to convince them that they could do complex thinking. He expected that if he simply presented them with historical documents and tried to get them to engage in historical thinking many of his students would shut down. So, he convinced himself and the kids that they already knew how to do this thinking: "They just need to be able to do it in school." Furthermore, he had to interpret student thinking on a broad spectrum that could include small steps toward progress rather than see target thinking as all-or-nothing.

A third factor that Craig considered in formulating his response was how to avoid distractions from the problem space. Craig's intention to maintain the problem space required a clear focus, a focus that he sought to maintain in order to not get pulled in many directions. For example, at the point in the lunchroom fight when a student suggested, "We need the multiple sources," Craig described thinking "he knew corroboration already [but] I did not want to get into that."

Even though Craig recognized that "this would be a good jumping off point to get into corroboration" he determined, "not yet." Craig described his feeling at the time, "I don't want him to confuse everyone else because I want to introduce each of these [skills] independently before we start mixing them together." Craig's explanation demonstrated his commitment to maintain the problem space he had established, even in the face of a generative student comment.

At another point, Craig avoided a distraction that many teachers might consider fundamental to the lesson: reading comprehension. Craig described a student reading problem that he did not see as a threat to his problem space. In the exchange, the student explained that the second document was "about mobilization." Craig brushed off the student comment saying, "It is about mobilization, but who wrote the 2nd document? …What government?" Eventually the students identified that it was written by the British government and the discourse proceeded to the reliability of the document. Craig's explained during the debrief:

> I don't think students really got into the documents … but I almost didn't want them to. I was trying to find the least complicated, least issue-based document to show sourcing on this question. If I could have just found a newspaper that says, "It's all Russia's fault" I would have used that.

Craig's explanation suggested that the student confusion about the content of the document was only relevant in as much as they could understand the basic argument and the author.

When one student in the totalitarianism lesson raised a question about why Hitler hated the Jews, Craig sidestepped it and promised her that he would return to it on another day. He told me after:

> I think it is great historical thinking of why….but I could not answer that question within that lesson...because we were looking at, "What's it like to live in a totalitarian

state?"....To answer [her] question would require so much depth and background and it's

…delicate.

Instead of answering her question, he decided to put it on hold for another day and return to the problem space he had created, student analysis of life under totalitarian regimes.

In other cases, Craig described responding to student thinking, not because of a distraction threat, but because of a confusion or misconception that threatened the problem space. For example, When Craig heard the student say that Hitler came to power because the "Germans were stupid," I thought it was a similar moment to the previous comment in which a student had asked about Hitler and the Jews. However, Craig explained later that the student thinking was more relevant to his goal for the day (understanding how Nazism was totalitarian). So, he chose to address the presentist misconception that people from that time were stupid. He asked them to consider what was happening at the time as a way to contextualize how this could have happened. During the debrief, he explained:

The student that made that comment was like straight against what I wanted them to do. So I needed to stop…where the girl, her comment was more like a question and she even displayed the positive side of it. So I was almost like, okay, hold on to that question.

In these ways, Craig described trying to maintain the problem space by offering enough support that the confusion did not threaten the students' ability to work in the problem space.

In the later lessons that I observed, Craig maintained a clear focus on the 'workspace,' and aggressively addressed confusions or distractions to the primary student thinking he aimed to facilitate, even more so than at the outset.

Change in Responding. Over the course of the remainder of the study, Craig demonstrated patterns of response that initially probed student thinking but frequently resulted in

Craig's reasoning. For example, in the lunchroom fight lesson, Craig's probed student ideas initially but could not often sustain discourse with his responses. This pattern began when the students were discussing the four potential sources of information about the fight. Craig explored student thinking about each of the four potential sources by first asking students whom they would trust. Then, he asked one student to explain why that source might be more credible than the others. Once Craig restated the basic argument, he turned to a student who disagreed and asked for their explanation. With the first three sources, however, he did not press the thinking beyond highlighting it for the class.

The security guard source elicited such enthusiasm that Craig departed from the established pattern and let students argue before he stepped in to summarize:

C: Ok, how about the security guard

SS: No...No...Yes…no! (Talking, shouting).

C: Interesting. The other class said 'no' too. Why not?

SS: (Several students talking all at once).

S3:   I say yes…

S4:   They don't know anything.

S5:   If you hang out in the office and get to know them, they will tell you everything.

SS: (more arguing).

C: I heard a lot of things. One was that maybe they don't want to give you all the
        information. Maybe he did not even notice that they were mouthing and pushing all
        period. But, on the other side, it is his job to figure this out.

Admittedly, Craig was engaged in responses that were not actually exploring an historical event. Some might argue that it should not, therefore, count as responding to thinking about historical

evidence. However, given that the discourse is directly analogous to discourse about the reliability of historical sources, I deemed this conversation wholly relevant to student thinking about the concept of evidence in historical investigation.

Significantly, these exchanges demonstrated a pattern evident in Craig's instructional responses in the later lessons. Craig often initially probed student answers but did not sustain discourse with those responses. That is, he initially responded to student ideas with a High II response. If the student responded with the point that Craig seemed to expect, he typically moved to another line of reasoning or built on the student reasoning himself. Although the questions and probes were often focused on substantive thinking rather than factual recall, Craig rarely sustained student discourse for more than a few back-and-forths.

This pattern showed up in the lunchroom fight lesson when the students transitioned to the World War I documents. For example:

C: S5, you had the first document right? Whose fault was WWI according to this document?

S5:   Russia because...(inaudible).

C: So S6, did you find that to be a credible source?

S6:   No.

C: Good. Why not?

S6:   Because the article is from a biased opinion and it is ....1914 (inaudible).

C: …S6 said no it is not because Germany was on the opposite side of the war. Right? They
        were fighting Russia. So S6, he made a good point. When was it written?

SS: (Murmuring).

C: Yeah, [the war] just started. So, at this point, Germany and Russia are actively fighting and killing each other. Why would you want to point the finger at Russia if you are Germany?

S3:   You don't want to get blamed.

C: Right, you don't want to get blamed. How about the second [document]?

Here was something that resembled revoicing of student ideas. However, Craig controlled the discourse through evaluation and a teacher-student exchange pattern. He took the student thinking about the date of the document and used it to build the reasoning himself in a High I response.

This pattern was an important part of Craig's fast-paced teaching style, a style that worked for keeping his students engaged. But, his frequent responses set up a pattern that commonly resulted in his doing much of the difficult reasoning, lest the discourse ground to a halt. Craig often followed a pattern of questioning and explaining, apparently intended to arrive at a target understanding, although not necessarily a right answer. Frequently this involved Craig's initial question, follow-up questions or probes and eventual unpacking by Craig, after a student had partially delivered an idea. Although engaging, this pattern limited interpretations and shepherded students toward Craig's target.

Craig's tendency to reason himself based on an underdeveloped student idea was evidenced in several of the rounds of discussion about the lunchroom fight. In the exchange below, he built on a nascent student idea:

C: What about the friend of the kid in the fight?

SS: Yes…no…yes.

C: Why yes?

S: Because if he is his friend then he was with him.

C: Ok, so he was likely there. He talked to the kid after. Maybe he does not have as much stake in who won.

C: Why no?

S: That is his best friend and he is like, you know…

C: That is his best friend and he is trying to support him. Good.

I could not help wondering if the students who participated in this exchange meant exactly what Craig seemed to want them to mean. It was impossible to know whether these students, or others in the room, were following the reasoning because Craig did not press their reasoning all the way to an explanation of reliability.

The pattern extended to exchanges beyond the allegory of the fight. For example, when the class discussed sourcing as an explicit skill, Craig asked:

C: Why would it matter when it (a document) was written?"

SS: (No response).

C: Why would it matter when it was written? (pause) If we are looking at a document for WWI … would we rather have a diary from 1916 or from 1960.

S1:   1916.

S2:   During the war.

C: Why would you rather during the war?

S: Because (inaudible).

C: Yeah, what did you eat for breakfast this morning? Think to yourself. Pop tarts? What did you eat for breakfast 4 years ago?

S: I don't know.

C: You're probably not sure. That is why it matters… Overall what are we trying to get to?

        … Is this source believable?

In this conversation, Craig shifted between a concrete example of sourcing's relevance and an

analogy to emphasize his point. While he invited and built on student thinking, Craig made the

inferences for the students instead of asking them to make those inferences. Had he asked the

students to apply that analogy to the World War I documents, the exchange could have been a

High II follow-up.

        In the lesson on totalitarianism, Craig continued to respond with mostly High I responses.

Typically, Craig anchored to an aspect of the student idea associated with his goals and then did

the thinking for the students. For example, at a point in the lesson when Craig was reminding

students to source the documents, he pressed them about why sourcing was important. After a

student suggested that the author of the document should be noted when sourcing, Craig

responded:

C: Person who wrote the document? Maybe, What else might you put there?

S4:   When it was made.

C: When. So, we could do who, when (wrote on board)….What are you trying to get to?

S3:   Legitimacy?

C: Legitimacy, maybe. What skill is this called?

S5:   Sourcing.

C: Do you remember that, sourcing? Remember that S11? Remember the fight in the

        cafeteria?

SS: (Laughter). We do.

C: What was the deal? Remember there was a fight and you wanted to figure out the story,

so you go around and …

S7:   Are you talking about credibility?

C: Yes, that is exactly what I am talking about. So, what is the point of asking these

questions?

S7/S4: Credibility of source.

C: Credibility (wrote on board). That is the key. So, in the source, especially if it is a

primary document, put down any information you can find…the name of the person who

wrote it…What type of source is it? Is it a speech, a letter…Whatever you have, put that

with the source…

Craig anchored to the students' prior understanding of sourcing. But, rather than allowing

students to unpack "credibility" once he made the connection, Craig unpacked it for them.

Although students immediately remembered the fight, it was not clear whether they actually

understood the relationship between sourcing and credibility as it related to the documents on

totalitarianism. A more responsive move might have been to turn to one of the documents and

ask a student to reason through its credibility.

When the class moved on to discussing the various sources, one student pointed out that

Mussolini was dressed in a military uniform. This offered a chance for Craig to press for why

that mattered. But, instead he seized the thinking and built on it: "Military uniform. Military is

big in fascism. He is like a general leading the state. You see a lot of flags. The eagle. How are

people acting?" Some might argue that Craig response above is an example of modeling the

required thinking. However, modeling in a way that would keep student thinking the focal point

would have required Craig to explicitly articulate a meta-conversation about his thinking and then return to the student idea in a way that allowed the students to follow suit.

On the whole, the later data suggest that Craig's responses often began as High II attempts but became High I as discourse preceded. That is, while he initially explored student thinking, he eventually took over the reasoning himself. As compared to the earliest data, it seems that Craig's responding changed in that he attended more closely to students' ideas as the year progressed, whether he chose to allow students to build on those ideas or not.

Summary of Findings

In this chapter, I have considered how Craig elicited, interpreted, and responded to student thinking about historical evidence. My first research question asked how Craig engaged in these practices at the outset of the study. The earlier data indicates that he sometimes was able to envision and launch generative tasks and in these tasks he elicited thinking about justification, contextualization, and corroboration. He tended to notice whether students were engaged in the problem space that he envisioned and was able to identify problems they faced as they worked in that problem space. As he monitored students, he sought to connect the target thinking to their everyday lives and to maintain a focus on the problem space he envisioned. His responses at the outset were inconsistently focused on building student reasoning.

My second research question asked how Craig's capabilities to engage in these practices changed during the study. In the later lessons, Craig's tasks were generative and he elicited sourcing, contextualization, and corroboration. Craig continued to notice historical thinking in his classroom and began describing that thinking as if on a trajectory from beginner to mastery. Because he felt that all his students could engage in the work, he focused on how to avoid confusions or distractions that would complicate the 'workspace.' He often responded to student

thinking in the later lessons with an initial High II follow-up and then moved to High I as

discourse continued.

Chapter 9: Cross-case Analysis and Conclusions

In this dissertation study, I set out to explore teacher candidates' eliciting, interpreting, and responding (EIR) during a history teacher preparation program. In conjunction with a history methods course instructor, I redesigned a methods course based on a core practice approach intended to develop EIR abilities specific to historical thinking. To explore how teacher candidates (TCs) understood, enacted, and later adapted EIR, I studied four TCs who each represented degrees of disciplinary understanding and placement context amenability. In the case findings, I considered the individual experiences of each TC as they learned to teach history.

In this final chapter, I first look across these four individuals to illuminate cross-case findings, or themes, that were apparent when I considered the four cases together. I then describe several conclusions and consider the conceptual framework that guided this study. Finally, I indicate a number of implications of the findings of this study for teacher education practice and for future research.

Cross-case Analysis

To this point in the study, I have considered only individual examples of the phenomenon of interest in this study. In what follows, I look back on each candidate and compare and contrast the ways that they elicited, interpreted, and responded to historical thinking. In doing so, I articulate the major patterns that I noticed across the data considered in this study.

Cross-case eliciting. In each of the individual cases, I considered the generativity of the tasks that TCs launched and the type of thinking that they elicited during the task. Below I describe patterns that suggest how TCs engaged in eliciting at the outset and in their later lessons.

Generativity. When I look across the four cases and consider the generativity of the tasks that the TCs were able to launch, three findings are apparent. First, TCs made adaptations to course IA structures that varied in generativity. Second, even when a TC designed a generative task, some were not able to launch that task in generative ways. Third, all four TCs continued to use primary sources and methods course IA structures in their lead teaching[8], but only some used these to launch generative tasks.

Adaptations and generativity. Even when TCs implemented a task designed in the methods course to promote interpretation of evidence and historical reasoning, they made adaptations to the design that resulted in varying levels of generativity. Prior to teaching the course-required discussion lesson, for example, three of the TCs (Craig, Kendra, and Sally) adapted a discussion IA to meet the perceived needs of their classrooms. Later lessons demonstrated similar adaptations to activity structures that had been introduced in the methods course.

Some of the TCs, however, adapted activity structures in more generative ways than others. It seemed that successfully adapting the IAs depended, in part, on a TC's ability to first envision a generative task and then consider how adaptations might impact the generativity of that task. As TCs moved into lead teaching, the ability to adapt an IA in generative ways seemed closely related to an ability to create generative tasks for lessons.

This was clearest in the case of Sally who, throughout the course of the study, adapted IA structures in such as way as to preclude generativity of the tasks. For example, in the New Deal lesson, she scaled back the consensus portion of the structured academic controversy SAC, and in so doing, eliminated a need to corroborate across the documents and determine a reasonable

---

[8] Lead teaching refers to the point in the internship that TCs took over all aspects of planning and instruction. Lead teaching occurred after the conclusion of the methods course.

conclusion based on all the available evidence. With the exception of Gabby's outset lesson, Craig and Gabby both adapted IA structures in ways that maintained generative designs. Kendra's early adaptations, as with the graphic organizer in the Lexington Green lesson, were founded on her own misunderstandings. In her later lessons, such as the Federalist 10 lesson, she adapted in ways that reduced the complexity but also the generativity of the tasks.

The relevant literature suggests that the TCs' adaptations of conceptual and practical tools learned in methods courses were typical (Grossman, Smagorinsky, & Valencia, 1999; Johnson, Thompson, Smagorinsky, & Fry, 2013; Thompson et al., 2013). TCs adapt tools for a variety of reasons and these adaptations are influenced by a variety of factors. The studies suggest that chief among these factors are the social context of learning, including the communities of practice with which the TC identifies and individual characteristics of the learner (Grossman et al., 1999; Thompson et al., 2013). If one considers the IAs as practical tools provided in the methods course, practical tools grounded in conceptual tools that were taught in the methods course, then it is possible to partly attribute the generativity of the TCs' tasks at the outset to aspects of individual characteristics and social contexts of each learner.

