Teaching Citizens: Exploring the Relationships Between Teacher Professional Learning, Interactive Civics, and Student Achievement on NAEP Civics

Author: Joshua Simon Littenberg-Tobias

Boston College
Lynch School of Education

Department of Educational Research, Measurement, and Evaluation

TEACHING CITIZENS: EXPLORING THE RELATIONSHIPS BETWEEN TEACHER PROFESSIONAL LEARNING, INTERACTIVE CIVICS, AND STUDENT ACHIEVEMENT ON NAEP CIVICS

Dissertation by

JOSHUA LITTENBERG-TOBIAS

submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

Dissertation Chair: Dr. Laura O'Dwyer

MAY 2015

Abstract

Teaching Citizens: Exploring the Relationships between Teacher Professional Learning, Interactive Civics, and Student Achievement on NAEP Civics

Joshua Littenberg-Tobias

Dr. Laura O'Dwyer, Chair

Youth civic participation is at alarmingly low levels. In 2014, nearly 80% of eligible 18-29 year-olds did not vote in the midterm election (CIRCLE, 2014). Other forms of civic engagement are also at starkly low levels: less than one in ten 18-29 year-olds report contacting a public official, boycotting a product, or frequently expressing political opinions on the internet (U.S. Census Bureau, 2014). Historically, schools have been tasked with preparing students with the knowledge and skills to be active democratic citizens. However, few studies have examined the role of teachers in fostering students' civic knowledge and skills.

This study used data from the 2010 National Assessment of Educational Progress (NAEP) 8th grade civics assessment to analyze the relationships between teacher participation in professional learning, use of interactive instructional practices, and student achievement in civics. Participation in professional learning significantly predicted both interactive instruction and student achievement: a one standard deviation increase in professional learning was associated with a predicted .32 standard deviation increase in interactive instructional practices, and a predicted .045 standard deviation increase in student achievement. There was no significant difference between more traditional and communities of practice based forms of professional development in their relationships with interactive instructional practices and student civic achievement. Interactive instructional practices were also significantly associated with increases in student achievement on NAEP civics, but the effect size was small: a one

standard deviation increase in interactive instruction was related to a predicted .03 standard deviation increase in student achievement. Moreover, the relationship between interactive instruction and student achievement was curvilinear; high levels of interactive instruction were associated with decreases in student achievement. The study did not find any evidence that teacher participation in professional learning increased the effectiveness of interactive instructional practices.

Acknowledgements

This dissertation is about civics and the creation of this dissertation was a civic endeavor that required the participation of many different individuals. I would like to thank the many people who helped bring this dissertation to fruition. To my chair, Laura O'Dwyer, thank you for providing thoughtful feedback and advice throughout the dissertation process and for ensuring that I stayed focused on the end goal. To Lauren Saenz, thank you for serving on my committee and offering critical feedback when it was needed. To Larry Ludlow, thank you for reading over the early drafts of the dissertation and for always being a supportive department chair. To Peter Levine, Kei Kawashima-Ginsberg, Felicia Sullivan, and the rest of the staff at CIRCLE, thank you for fostering my interest in civic engagement and for supporting me in my research. To Vincent Cho, thank you for being an incredible research advisor and mentor and for helping me develop as writer, researcher, and scholar.

I would also like to thank my parents, Susan and David, and my brother Ethan. You have always supported me in my studies, even when you weren't quite sure what it was I was studying in graduate school. I am so grateful to have such a loving and supportive family.

This dissertation is dedicated to my wife, Liza. Thank you for reading through the many drafts of this dissertation and always being supportive of my dissertation research, even when it meant that I had to spend hours holed up in my office. Thank you also for being an amazing, loving, and supportive partner—you always seems to know what I need, often before I even realize it! I am so thankful that I have been able to start a family with you and now that my dissertation has been completed I am looking forward to spending more time with you and our baby Stella.