One individual characteristic previously noted for the TCs was the importance of disciplinary understanding of history. It is striking that the two TCs with strong disciplinary understanding (Gabby and Craig) were able to adapt the IAs in ways that enabled students to engaging in historical discourse; Craig's lesson on totalitarianism and Gabby's 1920s SAC lesson are prime examples. Sally, on the other hand, a TC with lower disciplinary understanding, adapted the IAs in ways that interfered with historical discourse, reducing the activity to little more than sorting quotes. Kendra's limited disciplinary knowledge almost certainly was a factor in her discussion lesson troubles but this factor appeared to act in conjunction with the context of

her placement. That is, the expectations of Kendra's mentor played a significant role in the adaptations she made to the IA in her discussion lesson, although Kendra did not sufficiently grasp the lessons' purpose. In later lessons, Kendra appeared able (and willing) to launch generative tasks when using her mentor's materials but, when faced with designing and launching tasks on her own, these tasks lacked an intentional focus on argument building based on interpretation.

Social context also appeared to play a significant role in Sally's adaptations in that Sally's mentor and planning team seemed always to scale back the interpretive aspects of instructional tasks. A pervasive sense that most of the students were not capable of doing difficult thinking, reading, and writing pervaded Sally's descriptions of conversations with her colleagues. Craig reported observing teachers in his placement who taught lessons that positioned students to reason about evidence and build historical arguments, even though such tasks were rare for his mentor. Furthermore, teachers in Craig's school regularly discussed the challenges of a new curriculum and ways for approaching historical reasoning. Gabby, on the other hand, created one task after another that promoted student interpretation of primary documents and argument building without the support of a mentor who valued primary sources or historical discourse.

Necessary but not sufficient. Even when TCs created or adapted tasks in ways that maintained generativity, some TCs were unable to launch the task in generative ways. At the outset, only Craig was able to launch one or more generative tasks successfully without significant complications, as he did in the Industrial Revolution lesson (although not in the Enlightenment lesson). Both Gabby and Kendra attempted to launch tasks with generative designs but these did not result in generative tasks, at least initially. Gabby's robber barons and

captains of industry lesson turned around as Gabby checked in with each group individually and explained the differences between a quote and an explanation. Kendra's lesson on Lexington Green never turned around, and Sally did not envision a generative task in the first place.

As TCs began lead teaching, Craig continued to launch generative tasks and Gabby launched generative tasks every time I visited. Kendra, on the other hand, regularly envisioned generative tasks but struggled to launch them successfully. In these cases, then, an ability to design a generative task was related but distinct from actually launching those tasks. The literature is filled with examples of TCs who were unable to implement lessons in their internships that they designed in their methods courses (e.g., Feiman-Nemser & Buchmann, 1985). While Sally, and to some extent Kendra, demonstrated such challenges, Gabby and Craig were not unlike Monte-Sano and Budano's (2011) novice teachers who were able to envision generative tasks and launch those tasks in secondary classrooms.

Again, one can consider TCs' individual characteristics as potential explanations for ability to launch generative tasks. Beyond disciplinary understandings that distinguished between these TCs, other individual factors likely played a role in their abilities to launch successfully the tasks they designed. One might consider that although all of these TCs were novice history teachers, each was perhaps less experienced in some of the required competencies than in others. Although Craig and Gabby were novices, they appeared to not be novices in historical discourse when compared to Sally and Kendra. Presumably, something in Craig's and Gabby's experience prior to the methods course contributed to their ability to manage the complexities of historical discourse in their placement classrooms. Exactly what contributed, or what combination of factors contributed, to make these TCs 'less novice' than Sally and Kendra is, unfortunately, a question beyond the available data.

Primary sources and IAs. All four TCs used primary sources and methods course IAs in their outset lessons and continued to use these resources in their lead teaching, but only some used these to launch generative tasks. The fact that all TCs continued regularly to draw on primary sources and methods course IA structures as central resources of history instruction contributed to the possibility for generative tasks, whether those tasks were launched in generative ways or not.

Of course, the ways that the TCs used primary sources and IA structures were key determinates of the generativity of the tasks they launched. Where Gabby recognized the type of discourse that the IA and primary sources were intended to produce, Sally seemed only to value the form of the IA and the engaging nature of primary sources. That is, Gabby regularly used primary sources and IAs (particularly the SAC) to facilitate challenging discourse about historical problems. Whereas, the use of primary sources and course IAs for Sally seemed to just extend a focus on general argumentation and content, as in the Brown v. Board SAC.

Craig, on the other hand, saw primary sources, and to a lesser extent, the methods course IAs as a way to introduce his students to a type of discourse he believed they had not experienced in academic settings but would be able to accomplish. Kendra seemed to recognize the potential value of primary sources and the possibility of the IAs for promoting rich discourse but temporarily abandoned these commitments when faced with the challenges her students presented, such as reading comprehension difficulties.

I did not expect the frequency with which TCs used primary sources in their lead teaching because the research on secondary history instruction suggests that the regular use of primary sources remains anomalous in history classrooms (Fickel, 2006; Grant, 2003; 2006; Van Hoever, 2006). What is more, research suggests that even when teachers include primary

268

sources, many use them in ways that are simply extensions of the textbooks, more intended for student engagement than historical discourse (Hick, Doolittle, & Lee, 2004). While I observed instances of such use of primary sources by Sally, and on one occasion by Kendra, at least two TCs in this study leveraged primary sources to facilitate historical thinking in their classrooms.

The fact that TCs continued to use methods course IA structures suggests that TCs at least valued the form of the methods course IAs, even if they did not master those forms (or even understand them). This finding offers a second point of departure from much of the literature on TCs. Research suggests that TCs are more influenced by their placement contexts than by methods course instruction (McIntyre, Byrd, & Foxx, 1996) and sometimes even teach in ways contrary to understandings they demonstrated in their methods courses (Feiman-Nemser & Buchmann, 1985). TCs in this study regularly drew on the methods course IAs, sometimes even when those activity structures were directly at odds with the established norms of the placement context (as in the case of Gabby and Sally).

Only Kendra reported a placement context in which the mentor teacher facilitated the type of discourse the methods course IAs were designed to support. Yet, all four TCs used task structures that at least had the potential to elicit historical thinking. In the case of Gabby, Craig, and Kendra, these activity structures functioned to support the eliciting of historical thinking to varying degrees.

Type of Thinking Elicited. A second major focus of my analysis included the specific type of historical thinking the TCs elicited. When looking across that type of thinking, I found that 1) two of the TCs elicited historical thinking more often than the others and 2) some of the TCs considered justification a reasonable goal.

Variation in eliciting historical thinking. First, the TCs varied in the type and frequency of historical thinking they elicited. Three of these TCs were able to elicit historical thinking when provided with support from the methods course during the course-required IAs. Craig incidentally elicited contextualization in the Enlightenment Lesson and intentionally elicited corroboration in the Industrial Revolution Lesson. Despite the confusion that marked Kendra's lesson on Lexington Green, there was no shortage of historical thinking elicited. Gabby intentionally chose not to elicit historical thinking with her students in the robber barons and captains of industry lesson, one of the first times she taught in her new placement. Only Sally did not elicit historical thinking in the course-required lesson.

The same three TCs continued to elicit historical thinking during the later lessons, some more consistently than others. Gabby elicited some type of historical thinking in every later lesson I observed but typically focused on contextualization. She seemed to prefer to model historical thinking and offer questions and prompts to provoke it rather than teach it explicitly. Craig's approach was just the opposite in that he facilitated historical thinking in highly explicit and student-relevant ways. His treatment of sourcing, and to a lesser extent contextualization and corroboration, included explicit attempts to teach historical thinking skills. I observed Kendra elicit sourcing and contextualization in later lessons but she tended to withdraw her expectations when challenges arose. I never observed Sally explicitly ask her students to engage in sourcing, contextualization, or corroboration.

The findings demonstrate that three of the TCs were able to elicit historical thinking, first when provided with support from the methods course, and later independently. However, the findings also underscore the reality that more support might have helped Kendra and Sally elicit historical thinking in ways that could have benefited their students. As it was, historical

reasoning was either unattainable for the students (as it sometimes was in Kendra's class) or inconceivable for the teacher (as it was for Sally). Similar to Monte-Sano's (2011a) novice teachers, three of the TCs in this study were able to elicit historical thinking during classroom teaching.

Historical thinking optional. Second, some of the TCs considered justification a sufficient goal and viewed historical thinking as optional, even in the course-required discussion lesson. Both Sally and Gabby openly acknowledged that they did not intend to elicit historical reasoning, beyond justification, in their discussion lessons. This obscured the reality that Gabby's choice was strategic while Sally's was apparently incidental. As Gabby transitioned into lead teaching, she dramatically increased the expectations of historical thinking in her classroom. Sally's teaching and debriefs, on the other hand, continued to focus on justification.

Craig's later lessons demonstrated increasing attention to historical reasoning and I did not observe him focus on justification isolated from an authentic historical task. While later lessons demonstrated Kendra's commitment to historical thinking, sometimes she settled for tasks that required only justification.

Given the TCs' predilection for justification, it is worth considering again the difference in justification and historical reasoning. A teacher elicits justification when prompting students to support a statement with facts or arguments that come from a source text. This might include requests for direct quotes, paraphrases of text, or references to a particular document (Nokes, Dole, & Hacker, 2007). Justification is a part of historical thinking in as much as the student does one of these actions while interpreting that source in light of the sourcing information (sourcing), the contextual information (contextualizing), and/or while corroborating that source against another source (corroborating). Justification without these additions is simply general

argumentation, a skill required for historical thinking but not accomplishing historical thinking. Given that justification is so intertwined with historical thinking, it is not surprising that at least three of the TCs, not only Sally, often confused justification for historical thinking. As such, sometimes attempts to elicit historical thinking (always with Sally), were actually attempts to elicit justification alone.

At present, this distinction between justification and historical thinking is not clear across the literature. For example, Nokes, Dole, and Hacker's (2007) analysis of student historical reasoning included a section on "using documents as evidence," but did not require this use of documents to include historical heuristics. Nokes' (2012) most recent work continues to allow a distinction between general argumentation with evidence and historical argumentation to remain cloudy.

While some researchers make a clear distinction between general argumentation and historical thinking in their analytic methods (e.g., De La Paz, Ferreti, Wisinger, Yee, & MacAruther, 2012; De La Paz et al., 2014), this key difference remains unclear in much of the literature. Monte-Sano et al.'s (2014a) practitioner-oriented work emphasizes this distinction and provides materials to support teachers in the promotion of historical thinking, not only justification. Even with the distinction clear, I have not identified a study that explored an attempt to support teachers or TCs in noticing and addressing the difference between justification and historical thinking.

Cross-case Interpreting

With the delineation of interpreting, I sought to link TC eliciting and responding. Below I describe a number of themes I recognized according to the two analytic categories I used to define interpreting: noticing and factors considered for response.

Noticing. Three of the TCs noticed student thinking about evidence and most of the thinking that TCs noticed was relevant to a learning objective the TC had for the lesson. Noticing historical thinking. Three of the four TCs discussed noticing student thinking about evidence. In their debriefs, Craig and Gabby increasingly noticed it in later lessons while Kendra noticed it less frequently in later lessons. Sally hardly noticed it during the course of the study.

Three of the four TCs discussed noticing student thinking about evidence in the course-required discussion lesson. Although Craig's noticing at the outset was inconsistent, his later lessons, such as the lunchroom fight and totalitarianism lessons, demonstrated that he was able to notice historical thinking. Gabby also reported more noticing of student thinking about evidence as the year progressed. After Kendra's initial noticing binge, she continued to notice student thinking, but not always student thinking about evidence. Although I debriefed two observations with her after the Shay's Rebellion lesson, she did not bring up student historical thinking again as something she noticed. Only Sally seemed not to notice student thinking about evidence beyond students' identification of quotes that matched a particular argument.

These findings are significant, first, in that they demonstrate that these TCs attended to student performance. Research literature suggests that TCs and novice teachers focus attention on their own performance and their own thinking rather than that of students (Kagan, 1992). The TCs in this study noticed many of the things that typical novices notice including classroom management issues, time shortages, pressures from mentor or supervisor, and anxiety about student participation. However, all of the TCs in this study also noticed key student performance metrics and in the debriefs focused on specific aspects of student performance that they noticed.

More noteworthy, three of the four TCs described noticing student historical thinking, a complex form of student performance. Given that many researchers have argued that novice

teachers cannot pay attention to student thinking (Jacobs, Franke, Carpenter, Levi, & Battey, 2007; Shavelson, 2006; Sherin & Han, 2004), the findings of this study join other studies that provide evidence that some TCs are able to notice students' disciplinary thinking (Coffey, Edwards & Finkelstein, 2010; Grossman, 1992; Kazemi et al., 2009; Levin et al., 2009Sleep & Boerst, 2012; Thompson et al., 2013).

Like Monte-Sano's (2011a) and Monte-Sano and Budano's (2013) novice teachers, three of the TCs in this study reported noticing student historical thinking in classroom interactions, not only in decontextualized settings (c.f., Barton, McCully, & Marks, 2004; 2009; Seixas, 1994). The fact that some TCs in this study were able to notice student historical thinking while teaching bolsters claims from researchers in other subject areas that similarly demonstrated evidence that some TCs can notice complex aspects of student thinking in instructional interactions when provided with supports (e.g., Coffey, Edwards & Finkelstein, 2010; Kazemi et al., 2009; Levin et al., 2009; Levin & Richards, 2010; Singer-Gabella et al., 2015; Thompson et al., 2013). Craig, Gabby, and, to a lesser extent, Kendra, share similarities with novice history teachers in other studies who were able to notice key historical thinking (Monte-Sano, 2011a; Monte-Sano & Budano, 2013).

Goals impact noticing. Most of the student thinking that TCs noticed was relevant to the objectives they had for their lessons. TCs were 'tuned in' to student thinking they hoped to facilitate, and they tended to reference thinking that either exemplified or misrepresented their objectives when they discussed student thinking in the debrief. In those debriefs, TCs rarely mentioned historical thinking they noticed that was outside the scope of their objectives, without my prompting.

Sally was the least complex example of this phenomenon in that she simply envisioned a two-part analysis in which students would match quotes to claims. Craig, on the other hand, noticed whether students were grappling within the more complex historical problem space he created for students. Similarly, Gabby's noticing was often directed by her sustained attention on contextualization. One explanation for Kendra's decrease in noticing historical thinking is that she later prioritized other aspects of student performance she felt needed to happen prior to historical thinking.

The fact that TCs tended to notice student thinking relevant to lesson objectives suggests that noticing during actual classroom teaching is quite different from noticing in controlled environments, such as the pretest. In such controlled environments, TCs examined student thinking in a relative vacuum and had no teaching objective to guide their noticing, other than identifying salient elements in the essay or comments.

Noticing student thinking during classroom teaching, for these TCs, was apparently quite different from the pretest environment. For the most part, the student historical thinking that they noticed was thinking relevant to an objective they brought to the lesson. TCs were motivated to notice the thinking they wanted to see in the task, whether it was historical thinking or not.

Factors considered. Across the cases, I identified two themes related to the factors TCs considered for response: 1) objectives for student performance, and 2) characterization of student thinking.

Objectives for student performance**.** Just as objectives drove TC noticing, objectives for student performance were important factors that TCs considered for instructional responses. At the outset, Craig, Gabby, and Sally all had a clear sense of the problem space they were attempting to create, and a primary factor they considered involved how to focus students on that

problem space. Only Kendra seemed lost, at the outset, in the many factors she considered. As she proceeded through the year, however, she increasingly allowed her lesson outcomes to serve as the primary factor that would determine her responses to student ideas.

While objectives for student thinking served to help TCs in many ways, the way these objectives factored into TCs' response calculations may have also been detrimental to the advancement of student ideas and TC contemplation of these ideas. That is, the target thinking that TCs had in mind was often such a strong factor in determining their responses that it may have gotten in the way of understanding student thinking, even for those TCs whose responses appeared to value student thinking.