Table of Contents

Tables and Figures

Chapter 1

The conception of education as a social process and function has no definite meaning until we define the kind of society we have in mind

-John Dewey (1916)

There is growing concern about the state of civic life in American society. Compared with earlier generations, American adults are less likely to join community organizations (R.D. Putnam, 2001), participate in elections (Timpone, 1998), and trust public institutions (Uslaner & Brown, 2005). Part of the decline in civic engagement is generational. Millennials (18-29 year-olds) are less likely to join organizations, work with neighbors, attend community meetings, or read the news than older cohorts (CIRCLE, Harvard Institute of Politics, and Mobilize.org, 2012). Millennials are also less likely than their Baby Boomer parents to have participated in a protest or signed a petition (Caren, Ghoshall, & Ribas, 2011). Youth civic participation is at alarmingly low levels. From 2008-2013, less than 10% of 18-29 year-olds reported contacting public officials, boycotting products, or frequently expressing opinions over the internet in the past year[1] (U.S. Census Bureau, 2014).

Many young Americans are also uninformed about how the government functions and about basic public policy questions. For example, in 2008 only 54% of 16-24 year-olds knew that the Supreme Court decided if a law was constitutional and only 47% knew that a two-thirds majority of the House and Senate is needed to override a Presidential veto (U.S. Census Bureau, 2009). Another study found that only 29% of 18-24 year-old respondents knew that the government spent more money on Social Security than it did on foreign aid (CIRCLE Staff, 2013).

[1] Analysis conducted by the author.

Moreover, many youth are not developing the analytical and communication skills to have informed debates and discussions about civic issues (Campaign for the Civic Mission of Schools, 2011; Hess, 2009) From 2008-2013, only 27% of 18-29 year-olds reported frequently talking about politics with friends and family within the past year, and far fewer (9%) frequently shared political opinions on the internet (U.S. Census Bureau, 2014).

Nonetheless, what might be most troubling are the large gaps in civic engagement by race, socio-economic status, and educational attainment (CIRCLE, Harvard Institute of Politics, and Mobilize.org, 2012; Flanagan, 2004; Hart & Atkins, 2002; Levinson, No Citizen Left Behind, 2012). Millennials with college degrees are more likely to contact a public official, serve in a leadership position in an organization, or boycott a product (CIRCLE, Harvard Institute of Politics, and Mobilize.org, 2012). White Millennials are also more than twice as likely as African American and Hispanic youth to have contacted a public official or boycotted a product in the past year (U.S. Census Bureau, 2014).[2]

The current anemic level of youth civic engagement in the United States may have negative implications for the future of American democracy. One of the characteristics of a democratic society is that citizens engage in informed debate on public issues (Delli Carpini & Keeter, 1996; Gutmann & Thompson, 2004). When many citizens lack access to information and are disengaged from the political process, public debates are more likely to be dominated by divisiveness and inaction and make productive discourse impossible (Campaign for the Civic Mission of Schools, 2011). Additionally, large gaps in civic knowledge and participation by race and socio-economic status raise concerns about the diversity of civic society in this country (Campaign for the Civic Mission of Schools, 2011).

[2] Analysis conducted by the author.

Statement of the Problem

Researchers often argue for better school-based civic education as a way of increasing youth civic engagement (Carnegie Corportation of New York and CIRCLE, 2003; Kahne & Sporte, 2008; Torney-Purta, 2002). From a policy perspective this makes intuitive sense. Historically, one of the purposes of the public school system was to prepare students for democratic citizenship (Carnegie Corportation of New York and CIRCLE, 2003; Dewey, 1903; Hinde, 2008). One-fourth of state constitutions explicitly describe preparation for citizenship as the rationale for a system of public instruction and more than one-half have statutes explicitly addressing civic education (Tolo, 1999). These ideals continue to be enshrined in state civic education requirements. Forty-five states require students to complete a civics or government course in order to graduate from high school (CIRCLE, 2014)

Advocates describe the goal of civic education as twofold: increasing civic knowledge and developing civic skills (Carnegie Corportation of New York and CIRCLE, 2003; Torney-Purta, 2002). Civic knowledge can be defined as the understanding of how public policy is created through political institutions, processes, leaders and parties (Delli Carpini & Keeter, 1996). Some examples of civic knowledge include understanding how laws are passed by Congress or being able to identify a political party's position on an issue. Civic skills describe the abilities citizens need to think and act effectively in response to societal problems (National Assessment Governing Board, 2010). These include being able to evaluate arguments and articulate positions on public issues (National Assessment Governing Board, 2010; Torney-Purta, 2002) as well as being able to organize others to engage in collective action on social issues (Ginwright & Cammarota, 2002; Westheimer & Kahne, 2004).

Researchers theorize that these skills and knowledge may prepare students to be more active citizens in a number of ways. Understanding public policy issues allows citizens make informed decisions about which policies and candidates to support (Bartels, 1996). Students who understand how government works are also better prepared to advocate for issues that are important to them as citizens (Brady, Verba, & Schlozman, 1995; Westheimer & Kahne, 2004). Civic education also teaches students how to make reasoned arguments and collaborate with others, which are important civic qualities in democratic societies (Enslin, Pendlebury, & Tjiattas, 2001).

Empirical evidence largely supports these claims. Students with greater civic knowledge are more likely to indicate that they will vote when they become eligible (A. Cohen & Chafee, 2012; Torney-Purta & Amadeo, 2003). This relationship between civic knowledge and civic participation continues later in life; adults with higher levels of civic knowledge are more likely to vote, volunteer for political campaigns, and contribute money to a political candidate (Delli Carpini & Keeter, 1996; Popkin & Dimock, 1999). Civically knowledgeable adults are also more likely to be ideologically consistent in their opinions and more likely to support policies that reflect their self-interest (Bennett, 2003; Delli Carpini & Keeter, 1996). Adults with greater civic skills, such as planning meetings or communicating ideas, are also more likely to participate in politics (Brady, Verba, & Schlozman, 1995).

However, researchers have failed to find a consistent link between learning civics in school and civic attitudes or behavior (Manning & Edwards, 2013). In their seminal study, Langston and Jennings (1968) did not find any significant relationship between high school civics credits and students' political efficacy or interest in civic participation. More recent studies have reported mixed finding as to the relationship between civics course-taking and civic

engagement. For example, Callahan, Muller, and Schiller (2010) did not find a significant relationship between the number of social studies credits a student took and their likelihood of voting. In contrast, Bachner (2010) found a positive and significant relationship between a taking full-year civics course and the students' likelihood of voting when they became eligible. However, the relationship was modest (an increase of 3%-6%) and was not consistently statistically significant across all of the election years examined in the study (Bachner, 2010; Manning & Edwards, 2013).

Some researchers have argued that it is not civics courses alone, but interactive instruction, that are related to increased civic engagement (Callahan, Muller, & Schiller, 2010; CIRCLE, 2013; Kahne & Sporte, 2008; Torney-Purta, 2002). Civic educators have developed a number of interactive activities designed to teach civics such as mock trials, structured discussions of controversial issues, and community service projects. For example, the Boston Debate League works with teachers to set up classroom debates in order to teach students evidence-based argumentation (Belanger & Stein, 2013). Generation Citizen engages middle and high school students in identifying and taking action to address important issues in their local community (Pope, Stolte, & Cohen, 2011).

Advocates for these types of activities claim that they allow students to better understand civic concepts and develop civic skills than traditional lecture-based civics lessons (Hess, 2009; Levinson, 2012; McDevitt & Kiousis, 2006). They argue that exposing students to different forms of civic participation while they are in school will lead to increased civic engagement when they are adults (Finlay, Wray-Lake, & Flangan, 2010; Levine & Higgins-D'Alessando, 2010).