Although Sally was an extreme case of "How can I get them to see?", all of the TCs seemed to operate in a tension between understanding student thinking and getting students to a target thinking objective. This dilemma between "How do I get students to my objective?" and "How can I understand their thinking?" was not something that the TCs regularly articulated but seemed to run like an unspoken thread through the factors that Kendra, Gabby, and Craig weighed in their response considerations.

What is, perhaps, worthy of attention here is that none of the TCs regularly spoke of a focus on historical facts or narratives as primary factors they considered in calculating their responses. If secondary school history is typically an institutionalized exercise in memorizing a large number of historical facts (Cuban, 1991; Lee, 2005; VanSledright, 2008) and successful history students are those who can master an agreed-upon body of facts and narratives (Bain, 2005; Seixas, 1996; Wineburg, 2001), then these TCs appear to be irregular in their approach to instruction. That is, the objectives that were primary factors in their responses appeared to be

consistently different than the objectives of history teachers depicted in the literature, who focus on what students can recall about the content rather than what they can do with the content.

Characterization of thinking. The second theme I recognized across cases was related to the first: when TCs noticed student thinking misaligned with their objective, the way the TC characterized that thinking was an important factor that helped them calculate a response. Sally seemed to characterize thinking as complete or incomplete based on the two-part analysis and then determined a means to "help the students see" how to complete the analysis.

Craig, on the other hand, seemed increasingly to consider student historical thinking on a trajectory as the study proceeded, which led him to consider how to clear distractions from the problem space so that students could engage in the thinking he envisioned. For example, he viewed a student's suggestion that sourcing was copying and pasting a source as a stop on the way to understanding sourcing, rather than a wrong answer.

When Kendra noticed thinking misaligned with her objectives at the outset, she did not know what to do. In later lessons, she determined that the task needed to be broken into discrete steps. Gabby was almost opposite to Kendra in this regard in that she always expected capability and simply weighed how to best prompt them to successfully engage in the target thinking.

Responding

When looking across the four TCs responding patterns, I found that 1) three of the four TCs demonstrated at least some examples of High II responses, 2) all four of the TCs frequently exhibited a tendency to demonstrate their own reasoning in High I and Medium responses, and 3) only Sally consistently used Low responses.

High II responses. Three of the four TCs demonstrated at least some examples of High II responses, or those that made student reasoning the focal point of the discourse. At the outset,

Kendra was the most ambitious in her use of High II responses but she was unable to manage the complexity of student thinking. This may have contributed to her decision to use High II responses less frequently in later lessons. Although at the outset Craig used High II responses inconsistently; in the later lessons, he frequently pressed student thinking as an initial response to student ideas. While Gabby rarely responded in ways that explored student thinking in detail, Sally almost never did.

The inconsistent nature of High II response across the candidates is not surprising given the literature on novice and preservice teacher responsiveness (Hogan, Rabinowitz, & Craven, 2003). Many studies suggest that even TCs who are able to notice salient student thinking do not respond to that thinking in ways that promote deeper reasoning (Heritage, Kim, Vendlinski, & Herman, 2007). It is, therefore, more surprising that TCs ever responded with High II follow-ups, and it is especially surprising that Craig used such follow-ups regularly.

However inconsistent, these findings expand on the growing body of evidence that some TCs can respond to that reasoning in ways that maintain student reasoning as the focal point of the discourse, when given adequate support (e.g., Coffey, Edwards & Finkelstein, 2010; Kazemi et al., 2009; Levin et al., 2009; Levin & Richards, 2010; Singer-Gabella et al., 2015; Windschitl et al., 2011). Craig, in particular, appeared to respond in ways not yet documented in history education literature (c.f., Monte-Sano, 2011a).

High I and Medium Responses. All four of the TCs frequently exhibited a tendency to demonstrate their own reasoning, either in response to a student idea (High I) or simply as a means to make a point not directly related to a student idea (Medium). Commonly, these High I and Medium responses occurred when TCs seemed to have a point in mind that was associated with a target learning objective, a point that they felt needed to be highlighted.

In stark examples of such responding, Sally met student confusions or questions by simply talking through the target thinking step-by-step. Kendra, on the other hand, seemed to turn to High I and Medium responses when conversations lost a thread or when her intended point was in danger of being obscured. Like Gabby and Craig, Kendra's Medium responses were often vague reformulations of student reasoning in which the TC appeared to expand on a student idea but the idea actually originated with the teacher, such as the time that Gabby extracted the definition of "philanthropic" from a student's idea before moving on.

For Craig and Gabby, High I and Medium responding appeared to gently steer students toward the tasks' objectives. Kendra and Sally, alternatively, were not so gentle in their attempts to compel student thinking toward some desired end. However, both Craig and Gabby used Medium responses less frequently in the later observations. Perhaps their movement away from Medium responses and toward High I responses was indicative of an increasingly clear focus on student thinking objectives.

All of the TCs tended to stop exploring student thinking when they heard the thinking that the lesson targeted. This was most obvious in the case of Sally, who had a very clear notion of what it meant for students to "complete" the reasoning the task required. But, even for Gabby, there was a sense that once students had identified quotes and turned them into evidence with explanation, the thinking was complete. Craig too, only pursued student reasoning while the target idea remained unarticulated. Perhaps a reason Kendra demonstrated so many High II responses in the Lexington Green Lesson was because she was unsure what she was targeting and thus, responded to student thinking with one probe after another. As she clarified her "outcomes" in later lessons, her responses were less exploratory.

The TCs' penchant for High I and Medium responses is not unlike one of Singer-Gabella et al.'s (2015) preservice teachers who observed "instances in which student thinking was invited but not positioned as a lever for learning" (p. 7). The authors attributed this TC's tendency to limitations in her understanding of how student thinking could serve as a resource to advance mathematical understandings. My TCs also appeared more comfortable steering students toward their objectives rather than closely attending to that thinking and determining a means to build on what was available.

Few Low responses. Of the four TCs, only Sally consistently used Low responses. Sally's tasks offered few genuine opportunities for students to voice their thinking about the documents. Sally did not approach tasks that I observed as opportunities to explore student thinking and may have not even recognized that there was thinking to explore, beyond a capacity to match quotes to arguments. For many of Craig's, Gabby's, and Kendra's tasks, there were few easy answers, and as such, were limited opportunities for Low evaluative responses. Although I observed each of the TCs rebroadcast student statements in a way that suggested student thinking was on the right track, such responses were rarely isolated from more responsive follow-ups.

The limited numbers of Low responses among these TCs is perhaps indicative of the interconnected nature of EIR. That is, the lack of generativity in the tasks that Sally launched was associated with her Low responses. For the others, especially Gabby and Craig, the generativity of the tasks they launched precluded simple answers, and in some cases, confounded such answers.

Given the literature's depiction of the typical secondary history course already described above (Cuban, 1991; Downey & Levstik, 1991; Goodlad, 1984; VanSledright, 2002; Wineburg, 1991b; Wineburg & Martin, 2004), the limited frequency of Low response among these TCs was

unanticipated. Again, this finding is likely related to the nature of the tasks that TCs launched in their placements. In most of the lessons I observed in Gabby and Craig's class, and many I observed in Kendra's class, there were simply few opportunities to treat student ideas as "right" or "wrong," beyond reading comprehension and details of historical background.

Conclusions and Implications

In this dissertation study, I set out to investigate how teacher candidates elicit, interpret, and respond to student thinking about historical evidence during their teacher preparation program. The research questions that guided this inquiry were:

1. How do TCs engage in these practices at the outset of their program?

2. In what ways do TCs' capabilities to engage in these practices change during their program of study?

Three overall findings are most worthy of mention because they contribute to gaps in the research literature.

First, some of the TCs in this study elicited, interpreted and responded to student historical thinking (Gabby & Craig) while others did not (Sally) or did so only in certain scenarios (Kendra). This contributes to the literature by extending the findings of previous mathematics and science education studies that suggest some TCs are able to attend to student disciplinary thinking and respond in ways that maintain a focus on student reasoning (e.g., Coffey, Edwards & Finkelstein, 2010; Kazemi et al., 2009; Levin et al., 2009; Levin & Richards, 2010; Singer-Gabella et al., 2015; Windschitl et al., 2011). The present study strengthens these claims by demonstrating that some history TCs are able to engage in such practices when given certain supports.

Like several other studies (e.g., Singer-Gabella et al., 2015), including some in history education (e.g., Monte-Sano & Budano, 2011), the present study confirmed that some TCs cannot (or will not) engage in this practice, even when offered similar supports as the TCs who did EIR. Similar findings are evidenced in other subject area research as well. Together these findings suggest the question is not if TCs can engage in EIR but, which individual characteristics and social contexts of learning are influential in the development of EIR. The findings of this study confirmed that disciplinary understandings play in important role but other characteristics, unaccounted for in this study, surely contribute as well.

Second, TCs in this study regularly used and adapted conceptual and practical tools from the methods course to their internship teaching. This finding is significant because much of the literature suggests that methods courses have a insignificant impact on TC practice in internships (Clift & Brady, 2005). I found, in contrast, that the TCs regularly used practical tools offered in the methods course and frequently, during debriefs discussed conceptual tools directly related to methods course curriculum. Although this study was not designed to validate a core practice approach, the findings provided innumerable examples to suggest that the methods course influenced the development of TC practice in the internship. Furthermore, the findings demonstrated that a practice promoted in the methods course (i.e., EIR) was incorporated by some TCs in internships and retained during their lead teaching.

Third, the findings of this study suggest that attention to students' justification abilities (e.g., general argumentation) can be, but is not necessarily a harbinger of a TC's ability to attend to historical reasoning. This finding is significant because the distinction between justification and historical thinking is not clear across the literature and may play a role in the difficulties that TCs face as they learn to EIR historical thinking. As far as I can tell, the present study is the first

to include analysis of an attempt to develop TC's abilities to distinguish between justification and historical reasoning.

Reconsideration of the Conceptual Framework. The EIR conceptual framework for this study functioned as a lens to direct my attention throughout data collection, analysis, and reporting. The framework allowed me to approach EIR as a whole practice, an interconnected mechanism by which TCs were able to facilitate student reasoning in their classrooms. Additionally, the EIR framework allowed me to unpack this complex practice into component parts, each of which appeared to impact the other.

At the beginning of this study, the EIR framework depicted in Figure 1.1 represented my understanding of a broad base of literature and demonstrated my conceptualization of responsiveness in history instruction. Based on the findings of this study, I have revised the conceptual framework to demonstrate some of the key contributions of this study. Upon completion of my analysis, it was apparent that the original conceptual framework used for data collection and analysis depicted only a broad outline of the complexity of the phenomenon. Figure 9.1 offers a new framework that more definitively illustrates the phenomenon of interest in this study, in light of the findings.
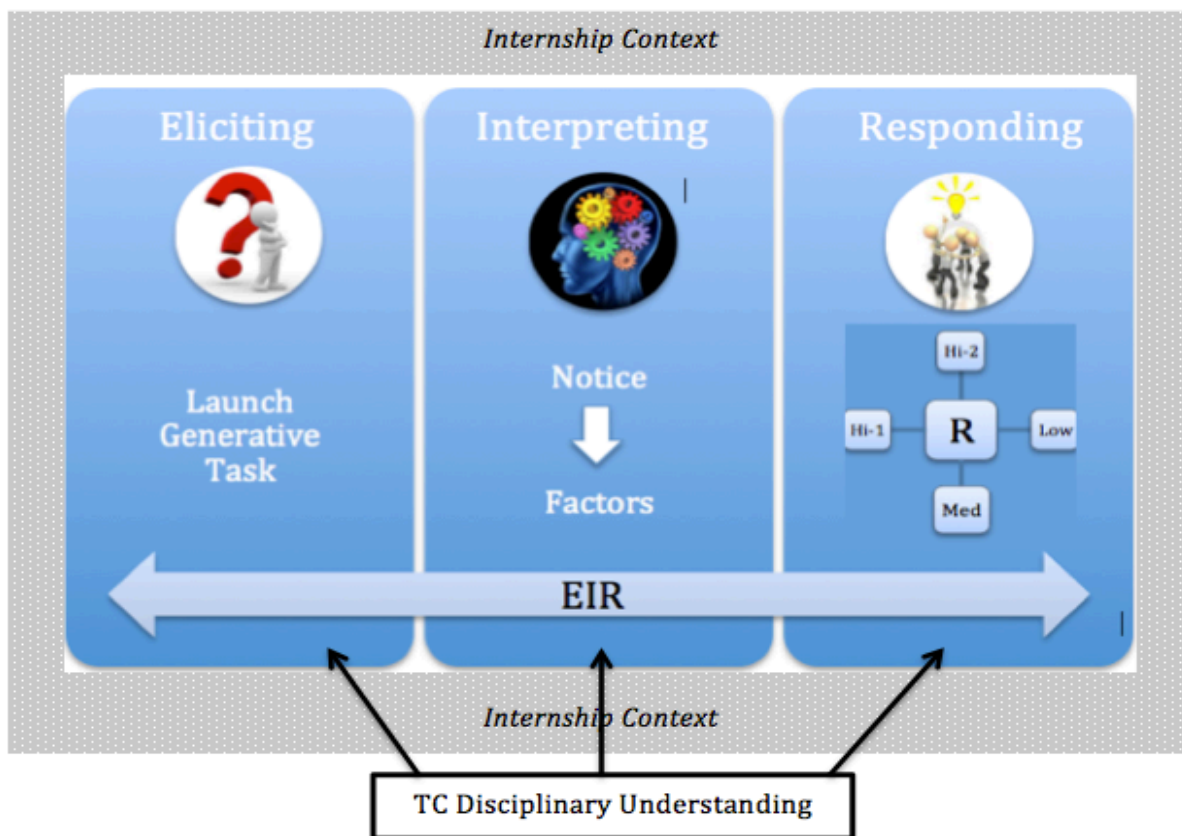
Figure 9.1: Revised Conceptual Framework

Context and disciplinary understanding. In Figure 9.1, teacher candidate EIR is depicted as occurring within a particular internship context and impacted by disciplinary understandings of the TC. While the internship is certainly not the only context that matters for TCs, it serves as a key laboratory for learning because it is the site of the most authentic and complex enactments of teaching practice. As such, it is portrayed in the revised conceptual framework as a canvas on which EIR is envisioned and enacted. Disciplinary understanding, an individual characteristic key to TCs capacity to EIR in this study, is depicted as impacting each aspect of EIR.

Eliciting. The eliciting column in the revised framework highlights the importance of generativity because the findings of this study suggest that the launch of a generative task was necessary for TCs to elicit student thinking about evidence and thus, to interpret and respond to

such thinking. The interpreting column in the revised framework highlights the relationship between what the TC claims to notice and what factors they claim to consider in formulating an instructional response.

Interpreting. While interpreting will always be a 'black box' of sorts, this study demonstrated the importance of exploring both the observed actions of the teacher and the TC's explanation of those actions. Explanation without observation would have left important holes in the narrative about what the TC noticed and considered in light of student thinking. Likewise, my assumptions about what a TC was thinking based only on the responses I observed were often insufficient and sometimes entirely wrong. Thus, the interpreting column of the revised conceptual framework depicts what the TC claimed to notice, in light of what the researcher assumed the TC noticed. Similarly, the graphic depicts the TC's description of the factors considered, in light of the researcher's observations. This expanded portrayal of interpreting establishes complex but rich ground for exploring TC interpretation of student thinking.

Responding. The responding column of the framework includes four analytic categories of response observed and detailed in this study. This revised portion of the framework highlights the reality that it is not only whether a TC responds but how a TC responds to students' thinking that is a key factor in promoting classroom disciplinary discourse.

The revised EIR framework above can advance a conversation in the field of social studies education about what it means to attend to student thinking in history, both for the purposes of teacher education pedagogy and research analysis. As far as teacher education is concerned, the framework helps to clarify the key aspects of practices and moves associated with attention to student thinking. As far as research is concerned, the framework articulates a

complex phenomenon not previously connected in the social studies education research literature and suggests analytic models that can be tested and adapted in future studies.

Implications for teacher education. Below, I speculate on what implications the findings of this study suggest for teacher education. The fact that some of the TCs in this study elicited, interpreted and responded to student historical thinking while others did not has implications for social studies teacher education. As studies in other subject areas suggest (Singer-Gabella et al., 2015; Thompson et al., 2012), TCs will require various types of support in order to engage in this practice. The findings of this study suggest potential directions for methods courses and internships, and ways in which teacher educators can better support the development of TCs' abilities to EIR.

Methods courses. Given that a generative design was necessary for the launch of a generative task in this study, methods courses can play a key role in developing TCs' abilities to identify aspects of generativity, adapt tasks in generative ways, and design generative tasks. However, the present findings make clear that instructors should never equate an ability to envision a generative lesson with an ability to launch such a lesson. As others have pointed out, methods courses must do more than simply support TCs in task design and analysis (e.g., Clift & Brady, 2005; Grossman, 2005; McConney, Schalock, & Schalock, 1998). TCs need opportunities to enact and then investigate key aspects that make the difference in the launch of generative tasks, such as clearly framing historical questions, providing sufficient background information, and connecting to students' prior knowledge. Without opportunities to practice these competencies in low-stakes settings of reduced complexity, such as rehearsals and modified classroom enactments, teacher educators should not expect to recognize many of the

gaps in TCs' understandings of EIR that will become painfully obvious when they face an hour in front of students.

Methods courses also have an important role to play in the development of TCs' abilities to analyze student thinking in controlled environments and fashion appropriate responses to that thinking. As with eliciting, however, a TC who can notice student thinking and formulate responses in decontextualized settings will not necessarily be able to do so in the classroom. Thus, methods courses should provide structured opportunities for TCs to work across settings of teacher education in order to practice interpreting and responding, first in low-risk settings of manageable complexity. These enactments, then, can become the subject of new analysis in the methods class and the foundations for new designs.

In light of the above implications, the findings of this study also suggest limitations, and perhaps failures, of the core practice approach that Meredith and I implemented in the methods course. For example, TCs needed more opportunities for structured analysis of their own attempts at EIR, in order to address problems that were evident early on but not diagnosed until later. Correspondingly, we needed a stronger mechanism for feedback on TC enactments than we originally envisioned. The limited amount of targeted feedback seemed to contribute to a perpetuation of problems observed early in TCs' internship performance.

Internships. The findings of this study suggest that internship placements may be a key factor for the development of EIR abilities, especially for TCs with weaker disciplinary understandings. Without support and resources offered at the placement site, it is unlikely that a TC with weak disciplinary understandings will develop a capacity to EIR. In particular, rich and meaningful internships will require strong mentors who have a strong grasp of the discipline and a vision for teaching in ways that leverage disciplinary literacy.

Thus, placement decisions for such TCs should prioritize mentors who regularly use historical thinking and have resources available to support TC needs. Conversely, the TCs in this study with strong disciplinary understandings seemed only to need an environment welcoming of new approaches, and not necessarily a mentor who could provide consistent support in teaching historical thinking.

Ambitious core practice designs need to attend closely to the contexts of the internships and the feedback and support that the TCs receive in these settings. Core practice approaches that are limited to sending TCs on "enactment errands" in their internships without designs that consider contextual elements are not likely to return promising outcomes. One approach that could support TCs continued development of core practices in the internship is based on a cognitive apprenticeship model and is described in the professional development work of Monte-Sano, De La Paz, and Felton (2014).

Teacher educators. As alluded to in Chapter 4, I was not sufficiently prepared to teach these TCs how to EIR historical reasoning. In retrospect, I recognize that my understanding of historical reasoning was limited in certain ways that required learning conceptual and practical tools either just before or while I taught. Furthermore, I realize that my understanding of how to develop expert thinking, both for students and TCs, was limited. I made many adaptations and revisions to improve the methods course, not only in response to student needs but also in response to my own development in understanding EIR and historical thinking.

Despite unexpected limitations in my understandings, my chief obstacle was a lack of experience engaging in EIR with secondary history students. Although I taught social studies for many years, I did not come to value student discourse in the ways I do now until graduate school. Even with Meredith's support, I feel my lack of practical experience engaging secondary

students' historical reasoning played a significant role in the gaps I later recognized in the TCs' performance in this study. Available literature on the practices of social studies teacher educators is mostly limited to their supervisory practices (e.g., Slick, 1998; Wilson & Saleh, 2000) and does not consider the demands of teaching EIR or historical reasoning.

To develop TCs' ability to EIR historical thinking, social studies teacher educators must have a strong understanding of "doing history" and a strong understanding of the varying levels at which TCs will encounter these ideas in actual classrooms. Social studies teacher educators should have significant experience engaging in EIR and in teaching historical thinking to secondary students. Ideally, teacher educators will play a role that allows modeling and support across the settings of teacher education, acting sometimes as classroom teacher, sometimes as classroom coach, and sometimes as methods course instructor. Given these expectations, the training and resources available to teacher educators should be a matter of concern. Despite some limited attention (e.g., Goodwin & Kosnik, 2013; Murray & Male, 2005; Zeichner, 2005a), the professional development of teacher educators, whether social studies or otherwise, has not garnered much attention to date. The preparation and continuing professional development of teacher educators must become a priority for curricular or structural reforms to have sustaining effects on teacher education.

The finding that TCs in this study regularly used and adapted conceptual and practical tools from the methods course in their internship teaching has implications for the field of teacher education because it demonstrates a link between methods course instruction and TC practice. Some social studies education researches have argued that methods course instruction impacts TC practice (e.g., Dinkelman, 2000; Fehn & Koeppen, 1998; Slekar, 1998) but links were often difficult to establish. TC use of methods course IAs in this study demonstrates a

direct link between the methods course instruction and TC practice in the internships. Practices that can "travel between coursework and internships" (Lampert et al., 2013, p. 228) offer an opportunity for teacher educators to assess the TCs' ability and willingness to engage the conceptual and practical tools offered in the methods course. The findings of this study suggest that certain practice-based approaches to teacher education, like the core practice approach described here, are worthy of further development and examination.

This study's finding that TC attention to students justification abilities can be, but is not necessarily a precursor of ability to attend to historical reasoning suggests implications for teaching TCs to EIR. If TCs are to learn how to promote historical discourse in their classrooms, curricula, including courses on reading and writing in the content area, need to emphasize the distinction between general argumentation and historical arguments grounded in the three heuristics highlighted in this study.

Correspondingly, the findings of this study emphasize the importance of disciplinary knowledge in teaching social studies, a subject area construct not united by a common discipline. Disciplinary heuristics and disciplinary literacy are fundamentally tied to notions of expertise and are clearly expected in the Common Core Framework and the new NCSS C3 Framework. However, the findings of this study suggest that teachers will not be able to develop disciplinary expertise in students if they are not already proficient themselves. At present, future economics teachers sit next to future history and geography teachers (and government, sociology, and psychology teachers) in university methods courses and complete an almost identical course of study.

Perhaps it is unrealistic to expect TCs to develop expertise across all of the subject areas of social studies. On the other hand, attempts to focus only on history at the expense of the other

social studies disciplines will likely meet the type of pushback described in Chapter 4, when TCs in government placements were frustrated by the historical emphasis of the methods course. Likewise, it is impracticable for teacher education to somehow facilitate all of these capacities in so brief a span of time and with the resource limitations most programs face. Long-term professional development solutions that draw on the resources of school districts, university social science and education faculty, and community resources may offer the best direction for the future of social studies teacher development.

Future Research. Because this study, like others in science and mathematics, suggest some TCs can EIR when offered support and resources, the field should move beyond stage-based notions of teacher development. A more helpful direction would be an exploration of the characteristics of TCs and social contexts of learning that are most suited to learning to EIR. Research that targets specifics of TCs disciplinary understanding and explores the implications for teachers' abilities to focus on and respond to student thinking in productive ways could advance the field's understanding of ambitious novice teaching in the social studies. Conversely, it will be important to explore why some history TCs are not able (or willing) to EIR. Given the apparent importance of context for the TCs in this study who struggled to EIR, studies that explore additional tools, coaching, and remediation can make important contributions to the gaps that remain.

Key questions remain unanswered regarding why TCs used some methods course tools and not others and how the TCs' contexts contributed to these decisions. While similar questions have been addressed in other subject areas (e.g., Thompson et al., 2012) and to some extent in history education (e.g., Fehn & Koeppen, 1998; Monte-Sano, 2011a), the inclusion of a particular methods course approach (such as the one described in Chapter 4) may bring new

relevance to such inquiries. One approach that could build on the findings of this dissertation would be a similar study that includes lesson plan "think-alouds" in addition to the debriefs as a means to investigate the vision that the TC had for student thinking in the designed task, and the origins of that vision.

Opportunities abound for researchers interested in exploring core practice approaches in social studies and teacher education broadly. The notion of "core practice," however, is not yet established enough to have a clear meaning, even within the bounds of a single subject area. As such, teacher educators, teacher education researchers, and teacher-researchers will benefit from studies that include both the implementation details of a core practice approach (e.g., Chapter 4 of this study) and targeted assessment of TC performance, in light of the approach. Descriptions of core practice pedagogies without investigation into the ways TCs understand and incorporate these practices into their teaching are not particularly helpful for the advancement of the field of teacher education, except as a precursor of inquiry into TC performance. If a core practice approach is indeed promising, research must demonstrate that this promise extends into the classrooms of TCs and program graduates. Teacher educators and teacher education researchers must work together to seek not only improvement of core practice approaches, but also improved methodology for researching what TCs understand and how they enact core practices introduced in the methods course.

This study's finding about the ambiguous relationship between historical reasoning and justification offers an additional, and related, line of research. Future research should explore the impact of teacher education and professional development that explicitly addresses this distinction. Studies can examine whether the distinction may be a lever for helping TCs begin to design generative tasks and consider nuances in students' disciplinary thinking. Likewise, future

research can explore whether the clarification of general argumentation and historical thinking can help practicing teachers recognize performance gaps in their students' abilities and in their own understandings.

Concluding remarks

This dissertation study fills a gap in the research on core practice approaches to teacher education. Specifically, the study demonstrated the ways that TCs elicited, interpreted, and responded, in light of a methods course designed to support the development of that core practice and found that some TCs enacted EIR in lead teaching. Furthermore, this study highlighted key aspects of EIR that are specific to history education and fill gaps in the history education literature related to attention to student thinking and historical argumentation. The findings provide a foundation for future research in a number of areas and point to the need for additional inquiry into teacher development in disciplinary understanding and attention to student thinking.

Appendices

Appendix A: Placement Charts

SCHOOL	Course	Main Placements	Intro Placements
XXMS	7th grade World	Jennifer	Maddie
YYMS	7th grade World (2 reg, 2 co-taught)	Maddie	Same
ZZMS	8th grade US	Francis	
XXHS	9th grade US	Gabby	
YYHS	9th grade US History	Sally	Same
ZZHS	2 Honors US; 1 Int'l HRts	Catherine	Same
ABHS	2 ESOL Gov't	Tom	Same
ABHS	2 Gov't, 2 World	Melanie	Same
ABHS	2 Gov't	Kendra	Same
ZZHS	4 Inclusion Gov't	Kary	Same
ABHS	Government	Melanie	Same
XXHS	AP World History	NO CANDIDATE	Jennifer
XXHS	AP World History	NO CANDIDATE	Francis
ZZHS	4 World (2 ESOL bridge)	Craig	Same
ZZHS	3 World History	Elliot	Same

Appendix B: Pre-test Disciplinary Understanding

Pretests were scored according to Nokes et al.'s (2007) rubric for disciplinary understanding. One limitation of this analysis approach was an emphasis on heuristic count rather than content of historical heuristics. To corroborate the scores, I assessed the pretests against an entirely different rubric (Seixas & Morton, 2014) and found similar patterns. In addition, insights drawn from candidate's scores were corroborated against questionnaires and coursework.

Candidate	Total Score	Heuristics Total (SCC)	Justification (Use of Doc)	Sourcing	Corroboration	Contextualization
Gabby	**35**	**15**	**20**	**6**	**6**	**3**
Craig	**24**	**13**	**11**	**7**	**2**	**4**
Kendra	**18**	**6**	**12**	**3**	**2**	**1**
Maddie	15	8	7	5	0	3
Sally	**15**	**5**	**10**	**2**	**1**	**2**
Elliot	15	10	5	5	0	5
Francis	13	6	7	3	1	2
Catherine	10	6	4	5	0	1
Jennifer	8	4	4	3	0	1

In evaluating these scores, I made a distinction between the heuristic score and justification score (the use of documents as evidence score). A candidate could cite justification in ways that might mirror the requirements of a DBQ (supporting argument with evidence) but be unable to frame evidence that involved historical reasoning. A candidate with extensive disciplinary understanding would be able to do both. In this way, Craig and Gabby emerged as clear leaders in disciplinary understanding and most of the other candidates' scores hovered in a similar range. Thus, candidates were sorted as relative to one another into two categories: extensive disciplinary knowledge and average disciplinary knowledge.
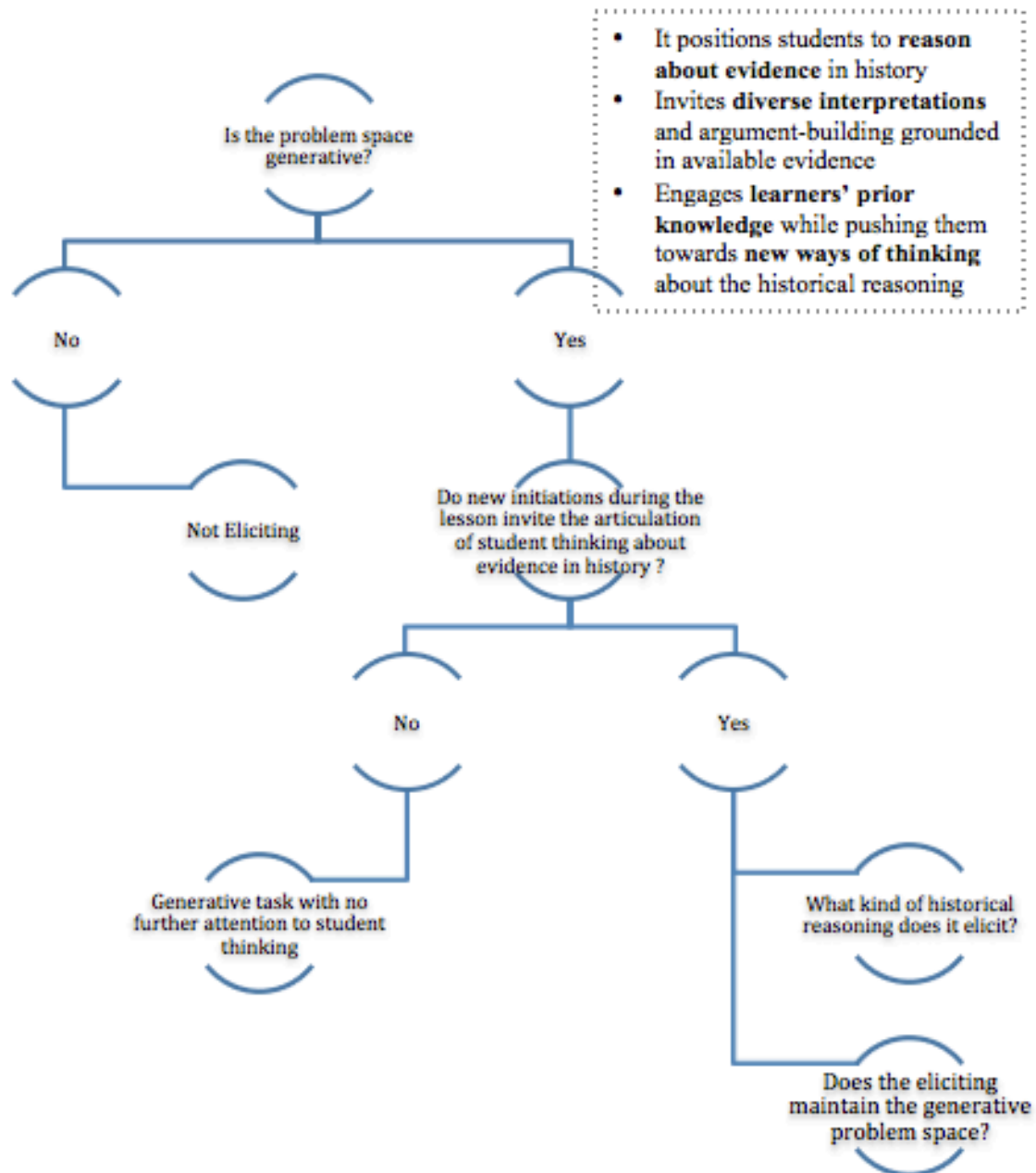
Appendix C: Observation Debrief Protocol

1. Did things go the way you expected them to go? (Explain)
2. How were things different than you expected them to be?
3. Tell me about the student thinking in the ideal version of this lesson (the one in your mind when you planned it).
4. How was the student thinking that you actually observed in the classroom similar or different from the ideal that you envisioned?
5. What did you notice about how the students used the documents/artifacts you gave them during the lesson?
6. I noticed that your inquiry question was _____. Can you talk me through your reasoning for choosing that question?