Indeed, research on interactive civics activities has found that students who report engaging in interactive civics activities have greater civic knowledge and skills, express more positive attitudes toward civic participation, and have higher levels of civic engagement (CIRCLE, 2013; Kahne & Sporte, 2008; Kawashima-Ginsberg, 2013; Torney-Purta, 2002; T. Zhang, Torney-Purta, & Barber, 2012). Moreover, numerous quasi-experimental and experimental intervention studies have found that exposing students to interactive civics activities increases civic outcomes (Kahne, Chi, & Middaugh, 2006; McDevitt & Kiousis, 2006; Syvertesen, et al., 2009).

Yet many teachers do not currently use interactive civics activities in their classrooms. A recent study found that only 53% of 8th graders reported discussing current events on a weekly basis and less than half (48%) reported having ever participated in a civic simulation such as a mock trial (Kawashima-Ginsberg, 2013). Another study found that 90% of U.S. students reported that the activities that they spent the most time on in civics classes was reading textbooks and doing worksheets (Kahne, Chi, & Middaugh, 2006). Opportunities to participate in interactive civics are also distributed unequally by race and socio-economic status. Poor students and students of color have fewer opportunities to participate in these types of activities (Kahne & Middaugh, 2008; Kawashima-Ginsberg, 2013).

The lack of interactive civics activities within schools and the wide disparities by race and socio-economic status have led some civic researchers and educators to call for more professional learning opportunities for civics teachers (Campaign for the Civic Mission of Schools, 2011; CIRCLE, 2013). Research on teacher professional development suggests that professional learning activities, such as discussing instruction with colleagues or joining educator networks, are related to more student-centered and interactive approaches to pedagogy (Bryk,

Camburn, & Louis, 1999; Louis & Marks, 1998; Lieberman & McLaughlin, 1992). However, there is a lack of research on the relationship between participating in professional learning activities and the use of interactive civics in the classroom. Additionally, there has been almost no research that examines whether teacher professional learning is related to increased student achievement in civics.

This dissertation examined the relationship between teacher professional learning, interactive civics, and students' civic skills and knowledge as measured by the National Assessment of Educational Progress (NAEP) civics assessment. It explored whether 8th grade teachers who engaged in professional learning activities were more likely to use interactive civics activities in their classes. Furthermore, it investigated whether teachers' engagement in professional learning activities were associated with higher levels of student civic skills and knowledge. Finally, it examined whether teacher professional learning moderated the relationship between teachers' use of interactive civics activities and their students' civic skills and knowledge.

Placement in the Field

The research described in this dissertation may contribute to the literature in a number of ways. First, it will provide additional insight into how interactive instructional practices in civics classes are related to increased student civic outcomes. Although there has been substantial research in the civic education field about the relationship between interactive civics experiences and student civic outcomes, most of the research on interactive civics activities has focused on individual classroom practices or interventions (e.g., Hess 2009; McDevitt & Kiousis 2006; Svyertesen et al., 2009). Relatively few studies have examined the relationship between teachers' use different forms of interactive civics activities and students' civic outcomes. As a result, there

is scarce empirical evidence about whether participating more frequently in interactive civics activities in school is associated with more positive student civic outcomes.

This dissertation will also connect two distinct areas of research that have not been previously examined together: teacher professional learning and civic education. Most existing studies on teacher professional learning have focused on reading or math (Lee & Smith, 1996; Lubienski, Lubienski, & Crane, 2008; Pil & Leana, 2009; Vescio, Ross, & Adams, 2008). Although there have been a few evaluations of teacher professional development program in civics (e.g., Barr, 2010) researchers have not previously investigated whether professional learning activities as a whole are associated with more interactive civics practices and increased student civic achievement.

Additionally, most civic education studies have been conducted within specific schools or districts. There have been very few studies that have examined civics instruction using a nationally representative sample of American students. NAEP uses random sampling procedures to generate a sample of American students that is nationally representative of the country as a whole (National Center for Education Statistics, 2010). This allows for inferences that have strong external validity to the population of American students within specific grades (Shadish, Cook, & Campbell, 2002).

Constructs and Frameworks

Civics Achievement

Measuring civic achievement is a controversial and fraught endeavor. Some argue that measures of civic achievement should focus on student understanding of basic facts about American government such as the names of the branches of the government (Farkas & Duffet, 2010; Pondisco, Sewall, & Stotsky, 2013). Others argue that civic assessments should measure

students' civic skills, such as their ability to deliberate on problems and take civic action on civic issues (McDevitt & Kiousis, 2006; Levinson, 2012). However, these types of tasks are often difficult to measure using standardized assessments.

This dissertation examined data from 2010 NAEP civics assessment. The NAEP civics assessment is a federally-funded study that assesses a representative sample of United States 4th, 8th, and 12th students on their civic skills and knowledge. The study also surveys teachers and principals about classroom instruction, school environment, and teacher professional development. The goal of the study is to "show how well American students are being prepared for citizenship in our constitutional democracy" (National Assessment Governing Board, 2010, p. vi).

The NAEP civics assessment measures civic knowledge and skills through multiple-choice and open-response questions. Some of the questions assess students' factual knowledge, such as whether they know the number of terms a President is allowed to serve (NAEP, 2011). However, many of the questions require students to analyze and interpret images and texts. For example, students might be asked to analyze the meaning of a political cartoon or to explain the importance of a Supreme Court decision (Levine, 2013; NAEP, 2011). Students are also asked to develop solutions to civic dilemmas such as overflowing garbage cans in a park to deciding what types of information are needed to justify U.S involvement in a conflict in a foreign country (NAEP, 2011). As a result, using students' scores on the NAEP civics assessment as an outcome measure allows this study to analyze not just students' factual knowledge of civics, but also their analytical skills as they relate to civic topics.

Nonetheless, NAEP civics is limited in the types of civic skills and knowledge that it can assess. For example, NAEP civics cannot provide any information about students' ability to

work in groups or to make effective oral presentations. Additionally, the civics knowledge and skills tested on NAEP civics tend to reflect the values of the test developers who are predominantly White, college-educated, and economically privileged (Levinson, 2010). These limitations may limit the types of inferences that can be made about differences in student achievement on the NAEP civics assessment.

Despite these limitations, NAEP civics data offers opportunities and advantages not available in other data sources. First, NAEP collects data on both civics teacher characteristics and on student achievement in civics and the data is set up to link teachers to particular students. This allows comparisons between teacher characteristics and student achievement on NAEP. Second, the NAEP sampling procedure produces a student sample that is nationally representative of the 8th grade students in public and private schools. NAEP's sampling procedure allows valid inferences about the civic knowledge and skills of all 8th grade students attending schools in the United States. Third, NAEP civics is the largest publically available dataset in K-12 civics education with nearly 10,000 students in the sample (National Center for Education Statistics, 2010). NAEP civics has enough statistical power to estimate small to medium sized effects of teacher and classroom factors. Finally, NAEP civics is a professional designed instrument that has been reviewed by experts for construct validity and has demonstrated strong psychometric properties (NAEP, 2011). The characteristics lend it strong credibility among policymakers and practitioners as a measure of civic achievement.

Interactive Civics

Civic researchers and educators have increasingly advocated for using student-centered activities such as simulations, service-learning, participatory action research, and discussions of controversial issues to teach civics (Billing, 2000; Hess, 2009; Levinson, 2012; McDevitt &

Kiousis, 2006). Yet scholars have used different terms to describe these types of civic activities. For example, Levinson (2012) describes the ideal form of participatory civic activities as "guided experiential civic education." Examples of these types of activities include mock trials, community organizing, and serving on student government. Levinson excludes service-learning from this definition because she argues that it promotes apolitical forms of civic participation that emphasize individual contributions rather than communal action.