I noticed that you asked Student 1 to _____. What made you want to ask that question?

I noticed that you suggested Student 2 should (specific task). What made you want to the student to do that task?

Appendix D: Eliciting Analysis Protocol

Is the problem space generative?

- It positions students to **reason about evidence** in history
- Invites **diverse interpretations** and argument-building grounded in available evidence
- Engages **learners' prior knowledge** while pushing them towards **new ways of thinking** about the historical reasoning

No

Yes

Not Eliciting

Do new initiations during the lesson invite the articulation of student thinking about evidence in history ?

No

Yes

Generative task with no further attention to student thinking

What kind of historical reasoning does it elicit?

Does the eliciting maintain the generative problem space?

Eliciting Codes (What kind of historical thinking?)

A. Eliciting Sourcing
Sourcing: An individual who uses sourcing looks at the source of a document before reading and keeps the source of the document in mind as he or she reads. The reader's understanding of the document is influenced by the document's source. Sourcing only occurs when the consideration of the source helps the individual make sense of the document.

It is not sourcing if an individual shows an awareness of the type of text, but does not tell why the type of text is important (unless it is obvious). It is also NOT sourcing if they comment on the text in general (lots of details/too confusing/syntax).

A1. AU Position Eliciting
If the teacher elicits thinking concerning the occupation, profession, level of training, or other credentials of the author of the document, it qualifies as eliciting sourcing.
Examples:
- "What is the author's role in the British army?"
- "Since Shaw was an officer in the British army, what might he have known?"
- "Consider the reliability of a document written by a person in his professional position."

A2. AU Motivation Eliciting
If the teacher elicits thinking regarding why an author might have written the document it qualifies as eliciting sourcing.
Examples:
- "Is there anything that Colonel Jackson might have had to gain by telling his commander about his success?"
- "Why would the author be trying to convince the readers that the Americans started the battle?"

A3. AU Participation Eliciting
If the teacher elicits thinking regarding the author's level of participation in an event, it qualifies as eliciting sourcing:
Examples:
- "Notice that Jones was a witness of the battle, does that make the document more or less reliable?"
- "How did Smith get his information about the battle?"

A4. AU Evaluation Eliciting
Teacher elicits thinking about any other consideration of the author
Examples:
- "Why do you say that the letter seems more truthful?"
- "Does the fact that the author admits that he can't remember make any difference in terms of reliability of this document?"

A5. Date Production Eliciting
If the teacher elicits thinking regarding when a document was created, it qualifies as eliciting sourcing.

Examples:
- "Does the date of the document tell you anything about the reliability of this information?"
- "This was written in his journal the day of the event. Does that make you think any differently about how you view this account?"

A6. Document Type Eliciting
If the teacher elicits thinking regarding the type of document it qualifies as eliciting sourcing.
Examples:
- "Notice that this statement was sworn before a justice of the peace. Does that suggest anything about the reliability?"
- "People usually write in their journals to keep a record for themselves. What does that suggest about reliability?"

A7. Evaluation of Document Eliciting
If the teacher elicits thinking about any other reason why the document is more or less reliable could be considered sourcing.
Example: "If textbooks tend to exaggerate the good about a country and leave out the bad, how might we rate this document's reliability?"

A8. Other Sourcing Eliciting
Only examples of eliciting sourcing that I am unsure how to categorize.

XA. Sourcing Related
Sourcing related information but not an example of eliciting sourcing

B. Eliciting Corroboration
An individual uses corroboration when he or she compares or contrasts information found in two or more specified documents. It is only corroboration when it helps the individual make sense of the event.

It is NOT corroboration if the teacher elicits a comparison or contrast with student background knowledge rather than information from another text. For example, "how does this compare with what you learned in eighth grade?"

B1. Direct Comparison Eliciting
If the teacher elicits thinking regarding a direct connection between similar information that was found in two or more documents, it qualifies as eliciting corroboration.
Examples:
- "Focus first on what all the authors seem to agree on."
- "Do Simpson and Smith have the same story about what happened after the soldiers arrived on the scene?"

B2. Direct Contrast Eliciting

If the teacher elicits thinking regarding information that was different between two documents, it qualifies as eliciting corroboration.
Examples:
- "What differences, if any, do you notice between these two accounts of where the shot was fired?"
- "Is Jones' account of what happened different from the others' accounts?"

B3. Claim Uniqueness Eliciting
If the teacher elicits thinking regarding information that was found in only one source, it qualifies as eliciting corroboration.
Examples:
- "Pay attention to what Valdez wrote that none of the others mentioned."
- "Does the textbook have information that we don't see in any of the other sources?"

B4. Claim Omission Eliciting
If the teacher elicits thinking regarding important details that a source left out but was found in other sources, it qualifies as eliciting corroboration.
Examples:
- "Smith was an eye-witness. But notice that he does not mention anything about hearing the command to fire."
- "Since Harper didn't include any information about the bad effects of the program, what might we conclude?"

B5. Corroborating Other Eliciting
Only examples of eliciting corroboration that I am unsure how to categorize.

XB. Corroboration Related
Corroboration related information but not an example of eliciting corroboration

C. Eliciting Contextualization
An individual uses contextualization when he or she attempts to place himself or herself in the specific context of the event that is taking place. He or she discusses specific details about the event that helps him or her understand why or how the event took place.

It is NOT contextualization when an individual inappropriately portray today's values or culture on the people of the past. For example, if a student argues that lots of women have short hair, so it shouldn't have been shocking for a woman to get her hair cut in 1920, this is a presentist analogy.

C1. Time or Location Awareness Eliciting
If the teacher elicits thinking regarding the chronology of an event, or specific features of the physical location of an event, it qualifies as eliciting contextualization.
Examples:
- "Can we make any inferences based on the date or location of the Boston Tea Party?"

- "Gettysburg is a hilly area with some forests around it. How does that connect with what we already know?"

C2. Culture or setting awareness Eliciting

If the teacher elicits thinking regarding the cultural values or common attitudes of the time period, or emotions that participants in an event may have been feeling, it qualifies as eliciting contextualization.
Examples:
- "Does the fact that the soldiers had been marching all night make any difference here?"
- "If I am a typical American in the 20s, what would I think a woman ought to be doing with her time?"

C3. Biographic Awareness Eliciting

If the teacher elicits thinking regarding the influence of prominent individuals who were involved in an event, it qualifies as eliciting contextualization.
Examples:
- "Everyone just seems to accept this… Whose idea was it in the first place?"
- "Let's imagine, Washington, being who he was at the time, walking into that room and proposing …"

C4. Historiographical Awareness Eliciting

If the teacher elicits thinking regarding the methods used by the author to understand the event that they write about, it qualifies as eliciting contextualization. In other words, if the teacher invites students to question whether a historian used effective methods to study an event before writing, it is eliciting contextualization.
Example: "Does it matter that the historian who wrote this didn't have access to most of the information we have looked at about the …?"

C5. Linguistic Awareness Eliciting

If the teacher elicits thinking regarding the different meanings of words over time, it qualifies as eliciting contextualization.
Examples:
- "How could we know whether Lincoln's use of the use of the word 'Negroes' was intended to be derogatory?"
- "Let's take a look at the language around the word 'misdemeanor' in order to see if the writers of the constitution meant what we mean when we use that word."

C6. Analogy Eliciting

If the teacher elicits a connection with information in the past by comparing it to current events or personal experiences, it qualifies as eliciting contextualization.
Example: "Could the debate over Prohibition be compared to a debate in our day? ... In what ways?"

C7. Contextualization Other Eliciting

Only examples of eliciting contextualization that I am unsure how to categorize.

XC. Contextualization RELATED

Contextualization related information but not an example of eliciting contextualization

D. Justification (Eliciting the use of documents as evidence)

An individual uses documents as evidence when she supports a statement that she makes with facts or arguments that come directly from one or more of the source texts. A relatively specific reference should be made in order to count as using documents as evidence.

It does NOT count as using documents as evidence when an individual makes a vague or general reference to the documents as a whole. For example, "After reading the documents, I believe that …"

D1. Direct Quote Eliciting

If the teacher elicits thinking regarding a direct quote from a document (either quoting the document or inviting a student quote) it qualifies as eliciting justification.
 Examples:
- "Wait a second, didn't Smith write, 'I warned the protesters three times before commanding my troops to fire?"
- "Patterson claimed that, 'Jefferson purchased two new slaves today.' What does that tell us?"

D2. General Citation Eliciting

If the teacher elicits a paraphrase of information found in a document (either paraphrasing the document or inviting a student to paraphrase) it qualifies as eliciting the justification.
Examples:
- "Can someone sum up what evidence we have that the British fired first?"
- "I recognize that you think Smith supports your argument but what is it that his statement shows?"

D3. Specific Reference Eliciting

If the teacher elicits thinking regarding a claim that a specific document supports an idea, it qualifies as eliciting the justification.
Examples:
- "Do we have a document that seems to agree with depiction in the picture?"
- "That would be a good argument if our documents verified it. Do they?"
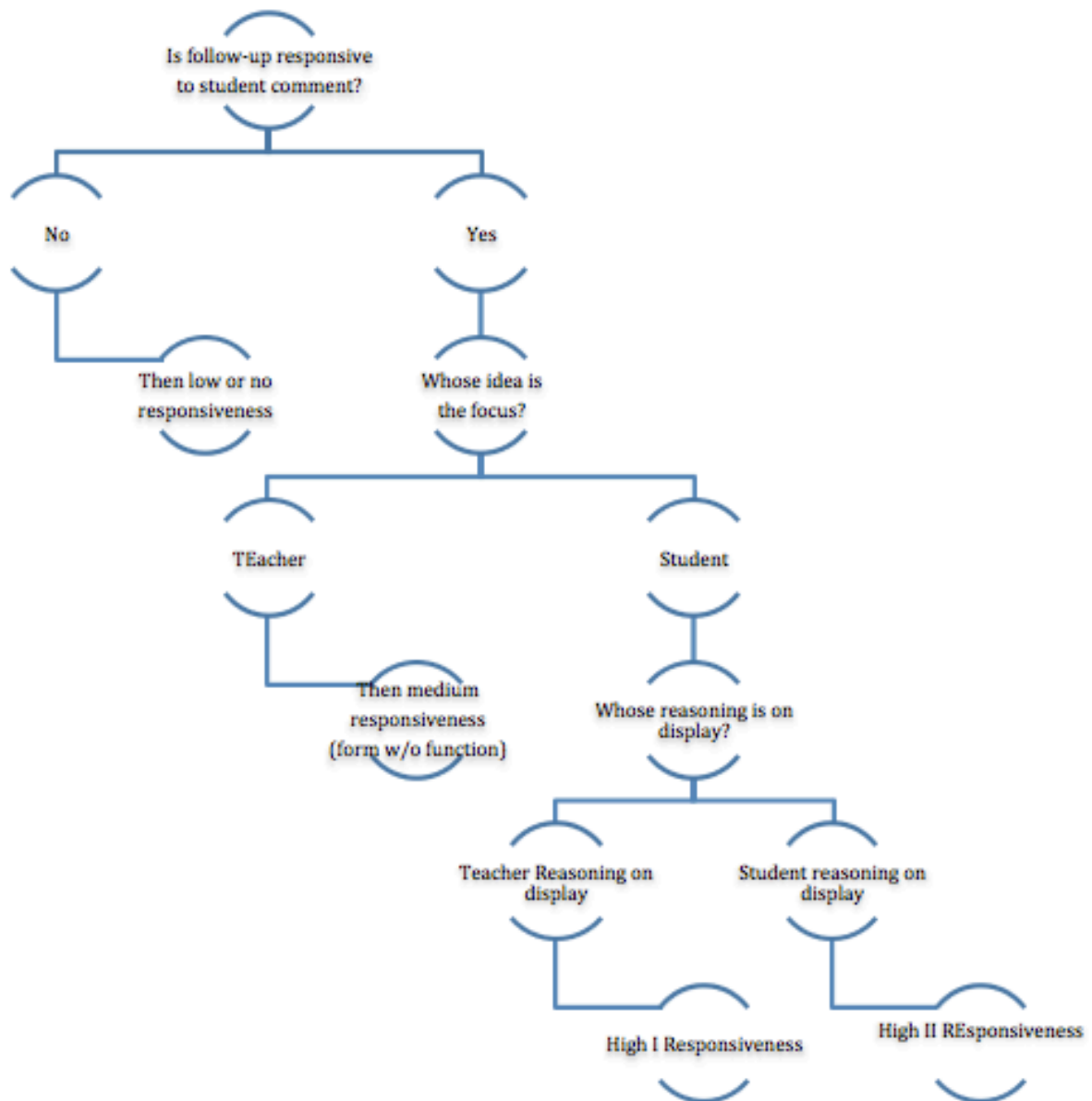
D4. Use of Doc Other Eliciting

Any other examples of a teacher eliciting references to specific texts to bolster an argument that I am unsure of how to categorize.

XD. Use of Doc as Evidence RELATED

Use of documents as evidence-related but not eliciting justification

Appendix E: Responding Analysis Protocol

303

Low Responsiveness: Follow-up that is not responsive to student's idea. Moves include evaluating, rebroadcasting, acknowledging, or making a related statement or question		
Evaluation	T tells S whether response is correct or incorrect. Can be explicit or implied.	S: The statement was written by the minutemen. T: Not the minutemen. S: The document is not reliable. T: Very good.
Rebroadcast	An echo which functions as an implicit evaluation of correctness.	T: Ok, so what about robber barons. S8, what do these pictures say about robber barons? How do they help explain it? S: They consider themselves bigger? T: Ok, bigger than everyone else.
Acknowledge	T acknowledges S response (Oh; ok; thank you).	T: You want to add something? S7: I want to add to what S3 was saying. He said, "to our knowledge no one was firing." How could the colonist fire when the colonists' backs were turned? T: So, you are looking at the story to see if the story makes sense. Very good. So, now what we are going to do in our groups…
Related Statement/Question	T asks question or makes statement related to S comment but does not build or incorporate on it. Continuation of T's line of thinking where **conversation might proceed in same manner regardless of S response.**	T: Ok, side B. S: This is written by Ida Tarbell who wrote the history of standard oil. T: So, write what questions you have or what you would need more information since your speed racers.