McDevitt and Kiousis (2006), advocate for what they term "deliberative learning" where students engage in the social construction of knowledge through interpersonal exchanges of diverse views. They argue that "deliberative learning" must be student-centered, interactive, and receptive to diverse views and opinions. The authors do not explicitly exclude other forms of civic education from their definition. It can be inferred, however, that classroom activities like dramatizations or service projects that do not include interpersonal exchanges of views among students would not qualify under their definition of deliberative learning.

This study will use the term "interactive civics" to describe classroom activities where students discuss issues with one another and engage in hands-on activities. Interactive civics includes discussion-based activities such as deliberation and debates on controversial issues. It also encompasses experiential activities such as service-learning or contacting a political representative. Although other studies have used the term "interactive" to describe these types of civic activities (Engelhardt & Steinbrink, 2001; Jackson, Hinde, & Hass, 2008) it has not been explicitly defined in the literature.

Interactive civics activities were defined based on three characteristics. First, interactive civics activities must be student-centered, meaning that students, not teachers, are responsible for engaging in their learning on their own and with their peers (Pederson & Liu, 2003). In student-

centered approaches, instructional time is devoted to independent and group work rather than teacher lectures (Pederson & Liu, 2003). Second, students need to be engaged in some form of interpersonal communication either written or spoken. Many learning theorists have argued that the process of learning is a social activity, where students learn how to participate in the discourse and practices of a particular community (Lave & Wegner, 1991; R.D. Putnam & Borko, 2000). These theories suggest that in order for students to learn how to be effective citizens they need to learn how to engage in conversations around civic topics (Homana, 2009). Third, interactive civics activities should engage students in authentic forms of civic participation that expose students to the types of civic activities that they might engage in as adults (Finlay, Wray-Lake, & Flangan, 2010; Levine & Higgins-D'Alessando, 2010). Civically engaged adults participate in service activities in their community, discuss issues with others, and join community organizations (Carnegie Corportation of New York and CIRCLE, 2003; R.D. Putnam, 2001). Advocates for interactive forms of civic education argue that exposing students to these activities in classrooms will help students develop the desire and skills to continue engaging in these civics activities as adults (Levine & Higgins-D'Alessando, 2010).

Teaching using interactive civics activities may foster civic engagement in a number of ways. Some have found that interactive civics activities more engaging for students than traditional forms of civics instruction, contextualizing civic knowledge and making it easier for students to absorb new knowledge (Hess, 2002; Kahne, Chi, & Middaugh, 2006). Furthermore, interactive activities provide students with the opportunity to develop civic skills, such as organizing peers and solving dynamic problems, that are more difficult to teach in other modes of instruction (Levinson, 2012). Finally, interactive civics activities change the orientation of civics instruction from the teacher or textbook to the student (E.W. Ross, 1997).

Although the goal of interactive civics is to develop students' citizenship skills, interactive civics is not limited to social studies classes focused specifically on civics and government (Levine, 2013). For example, students may learn about the structures of government by engaging in a mock debate between the Federalists and Anti-Federalist in a U.S. history class or about learn about civic activism by working on a group project about Gandhi's salt march in a world history class. As a result, this study will examine interactive civics within all middle school social studies classes, not only classes that are specifically focused on civics or government. This is consistent with other educational studies of students' civic knowledge and civic participation that include students who are not necessarily enrolled in a formal course in civics or government (A.Cohen & Chafee, 2012; Kahne & Sporte, 2008; Niemi & Junn, 1998).

Teacher Professional Learning

Although many studies have examined the relationship between interactive classroom activities and student civic outcomes, relatively few studies have examined how teacher characteristics might affect students' civic knowledge and skills. This lack of research on teacher impact comes at a time when there is increasing evidence about the roles teachers play in improving student academic achievement (Hanushek, 2011; Kane & Steiger, 2008; Wright, Horn, & Sanders, 1997). Studies on teacher effectiveness have found that some teachers are effective at improving student academic outcomes consistently across a school year whereas as others teachers are not as effective (Hanushek, 2011).