Medium Responsiveness: Follow-up that is minimally responsive to student; has the form but does not function as responsive. **Focus on T's thinking often in a recitation style interaction** (S provides basic information T incorporates in her response). Also includes corrective moves and co-opting S response (or peripheral part of it) to make desired point.		
Give Correct Answer or Hint	T's response to incorrect answers; provides brief answer or info (hint) with no substantive explanation (not tailored to specific misconception). "Here **let me tell you the right answer"** is what is communicated. **Does not address S thinking.**	S: So, because this document was written by the British commander… T: No, the other way around. T: So, do we know how many people there were? S: The British had more. T: Do we know? I don't think we know. Do we? S: No, it does not say.
Vague Reformulation	T reformulates S comment but connection is vague; initial idea is transformed or T adapts one part of S response (loosely derived) to impart the T's desired knowledge. T might expand on S response but **idea originated from T** because follow-up is embedded in recitation-script. Can also include when T explains S answer.	S: I don't understand. T: So, what would be your evidence that the new deal was not a success (teacher idea)? S: Because she has no clothes? T: Look at the date on the letter. Now, when did the new deal begin? S: 1933. T: So, the fact that she is writing this letter three years after the new deal started suggests things are better or still bad?
Minimal	Includes moves where S provides scripted info in response to known-answer question. T incorporates yes/no answer, like a fill-in-the-blank response. Also includes if T does not engage with S question (brushes it off) or prompts S to provide additional info to complete an answer. Notice that just using the language of a probing question does not necessarily mean it is a high response.	S: The WPA is providing hot well-balanced meals for the children. T: Ok, and how is that a good thing? S: I don't know. T: What document did you use? The WPA right? Ok, so the WPA provided hot meals for schoolchildren. So remember at the beginning of class we talked about breadlines and people not having a lot of food? So if the WPA is providing nutritious food for people, how does that help? S: Because they don't have to go hungry. T: Yes, very good, thank you. Brush off: S: But how come I can't just say that it is not reliable? T: Because you have to consider all arguments. Ok now everybody turn the page. (T does not engage with a substantive S question). S15: In Document D people are adding

		electricity to rural America.
		T: Ok, how is that helpful that they adding electricity to rural areas and building highways? How is that helpful?

High I Responsiveness: Follow-up that is responsive to S idea, question, or perceived misconception. The teacher's thinking is on display, but in response to the student's idea. This includes answering student questions and responding to student misunderstandings. T can expand on S comment but takes over S's idea and puts his/her thinking as focus.

Question Response	**T responds to S question** – can be conceptual question or requests to repeat a response, give basic info, perform historical reasoning, or a question clarifying directions/ instructions.	S: Wouldn't document A and document B serve as the same evidence since they were created by the same author? T: You might think so. But, notice the dates and consider how the story might have changed over time. Thirty years is a long time for a story to evolve.
Correct a misconception	T response more corrective in nature. Pursues S thinking in an attempt to correct perceived misconception by revealing T's reasoning process (T guides reasoning by asking specific sequence of questions – they can be contradictory or counterexamples). Specific feedback targeted to particular misunderstanding. May need to look at whole sequence to determine whose reasoning is being displayed.	T: A bad thing. SO, what does that tell us about whether Shays and his men were justified in doing what they did. S1: If you look at one point (Reads quote…It is impossible to bring them to court…revenge, hatred etc). I think what he is trying to say is that their actions are depending on how they want to end this. Like, if they want to go to court they have to have a lot of evidence for what is going on. T: You are right, if the court system is working correctly, but what he is saying is that it is not possible to bring them to court. Even if they had evidence. So, is that an argument for or against the court systems being just?
Teacher Reasons about student's idea	T expands on S idea. What differentiates this from Medium responsiveness is that the **idea originated with the student** but the teacher's reasoning is on display T takes S idea and she does the analysis, arguing, or rebutting.	S: I would guess that the minutemen were scared because it was at night and they did not know what was happening. T: Notice the context that S2 is building. The minutemen are scared, it is dark, they don't know what is happening. And perhaps they are thinking about firing their weapons for safety's sake (notice that T makes the inference to answer the inquiry question).

Stabilizing the Context with T inference	The teacher pauses reasoning to review relevant content knowledge and then follows up the stabilization by turning the contextual detail into evidence by making the inference for the S.	T: very good. Own and operate their own businesses. So, he talks about in the, basically in the 2nd paragraph he talks about AA making self-sufficient communities. What does that mean? Self-sufficient communities? S5: You can support yourself T: Right, so everything is in their own communities. Remember when we talked about Levittown. They had their own stores and everything and they did not need to go outside their communities. So, Malcolm X thinks African Americans should form these types of communities.
Response Modeling	In response to a student statement or question, the teacher thinks out loud and exposes how he/she works through a complicated idea in a text (e.g., how to reconcile a contradictory statement, decipher difficult prose, or interpret evocative language).	T: Ok, by helping the state, what does she mean? Does she mean the physical land border? Raise hand if you think it is something else? If they are not talking about physical borders, what else could it be? S3: London? T: good guess but no S3: the areas that the rebellion does not go to yet? T: Listen to the sentence if I add a word (adds word state) How does that change your interpretation if I add that word?

High II Responsiveness: Follow-up explores student thinking and allows their reasoning to be the focal point. Uptake in the true sense of the word – responding to and building on a student's idea so that his/her thinking is on display. Includes invitations for students to make sense of one another's ideas; probing S thinking; expanding, clarifying, or giving an example based on S idea; or T asks clarifying question to establish a joint focus of attention.		
Substantive Probe	T invites S to further explain his/her thinking with probing questions – often focused on analyzing, arguing, or rebutting.	S: No, not in great Britain. T: Oh, that is interesting. Listen (to class). Give me more. S: (inaudible) T: Why you say that?
Contradiction/Counter-claim or prompting counter claim	Challenges S thinking by asking Q or making argument/rebuttal that contradicts S conjecture (focus still S thinking and response doesn't illustrate T's reasoning).	T: Living conditions. Not just where they worked but where they lived. what were they? S: Windows broken stuffed with rags, floors wet, (Citing from document) T: Good. So, picture that. Which document was that again? S: document D T: Imagine that – Look at Document D for a second. Close your eyes and think

		about that. Look at the last sentence – kids rolling around "in the filthy moisture of the street oozing up." That is not positive. I thought the whole point about factories was that it made our life better..
		T: Does anyone have a different interpretation of that passage? Who disagrees?
Asking for textual support	The teacher requires that students back their claims with evidence from the documents (from original student or others).	S: (inaudible) T: Yeah, so what is your new answer to this question. That is what I am trying to ask. S1: (inaudible) T: Can anyone who agrees with S1 show me where S1 got that answer from documents?
Make Sense of Other's ideas	Invitation for S's to make sense of one another's thinking (agree/disagree, give hints, etc.).	What do we think about what S8 said? Can both be true?
Uptake/Revoicing	S ideas taken up through revoicing, expanding, clarifying, giving an ex or illustration. Highlights S thinking and connection is clear. Also includes requests for S to repeat idea (when done to emphasize).	S1: they used inventions to hide what they don't want to see T: who is hiding what who does not want to see. Be specific for me. S1: Like it is good for rich people (inaudible) T: So, maybe you agree with what S15 was saying earlier about it being good for the wealthy. Then, you are saying it is hiding what the non-wealthy life is like.
		T: That sounds a little bit like what S1 was saying earlier…do you see any connection?
Stabilizing the Context w/ invitation for student inference	The teacher pauses reasoning to review relevant content knowledge and then allows students to turn that context into evidence. Or, after stabilizing the context, returns directly to the student reasoning.	S5: I just was going to say that kids are not allowed to work. T: Kids are not allowed to work? When are kids not allowed to? S5: Now. T: Like are you saying today? S5: No like ...(inaudible) T: Yeah, so This is a good moment. Lets take a step out here. So, at this time, During the industrial revolution, were kids allowed to work in factors? SS: no, yes M: Yes they were. And, they did. They

		did work and the pay went to the family.
		S8: (something about pay - inaudible)
		T: Good. yes they worked for little pay and they worked just as long of hours. You are right S5 this eventually changes. After they start seeing these problems they start making reforms but at this time, No. Children are working in these factories
		M: why is that negative?

Adapted from Pierson 2009

Incorporates Reisman's moves: http://teachinghistory.org/teaching-materials/teaching-guides/25620

Appendix F: Observation Lesson Topics

Sally Lessons

Observation	Lesson Topic
0[9]	Progressive Reformers (did not use primary sources)
1	The New Deal
2	The Holocaust
3	Brown v. Board of Education
4	The Philosophies of MLK and Malcolm X

Gabby Lessons

Observation	Lesson Topic
0	The Rise of Ancient Rome (did not use primary sources)
1	Robber Barons and Captains of Industry
2	Imperialism
3	The Great Gatsby
4	The 1920's Prosperity SAC

Kendra Lessons

Observation	Lesson Topic
1	The Battle at Lexington Green
2	Shay's Rebellion
3	Federalist 10
4	Factory Conditions

Craig Lessons

Observation	Lesson Topic
1	The Enlightenment Thinkers
2	The Industrial Revolution
3	The Lunchroom fight - Fault for Starting World War I
4	Totalitarianism

---

[9] Observations marked "0" were marked thus because the TC and I had a miscommunication and they did not use primary sources in the lesson. I did not include these in the analysis because it was not congruent with observations of the other TCs.

Dear Teacher Candidates*:

This questionnaire is intended to capture a snapshot of your understandings of various aspects of teaching social studies. While the substance of your answers will in no way impact your grade in this course, we are analyzing these carefully to inform the direction of our time together over the coming year. Thus, we ask that you give this task your full attention and effort. Please ensure that your answers be entirely yours. In addition to serving as a tool to help track your learning over time, the data from this questionnaire will help us make improvements to the curriculum in the coming months and years.

Thanks! Michael and Catherine

1. In what ways do you expect your students to be prepared for thinking, reading and/or writing about history (or the subject you are teaching) at the beginning of the school year?

2. In what ways do you expect your students to be unprepared for thinking, reading and/or writing about history (or the subject you are teaching) at the beginning of the school year?

3. When you envision an ideal social studies classroom, what do you see? What are students doing? What is the teacher doing?

4. Students were given the newspaper stories below (newspaper #1 & newspaper #2) and asked to think aloud as they read. Read the newspaper stories below and the comment Matt made as he read the second newspaper. Then answer question A-B below.

   a. What are Matt's strengths and weaknesses in reading these two documents? Where possible cite evidence from Matt's comment for your assertions.

   b. What would your next steps be with Matt if you were his teacher?

*This pretest was generously provided by Susan De La Paz and is an adapted form of the pretest used in De La Paz, Monte-Sano, & Felton, (2014b).

Newspaper #1

DESTRUCTION OF THE WAR SHIP MAINE WAS THE WORK OF AN ENEMY…NAVAL OFFICERS THINK THE MAINE WAS DESTROYED BY A SPANISH MINE

Assistant Secretary Roosevelt Convinced the Explosion of the War Ship Was Not an Accident…George Eugene Bryson, the Journal's special correspondent at Havana, cables that it is the secret opinion of many Spaniards in the Cuban capital, that the Maine was destroyed and 258 men killed by means of marine mine or fixed torpeda. This is the opinion of several American naval authorities. The Spaniards, it is believed, arranged to have the Maine anchored over one of the harbor mines. Wires connected the mines with a... magazine, and it is thought the explosion was caused by sending an electric current through the wire. If this can be proven, the brutal nature of the Spaniards will be shown by the fact that they waited to spring the mine after all the men had retired for the night…

Source: Excerpt from *New York Journal and Advertiser*, February 17, 1898

Newspaper #2

MAINE'S HULL WILL DECIDE.

Divers to Find Whether the Force of the Explosion Was from the Exterior or Interior.

WASHINGTON, Feb. 16 – After a day of intense excitement at the Navy Department and elsewhere, growing out of the destruction of the battleship Maine in Havana Harbor last night, the situation at sundown, after the exchange of a number of cablegrams between Washington and Havana, can be summed up in the words of Secretary Long, who when asked as he was about to depart for the day whether he had reason to suspect that the disaster was the work of the enemy, replied: "I do not. In that I am influenced by the fact that Capt. Sigsbee has not yet reported to the Navy Department on the cause. He is evidently waiting to write a full report. So long as he does not express himself, I certainly cannot. I should think from the indications, however, that there was an accident – that the magazine exploded. How that came about I do not know. For the present, at least, no other warship will be sent to Havana."

Source: Excerpt from *New York Times*, February 17, 1898.

Matt's Comment

"Actually, I think it's kind of interesting, because the first document was saying how all the navy officials were certain that it was a mine and now, I guess the next day, they're like, oh we didn't say that. Kind of interesting switch."

5. Read the essay prompt and Lerissa's essay. Then answer questions A-B (see below).

    a. What are the strengths and weaknesses you notice in Lerissa's essay? Where possible cite evidence from Lerissa's essay for the assertions that you make.

    b. What would your next steps be with Lerissa if you were her teacher?

Essay prompt

"The explosion of the U.S.S. Maine caused the United States to invade Cuba in 1898." Use the documents provided and your own knowledge to evaluate this statement. Do you agree with this explanation of the causes of the Spanish American War? Why or why not? Use and cite evidence from the documents to support your analysis of this statement.

Lerissa's Essay

The Spanish American War was a good thing, we got Cuba as a territory, we showed everyone that we are not pushovers, and we showed that we are a dominating force in the world. I agree with the causes of the Spanish American War. Many Cubans were being treated badly in camps said to protect the Cubans. The Maine, a proud ship was sunk, sailors with lives were taken by the "treacherous butchers paid by Spain". The sight of four hundred and sixty women and children thrown on the ground, bodies piled along the ground so much that it is impossible to take one step without walking over a body. The Spanish American War was a necessity and was a good thing.

The truth is, that we are a dominating force in the world today, we control an abundance of land and we are loved by many countries. The fact that we went to war with Spain shows that we can do things for good and not just for ourselves. We can do many things not only for the People of the United States of America but for Cuba, a lonely country needing a defender from the Spanish tyrants. If England, and Germany can govern foreign land so can we.

---

6. Read and think carefully (e.g., mark-up) the following two documents (document 1 & document 2) as you consider the question, "Why did the U.S. invade Cuba in 1898?" Then answer questions A-F (see below).

Using the documents, write an argument in one paragraph in response the following question: "Why did the U.S. invade Cuba in 1898?"

    a. How do you know if your argument is right?

    b. Which document was most helpful to you in responding to the essay question? Why?

c. Which document was least helpful to you in responding to the essay question? Why?

d. Which document is most trustworthy? Why?

Briefly, outline a lesson you could use with your students that includes one or more primary source documents. Include activities or goals for facilitating students' historical thinking, reading, and writing.

Document 1: Reconcentration Camps

By the late 1800s, the Spanish were losing control of their colony, Cuba. Concerned about guerilla warfare in the countryside, they moved rural Cubans to "reconcentration" camps where the Spanish claimed they would be better able to protect them. However, people around the world saw newspaper reports that described horrible conditions in the camps for the Cuban people, who were called "reconcentrados." This account was forwarded to Washington, D.C., by Fitzhugh Lee, U.S. Consul-General in Havana, who said its author was "a man of integrity and character."

SIR: . . .[W]e will relate to you what we saw with our own eyes:
Four hundred and sixty women and children thrown on the ground, heaped pell-mell as animals, some in a dying condition, others sick and others dead, without the slightest cleanliness, nor the least help.

Among the many deaths we witnessed there was one scene impossible to forget. There is still alive the only living witness, a young girl of 18 years, whom we found seemingly lifeless on the ground; on her right-hand side was the body of a young mother, cold and rigid, but with her young child still alive clinging to her dead breast; on her left-hand side was also the corpse of a dead woman holding her son in a dead embrace . . .

The circumstances are the following: complete accumulation of bodies dead and alive, so that it was impossible to take one step without walking over them; the greatest want of cleanliness, want of light, air, and water; the food lacking in quality and quantity what was necessary to sustain life . . .

From all this we deduct that the number of deaths among the reconcentrados has amounted to 77 per cent.

Source: Excerpt from unsigned enclosure included with telegram sent by Fitzhugh Lee, U.S. Consul-General in Cuba, November 27, 1897. Havana, Cuba..

Document 2: March of the Flag

Beveridge gave this speech while he was campaigning to become a senator for Indiana. The speech helped him win the election and made him one of the leading advocates of American expansion.

Fellow citizens, it is a noble land that God has given us; a land that can feed and clothe the world;. . . It is a mighty people that he has planted on this soil . . . It is a glorious history our God has bestowed upon his chosen people; . . .a history of soldiers who carried the flag across the blazing deserts and through the ranks of hostile mountains, even to the gates of sunset; a history of a multiplying people who overran a continent in half a century. . . .

. . . William McKinley is continuing the policy that Jefferson began . . .

The Opposition tells us that we ought not to govern a people without their consent. I answer, The rule of liberty that all just government derives its authority from the consent of the governed, applies only to those who are capable of self-government. I answer, We govern the Indians without their consent, we govern our territories without their consent, we govern our children without their consent. . . .

They ask us how we will govern these new possessions. I answer: . . . If England can govern foreign lands, so can America. If Germany can govern foreign lands, so can America. If they can supervise protectorates, so can America. . . .

What does all this mean for every one of us? It means opportunity for all the glorious young manhood of the republic --the most virile, ambitious, impatient, militant manhood the world has ever seen. It means that the resources and the commerce of these immensely rich dominions will be increased . . .

In Cuba, alone, there are 15,000,000 acres of forest unacquainted with the axe. There are exhaustless mines of iron. . . . There are millions of acres yet unexplored. . . .

It means new employment and better wages for every laboring man in the Union. . . .

Ah! As our commerce spreads, the flag of liberty will circle the globe . . . And, as their thunders salute the flag, benighted peoples will know that the voice of Liberty is speaking, at last, for them; that civilization is dawning, at last, for them --Liberty and Civilization, those children of Christ's gospel . . .

Fellow Americans, we are God's chosen people. . . .

Source: Excerpt from Albert J. Beveridge's Senate campaign speech, September 16, 1898.

This pretest was generously provided by Susan De La Paz and is an adapted form of the pretest used in De La Paz, Monte-Sano, & Felton, (2014b).

Bibliography

Adler, S. (1991). The education of social studies teachers. In J. P. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 210-221). New York, NY: Macmillan.

Alexander, R. J., Osborn, M., & Phillips, D. (2000). Learning from comparing: New directions in comparative educational research (Vol II): Policy, professionals and development. Oxford, England: Symposium Books.

Ashby, R., Lee, P., & Shemilt, D. (2005). Putting principles into practice: Teaching and planning. In M. S. Donovan & J. D. Bransford (Eds.), *How students learn: History in the classroom* (pp. 79–178). Washington, DC: National Academies Press.

Bain, R. (2005). "They thought the world was flat?" Applying the principles of how people learn in teaching high school history. In M. S. Donovan & J. D. Bransford (Eds.), *How students learn: History in the classroom* (pp. 179-213). Washington, D.C.: The National Academies Press.

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal, 93*(4), 373–397.

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond and G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3–31). San Francisco, CA: Jossey-Bass.

Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of teacher education*, *60*(5), 497-511.

Ball, D. L., Sleep, L., Boerst, T., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *Elementary School Journal, 109*(5), 458-474.

Banilower, E., Smith, P. S., Weiss, I. R., & Pasley, J. D. (2006). The status of K-12 science teaching in the United States: Results from a national observation survey. In D. Sunal & E. Wright (Eds.), *The impact of the state and national standards on K-12 science teaching* (pp. 83 − 122). Greenwich, CT: Information Age.

Barton, K. C., McCully, A. W., & Marks, M. L. (2004). Reflecting on elementary children's understanding of history and social studies: An inquiry project with beginning teachers in Northern Ireland and the United States. *Journal of Teacher Education*, 55(1), 70–90.

Barton, K. C., & Levstik, L. S. (2004). *Teaching history for the common good*. Mahwah, NJ: Lawrence Erlbaum.

Becker, H. S. (1990). Generalizing from case studies. In E. W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate (*pp. 233-242). New York: Teachers College Press.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5-25.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-144.

Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theory and methods*. New Jersey: Pearson.

Bohan, C. H., & Davis, O. L. (1998). Historical constructions: How social studies student teachers' historical thinking is reflected in their writing of history. *Theory & Research in Social Education, 26*(2), 173-197.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). Learning and transfer. In *How people learn: Brain, mind, experience, and school* (pp. 31–78). Washington, DC: National Academy Press.

Bransford, J., Derry, S., Berliner, D., Hammerness, K., with Beckett, K. (2005). Theories of learning and their roles in teaching. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do.* (pp. 40-87). San Francisco, CA: Jossey-Bass Publishers.

Brenner, M. E. (2006). Interviewing in educational research. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 357-370). Mahwah, NJ: Lawrence Erlbaum.

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

Cochran, M.M. (2010). *Teaching historical thinking: The challenge of implementing reform-minded practices for three first year teachers*. Retrieved from ProQuest Digital Dissertations. (UMI 3443551).

Cochran-Smith, M., & Lytle, S. L. (1999). Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education, 24,* 249-305.

Coffey, J. E., Edwards, A. R., & Finkelstein, C. (2010). Dynamics of disciplinary understandings and practices of attending to student thinking in elementary teacher education. Proceedings of the 9th International Conference of the Learning Sciences.

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching, 48*(10), 1109-1136.

Conant, J. B. (1963). *The education of American teachers*. New York, NY: McGraw-Hill.

Cruickshank, D. R., & Metcalf, K. K. (1990). Training within teacher preparation. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 469-497). New York: Macmillan Publishing Co.

Cuban, L. (1984). *How teachers taught: Constancy and change in American classrooms, 1890-1980.* New York, NY: Longman.

Cuban, L. (1991). History of teaching in social studies. In J. P. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 197-209). New York, NY: Macmillan.

Davidson, J. W., & Lytle, M. H. (2004). *After the fact: The art of historical detection* (5th ed.). Boston, MA: McGraw Hill.

De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology*, *97*(2), 139.

De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, *35*(3), 174-192.

De La Paz, S., Ferretti, R., Wissinger, D., Yee, L., & MacArthur, C. (2012). Adolescents'
disciplinary use of evidence, argumentative strategies, and organizational structure in
writing about historical controversies. *Written Communication*, *29*(4), 412-454.

De La Paz, S., Felton, M., Monte-Sano, C., Croninger, R., Jackson, C., Deogracias, J. S., &
Hoffman, B. P. (2014). Developing historical reading and writing with adolescent
readers: Effects on student learning. *Theory & Research in Social Education, 42*(2), 228-
274.

Dewey, J. (1904). The relation of theory to practice in education. In C. A. McMurry (Ed.), *The
third yearbook of the National Society for the Scientific Study of Education* (pp. 9-30).
Chicago: University of Chicago Press.

Downey, M. T., & Levstik, L. S. (1991). Teaching and learning history. In J. P. Shaver (Ed.),
Handbook of research on social studies teaching and learning (pp. 400- 410). New
York, NY: Macmillan.

Elton, G. R. (1970). What sort of history should we teach. In Ballard, M. (Ed.), *New movements
in the study and teaching of history* (pp. 221-230). Bloomington: Indiana University
Press.

Engle, R. (2011). The productive disciplinary engagement framework: Origins, key concepts,
and developments. In D. Dai (Ed.), *Design research on learning and thinking in
educational settings*. New York: Taylor & Francis, Educational Psychology Press.

Epstein, T. (2000). Adolescents' perspectives on racial diversity in U. S. history: Case studies
from an urban classroom. *American Educational Research Journal, 37*(1), 185-214.

Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (Eds.). (2006). *The Cambridge
handbook of expertise and expert performance*. New York: Cambridge University Press.

Fehn, B., & Koeppen, K. E. (1998). Intensive document-based instruction in a social studies

    methods course and student teachers' attitudes and practice in subsequent field

    experiences. *Theory & Research in Social Education, 26*(4), 461-484.

Feiman-Nemser, S., & Buchmann, M. (1985). Pitfalls of experience in teacher preparation.

    Teachers College Record, *87*(1), 53-65.

Fennema, E., Franke, M. L., Carpenter, T. P., & Carey, D. A. (1993). Using children's

    mathematical knowledge in instruction. *American Educational Research Journal, 30*(3),

    555 − 583.

Fickel, L. H. (2006). Paradox of practice: Expanding and contracting curriculum in a high stakes

    climate. In S. G. Grant (Ed.), *Measuring history: Cases of state-level testing across the*

    United States (pp. 75-103). Greenwich, CT: Information Age Publishing.

Fielding, N., & Fielding, J. (1986). *Linking data: The articulation of qualitative and quantitative*

    methods in social research. Beverly Hills, CA: SAGE.

Fogo, B. (2012). Core Teaching Practices: Delphi Study. Retrieved from

    https://cset.stanford.edu/research/core-practices.

Forzani, F. (2014). Understanding "core practices" and "practice-based" teacher education:

    Learning from the past. *Journal of Teacher Education, 65*(4), 357-368.

Franke, M. L., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative

    growth: A follow-up study of professional development in mathematics. *American*

    Educational Research Journal, 38, 653-689.

Fraser, J. W. (2007). *Preparing America's teachers: A history*. New York: Teachers College

    Press.

Ghousseini, H. (2009). Designing opportunities to learn to lead classroom mathematics

    discussions in pre-service teacher education: Focusing on enactment. In D. Mewborn &

    H. Lee (Eds.), *Scholarly practices and inquiry in the preparation of mathematics*

    teachers (pp. 147–158). San Diego, CA: Association of Mathematics Teacher Educators.

Ghousseini, H., & Sleep, L. (2011). Making practice studyable. *ZDM Mathematics Education,*

    43(1), 147 – 160.

Gipps, C. (1994). *Beyond testing: Toward a theory of educational assessment*. London: Falmer

    Press.

Gitlin, T. (1995). *The twilight of common dreams: Why America is wracked by culture wars*.

    New York, NY: Metropolitan.

Goetz, J. P., & LeCompte, M. D. (1984). *Ethnography and qualitative design in educational*

    research. Orlando: Academic Press Inc.

Goodlad, J. (1984). *A place called school.* New York, NY: McGraw-Hill.

Goodlad, J. I. (1990). *Places where teachers are taught*. San Francisco, CA: Jossey Bass.

Goodwin, A. L., & Kosnik, C. (2013). Quality Teacher Educators= Quality Teachers?

    Conceptualizing essential domains of knowledge for those who teach teachers. *Teacher*

    Development, *17*(3), 334-346.

Grant, S. G. (2001). It's just the facts, or is it? The relationship between teachers' practices and

    students' understandings of history. *Theory & Research in Social Education, 29*(1), 65-

    108.

Grant, S. G. (2003). *History lessons: Teaching, learning, and testing in US high school*

    classrooms. Mahwah, N.J.: Lawrence Erlbaum Associates

Grant, S. G. (Ed.). (2006). *Measuring history: Cases of state-level testing across the United States*. Greenwich, CT: Information Age Publishing.

Grossman, P. L. (1992). Why models matter: An alternate view on professional growth in teaching. *Review of Educational Research*, 171-179.

Grossman, P. (2005). Research on pedagogical approaches in teacher education. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher education* (pp. 425-476). Washington, D.C.: American Educational Research Association

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record, 111*(9), 2055-2100.

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers & Teaching, 15*(2), 273-289.

Grossman, P., & McDonald, M. (2009). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal, 45*(1), 184-205.

Grossman, P., Smagorinsky, P., & Valencia, S. W. (1999). Appropriating tools for teaching English: A theoretical framework for research on learning to teach. *American Journal of Education, 108*(1), 1-29.

Grossman, P., Valencia, S. W., Evans, K., Thompson, C., Martin, S., & Place, N. (2000). Transitions into teaching: Learning to teach writing in teacher education and beyond. *Journal of Literacy Research, 32*(4), 631-662.

Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a theory of teacher community. *Teachers College Record, 103*(6), 942-1012.

Hallam, R. N. (1967). Logical thinking in history. *Educational Review*, *19*(3), 183-202.

Hartzler-Miller, C. (2001). Making sense of "best practice" in teaching history. *Theory & Research in Social Education, 29*(4), 672-695.

Hatch, T., & Grossman, P. (2008). Learning to look beyond the boundaries of representation: Using technology to examine teaching. *Journal of Teacher Education, 60*(1), 70-85.

Hawkey, K., & Prior, J. (2011). History, memory cultures and meaning in the classroom. *Journal of Curriculum Studies, 43*(2), 231-247.

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment*? Educational Measurement: Issues Practice, 28*(3), 24-31.

Hicks, D., Doolittle, P., & Lee, J. K. (2004). Social studies teachers' use of classroom-based and web-based historical primary sources. *Theory & Research in Social Education*, *32*(2), 213-247.

Hogan, T., Rabinowitz, M., & Craven, J. (2003). Representation in teaching: Inferences from research of expert and novice teachers. *Educational Psychologist, 38*(4), 235–247.

Holt, T. (1990). *Thinking historically: Narrative, imagination, and understanding*. New York, NY: College Board.

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for research in mathematics education*, 258-288.

Janssen, F. J. J. M., Westbroek, H. B., Doyle, W., & Van Driel, J. H. (2013). How to make innovations practical. *Teachers College Record*, *115*(7), 1-43.

Johnson, T. S., Thompson, L., Smagorinsky, P., & Fry, P. G. (2003). Learning to teach the five-paragraph theme. *Research in the Teaching of English*, 136-176.

Kagan, D. M. (1992). Professional growth among preservice and beginning teachers. *Review of Educational Research, 62*(2), 129-169.

Kammen, M. (1989). History is our heritage: The past in contemporary American culture. In P. Gagnon & B. Commission (Eds.), *Historical literacy: The case for history in American education*. Boston, MA: Houghton Mifflin.

Kazemi, E., Lampert, M., & Franke, M. F. (2009). Developing pedagogies in teacher education to support preservice teachers' ability to enact ambitious teaching. In R. Hunter, B. Bickness & T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 12-30). Palmerston North, NZ: MERGA.

Kazemi, E., Lampert, M., & Ghousseini, H. (2007). Conceptualizing and using routines of practice in mathematics teaching to advance professional education. Chicago: Report to the Spencer Foundation.

Kazemi, H. (2015, January). *Organizing for teacher learning: Innovations in methods instruction and school-based professional development in elementary mathematics*. Invited Presentation at the University of Maryland CfME Colloquium, College Park, MD.

Kennedy, M. (2005). *Inside teaching: How classroom life undermines reform*. Cambridge, MA: Harvard University Press.

Koeppen, K. E. (1998). The experiences of a secondary social studies student teacher: Seeking security by planning for self. *Teaching and Teacher Education, 14*(4), 401-411.

Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing*. London: SAGE.

Labaree, D. F. (2004). *The trouble with ed schools*. New Haven, CT: Yale University Press.

Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review, 55*(2), 178-194.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven: Yale University Press.

Lampert, M. (2009). Learning teaching in, from, and for practice: What do we mean? *Journal of Teacher Education, 61*(1-2), 21-34

Lampert, M., Beasley, H., Ghousseini, H., Kazemi, E., Franke, M. (2010) Using designed instructional activities to enable novices to manage ambitious mathematics teaching (pp. 129-141). In M.K. Stein & L. Kucan (Eds.) *Instructional explanations in the discipline.* New York: Springer.

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Chan Turrou, A., Beasley, H., . . . Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education, 64*(3), 226- 243.

Lampert, M., & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning. *The Elementary School Journal*, *109*(5), 491-509.

Lapp, M. S., Grigg, W., & Tay-Lim, B. S.-H. (2002). *The nation's report card: U.S. history 2001.* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/main2001/2002483.pdf

Lee, J., & Weiss, A. (2007). The nation's report card: U.S. history 2006 (NCES 2007–474). Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Lee, P. (2005). Putting principles into practice: Understanding history. In S. Donovan & J. Bransford (Eds.), *How students learn: History in the classroom*. Washington D.C.: National Academic Press.

Lee, P., & Dickinson, A. J. (1984). Making sense of history. In P. Lee & A. J. Dickinson (Eds.), *Learning history* (pp. 117–153). London: Heinemann.

Lee, P., & Shemilt, D. (2003). A scaffold, not a cage: Progression and progression models in history. *Teaching History, 113*(1), 13-23.

Leinhardt, G., & Steele, M. (2005). Seeing the complexity to standing to the side: Instructional dialogues. *Cognition and Instruction,* 23(1), 87 – 163.

Lemov, D. (2010). *Teach like a champion: 49 techniques that put students on the path to college*. San Francisco, CA: Jossey-Bass.

Levesque, S. (2008). *Thinking historically: Educating students in the 21st century.* Toronto, Canada: University of Toronto Press.

Levin, D. M., & Richards, J. (2010). Exploring how novice teachers learn to attend to students' thinking in analyzing case studies of classroom teaching and learning. Proceedings of the 9th International Conference of the Learning Sciences.

Levin, D. M., Hammer, D., & Coffey, J. E. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education, 60*(2), 142-154.

Levstik, L., & Barton, K. (1997). Doing history: Investigating with children in elementary and middle schools. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Lowenthal, D. (1985). *The past is a foreign country*. Cambridge, UK: Cambridge University Press.

Matsko, K. K., & Hammerness, K. (2014). Unpacking the "urban" in urban teacher education: Making a case for context-specific preparation. *Journal of Teacher Education, 65*(2), 128-144.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach*. Washington DC: SAGE.

McConney, A. A., Schalock, M. D., & Schalock, H. D. (1998). Focusing improvement and quality assurance: Work samples as authentic performance measures of prospective teachers' effectiveness. *Journal of Personnel Evaluation in Education*, *11*(4), 343-363.

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.

McDonald, M., Kazemi, E., Kelley-Petersen, M., Mikolasy, K., Thompson, J., Valencia, S. W., & Windschitl, M. (2014). Practice makes practice: Learning to teach in teacher education. *Peabody Journal of Education, 89*(4), 500-515.

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

Metzger, M. (1998). Teaching reading. *Phi Delta Kappan*, *80*(3), 240-248.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* Sage.

Mink, L. O. (1987). *Historical Understanding*. Cornell Press: Ithica, NY.

Monte-Sano, C. (2008). Qualities of effective writing instruction in history classrooms: A cross-case comparison of two teachers' practices. *American Educational Research Journal, 45*(4), 1045-1079.

Monte-Sano, C. (2011a). Learning to open up history for students: Preservice teachers' emerging pedagogical content knowledge. *Journal of Teacher Education, 62*(3), 260-272.

Monte-Sano, C. (2011b). Beyond reading comprehension and summary: Learning to read and write in history by focusing on evidence, perspective, and interpretation. *Curriculum Inquiry, 41*(2), 212-249.

Monte-Sano, C., & Budano, C. (2013). Developing and enacting pedagogical content knowledge for teaching history: An exploration of two novice teachers' growth over three years. *Journal of the Learning Sciences, 22*(2), 171-211.

Monte-Sano, C., & Cochran, M. (2009). Attention to learners, subject, or teaching: What takes precedence as preservice candidates learn to teach historical thinking and reading? *Theory & Research in Social Education, 37*(1), 101-135.

Monte-Sano, C., De La Paz, S., & Felton, M. (2014a). *Reading, thinking, and writing about history: Teaching argument writing to diverse learners in the common core classroom, grades 6-12*. New York: Teachers College Press.

Monte-Sano, C., De La Paz, S., & Felton, M. (2014b). Implementing a disciplinary-literacy curriculum for US history: Learning from expert middle school teachers in diverse classrooms. *Journal of Curriculum Studies, 46*(4), 1-36.

Monte-Sano, C., & Harris, K. (2012). Recitation and reasoning in novice history teachers' use of writing. *Elementary School Journal, 113*(1), 105-130.

Murray, J., & Male, T. (2005). Becoming a teacher educator: Evidence from the field. *Teaching and teacher education*, *21*(2), 125-142.

National Center for History in the Schools. (1994). *National standards for United States history: Exploring the American experience.* Los Angeles, CA: Author.

National Council for the Social Studies [NCSS]. (n.d.). *The College, Career, and Civic Life (C3) Framework for Social Studies State Standards: Guidance for Enhancing the Rigor of K-12 Civics, Economics, Geography, and History* (Silver Spring, MD: NCSS, 2013).

Nokes, J. D., Dole, J. A., & Hacker, D. J. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology, 99*(3), 492–504.

Nokes, J. D. (2012). *Building Students' Historical Literacies: Learning to Read and Reason with Historical Texts and Evidence*. New York: Routledge.

Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S., & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse processes*, *35*(2), 135-198.

Ogren, C. A. (2005). *The American state normal school.* New York, NY: Palgrave Macmillan.

Patton, M. Q. (1990). *Qualitative evaluation and research methods.* Newbury Park, CA: SAGE.

Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative Social Work*, *1*(3), 261-283.

Peck, C., & Seixas, P. (2008). Benchmarks of historical thinking: First steps. *Canadian Journal of Education, 31*(4), 1015-1038.

Pierson, J. (2008). *The relationship between patterns of classroom discourse and mathematics learning*. Unpublished doctoral dissertation, University of Texas at Austin.

Popham, J. W., (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Ravitch, D. (1990). Multiculturalism: E pluribus plures. *American Scholar, 59*(3), 337–354.

Ravitch, D., & Finn, C. (1987). *What do our 17-year-olds know?* New York, NY: Harper and Row.

Reisman, A. (2010). *Document-based whole class discussion.* TeachingHistory.org. Retrieved from http://teachinghistory.org/teaching-materials/teaching-guides/25620

Reisman, A. (2012a). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction*, *30*(1), 86-112.

Reisman, A. (2012b). The "document-based lesson': Bringing disciplinary inquiry into high school history classrooms with adolescent struggling readers. *Journal of Curriculum Studies, 44*(2), 233-264.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *Handbook of research on teacher education* (pp. 102-119). New York: Simon & Schuster/Macmillan.

Robertson, A. D., Atkins, L. J., Levin, D. M., & Richards, J. (in press). What is responsive teaching? To appear in A. D. Robertson, R. E. Scherr, & D. Hammer (Eds.), *Responsive teaching in science*. London: Routledge.

Rosenzweig, R. (2000) How Americans use and think about the past: Implications from a national survey for the teaching of history. In P. N. Stearns, S. S. Wineburg, & P. C. Seixas (Eds.), *Knowing, teaching and learning history: National and international perspectives* (262–283). New York: New York University Press.

Ryals, K. (1975). *Competency-based teacher education edited by R. Houston and R. Howsam.* Chicago: Science Research Associates.

Sarason, S. B. (1993). *The case for change: Rethinking the preparation of educators.* San Francisco, CA: Jossey-Bass.

Schlesinger, A. M., Jr. (1992). *The disuniting of America.* New York, NY: Norton.

Seixas, P. (1994). Confronting the moral frames of popular film: Young people respond to historical revisionism. *American Journal of Education, 102*(3), 261-285.

Seixas, P. C. (1996). Conceptualizing the growth of historical understanding. In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development: New models of learning, teaching, and schooling.* Cambridge, MA: Blackwell Publishers.

Seixas, P. (1998). Student teachers thinking historically. *Theory & Research in Social Education, 26*(3), 310-341.

Seixas, P. C. (2000). Schweigen! Die kinder! In P. N. Stearns, P. C. Seixas, & S. S. Wineburg (Eds.), *Knowing, teaching, and learning history: National and international perspectives.* New York: New York University Press.

Seixas, Peter (2006). *Benchmarks of historical thinking: A framework for assessment in Canada.* University of British Columbia: Centre for the Study of Historical Consciousness. Retrieved from http://historicalthinking.ca/sites/default/files/Framework.Benchmarks.pdf

Seixas, P., & Morton, T. (2013). *The big six.* Toronto, Ontario: Nelson Education.

Shavelson, R. J. (2006). *On the integration of formative assessment in teaching and learning with implications for teacher education.* Paper prepared for the Stanford Education Assessment Laboratory and the University of Hawaii Curriculum Research and Development Group. Retrieved from: www.stanford.edu/dept/SUSE/SEAL.

Shemilt, D. (1983). The devil's locomotive. *History and Theory*, *22*(4), 1–18.

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

Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education, 20*(2), 163-183.

Singer-Gabella, M., Cartier, J., Forman, E., Knapp, N., Kannan, P., Shahan, E., Lancaster, L., Barrick, N. (2009). Contextualizing learning progressions for prospective elementary teachers of mathematics and science. Paper presented at the Annual Meeting, American Educational Research Association, San Diego, CA.

Singer-Gabella, M., Stengel, B., Shahan, E., & Kim, M. J. (in press). Learning to leverage student thinking: What novice approximations teach us about ambitious practice. *The Elementary School Journal*.

Sleep, L., & Boerst, T. (2012). Preparing beginning teachers to elicit and interpret students' mathematical thinking. *Teaching and Teacher Education, 28*(7), 1038-1048.

Slick, S. K. (1998). The university supervisor: A disenfranchised outsider. *Teaching and Teacher Education*, *14*(8), 821-834.

Smagorinsky, P., Cook, L., & Johnson, T. (2003). The twisting path of concept development in learning to teach. *Teachers College Record*, *105*(8), 1399-1436.

Smith, B. O. (1980). *A design for a school of pedagogy*. Washington, DC: US Government Printing Office.

Solomon, J. (2009). The Boston teacher residency: District-based teacher education. *Journal of Teacher Education, 60*(5), 478-488.

Stake, R. E. (1995). *The art of case study research*. London: SAGE.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive

      mathematical discussions: Helping teachers learn to better incorporate student thinking.

      Mathematical Thinking and Learning, *10*(4), 313-340.

Stanford History Education Group (n.d.). *Core Assessments*. Retrieved from:

      http://beyondthebubble.stanford.edu/assessments

Stein, M.K., Engle, R.A., Smith, M.S., & Hughes, E.K. (2008). Orchestrating productive

      mathematical discussions: Helping teachers learn to better incorporate student thinking.

      Mathematical Thinking and Learning, 10(4), 313-340.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for

      improving education in the classroom.* New York: Free Press.

Teachingworks.org (2013). *High Leverage Practices.* Retrieved from

      http://www.soe.umich.edu/academics/bachelors_degree_programs/uete/uete_hl/.

TEDD: The Design. (2014, January 1). Retrieved May 1, 2014, from http://tedd.org/approach/

Thompson, J., Windschitl, M., & Braaten, M. (2013). Developing a theory of ambitious early-

      career teacher practice. *American Educational Research Journal, 50*(3), 574-615.

Valencia, S. W., Martin, S. D., Place, N. A., & Grossman, P. (2009). Complex interactions in

      student teaching: Lost opportunities for learning. *Journal of Teacher Education, 60*(3),

      304-322.

Valli, L., & Rennert-Ariev, P. (2002). New standards and assessments? Curriculum

      transformation in teacher education. *Journal of Curriculum Studies, 34*(2), 201-225.

Van den Berg, R. (2002). Teachers' meanings regarding educational practice. *Review of

      Educational Research*, *72*(4), 577-625.

van Hover, S. (2006). Teaching history in the old dominion: The impact of Virginia's accountability reform on seven secondary beginning history teachers. In S. G. Grant (Ed.), *Measuring history: Cases of state-level testing across the United States* (pp. 195-220). Greenwich, CT: Information Age Publishing.

van Hover, S., & Yeager, E. (2003). Making students better people? A case study of a beginning history teacher. *The International Social Studies Forum, 3*(1), 219-232.

VanSledright, B. (2002). Confronting history's interpretive paradox while teaching fifth graders to investigate the past. *American Educational Research Journal, 39*(4), 1089-1115.

VanSledright, B. (2002a). Fifth graders investigating history in the classroom: Results from a researcher-practitioner design experiment. *The Elementary School Journal, 103*(2), 131-160.

VanSledright, B. (2002b). *In search of America's past: Learning to read history in elementary school.* New York, NY: Teachers College Press.

VanSledright, B. (2004). What does it mean to think historically... and how do you teach it? *Social Education, 68*(3), 230-233.

VanSledright, B. (2008). Narratives of nation-state: Historical knowledge, and school history education. *Review of Research in Education, 32*(1), 109-146.

VanSledright, B. A., & Limón, M. (2006). Learning and teaching social studies: A review of cognitive research in history and geography. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 545–570). Mahwah, NJ: Lawrence Erlbaum.

Wiggins, G. P., & McTighe, J. (2005). *Understanding by design.* Alexandria, VA: Association for Supervision and Curriculum Development.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 18-40). New York, NY: Routledge.

Wilson, E. K., Konopak, B. C., & Readence, J. E. (1994). Preservice teachers in secondary social studies: Examining conceptions and practices. *Theory & Research in Social Education, 22*(3), 364-379.

Wilson, S. M., & McDiarmid, G. W. (1996). Something old, something new: What do social studies teachers need to know? In F. B. Murray (Ed.), *The teacher educator's handbook: Building a knowledge base for the preparation of teachers* (pp. 295-319). San Francisco: Jossey-Bass Publishers.

Wilson, E. K., & Saleh, A. (2000). The effects of an alternative model of student teaching supervision on clinical master teachers. *Action in Teacher Education*, *22*(2A), 84-90.

Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education, 96*(5), 878-903.

Wineburg, S. (1991a). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*(1), 73-87.

Wineburg, S. (1991b). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal, 28*(3), 495-519.

Wineburg, S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past.* Philadelphia, PA: Temple.

Wineburg, S. (2004). Crazy for history. *Journal of American History, 90*(4), 1401-1414.

Wineburg, S., & Martin, D. (2004). Reading and rewriting history. *Educational Leadership, 62*(1), 42-45.

Wood, T. (1998). Funneling or focusing? Alternative patterns of communication in mathematics class. In H. Steinbring, M. G. Bartolini-Bussi, & A. Sierpinska (Eds.), *Language and communication in the mathematics classroom* (pp. 167-178). Reston, VA: National Council of Teachers of Mathematics.

Yeager, E. A., & Davis, O. L. (1995). Between campus and classroom: Secondary student-teachers' thinking about historical texts. *Journal of Research and Development in Education, 29*(1), 1-8.

Yeager, E. A., & Davis, O. L. (1996). Classroom teachers' thinking about historical texts: An exploratory study. *Theory & Research in Social Education, 24*(2), 146-166.

Yin, R. K. (2003). *Case study research: Designs and methods* (3 ed.). London: SAGE.

Yin, R. K. (2006). Case study methods. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 111-122). Mahwah, NJ: Lawrence Erlbaum.

Zeichner, K. (2005). Learning from experience with performance-based teacher education. In F. P. Peterman (Ed.), *Designing performance assessment systems for urban teacher prep* (pp. 3-19). Mahweh, NJ: Lawrence Erlbaum.

Zeichner, K. (2005a). Becoming a teacher educator: A personal perspective. *Teaching and Teacher Education*, *21*(2), 117-124.

Zeichner, K. (2012). The turn once again toward practice-based teacher education. *Journal of Teacher Education*, *63*(5), 376-382.