These differences in teacher effectiveness can be explained in part by the types of instructional practices that teachers use in the classroom (Spillane & Louis, 2002). For example, a study of 3rd grade elementary school teachers in Virginia found that teachers who were on the top-quartile in their value-added scores were more likely to use a broader range of instructional

strategies than teachers who scored in the bottom quartile (Stronge, Ward, Tucker, & Hindman, 2007). Effective teachers were also more likely to ask questions that addressed higher-order thinking skills than ineffective teachers (Stronge et al., 2007). Another study found that English Language Arts teachers with value-added scores in the top quartile were more likely to teach explicit strategies for writing, that could be applied across a range of tasks, than teachers in the second quartile (Grossman, et al., 2010).

The relationship between instructional practices and teacher effectiveness has led many educational leaders and policymakers to call for more high-quality professional development opportunities for teachers (Duncan, 2009; Weingarten, 2010). Proponents of teacher professional development argue that providing teachers with access to high-quality professional development will improve teachers' instructional practices and therefore make them more effective in promoting student learning (Borko, 2004; Darling-Hammond & McLaughlin, 1995; Desimone, 2009; H. Hill, 2009).

Traditionally, professional development was conducted through direct instruction in a specific content or pedagogical area (Desimone, 2009; J.W. Little, 1993; Wilson & Berne, 1999). For example, a school might bring in an outside provider to teach a workshop or send a group of teachers to a summer institute or conference (Desimone, 2009). Although this "traditional" model is still the most commonly used form of teacher professional development, it is often criticized by scholars and practitioners as being too sporadic and decontextualized to be useful to teachers (Grossman, Wineburg, & Woolworth, 2001; J.W. Little, 1993; Wilson & Berne, 1999).

Consequently, the traditional model of professional development is increasingly being challenged by "communities of practice" approaches toward professional development (Grossman, Wineburg, & Woolworth, 2001; Wenzlaff & Wieseman, 2004). Communities of

practice learning theorists argue that individuals learn new skills through interacting with more experienced participants within a particular context (Brown & Duguid, 1991; Lave & Wegner, 1991). Based on these theories, advocates of communities of practices" approaches to teacher learning claim that teacher learning is most sustainable when it is conducted an ongoing basis with other educators (Darling-Hammond & McLaughlin, 1995; Grossman, Wineburg, & Woolworth, 2001). Some examples of this approach to professional learning include "professional learning communities" (PLCs) where teachers engage in frequent conversations with colleagues about instructional strategies, assessments, and student data (DuFour, 2004; McLaughlin & Talbert, 2006); "teacher online networks" where teachers share resources and collaborate with other teachers outside of the school through online platforms (Berry, Norton, & Byrd, 2007; Lock, 2006); and "coaching" interventions where fellow teachers provide detailed feedback and mentoring about improving their instructional practices (Guiney, 2001).

However, few studies have examined whether traditional or communities of practice models of professional development are more strongly associated with student learning outcomes. Most of the literature on teacher professional development focuses on changes in teachers' knowledge and practices rather than on changes in students' academic achievement (Guskey, 2003; Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Additionally, studies that do examine the relationship with student achievement often focus on more traditional forms of professional development (Yoon et al., 2007).

This study used a framework of subject-matter professional learning that incorporates both traditional and communities of practice models. First, the study examined how overall engagement in teacher professional learning activities is related to teachers' instructional practices and students' achievement on the NAEP civics assessment. Then the study separately

examined communities of practice and traditional forms of professional development. This allowed for contrasts between the different models of professional development and their relationships to interactive civics practices and student achievement on the NAEP civics assessment.

Population and Sample

The target population for this study was 8th grade students who attended public and private schools within the United States during the 2009-2010 academic year (NAEP, 2011). NAEP uses a multi-stage stratified sampling procedure to generate a sample of schools and students that is nationally representative of American public and private schools (National Center for Education Statistics, 2010). In total, 9,630 8th grade students from 470 schools participated in the assessment.[3] The analysis was conducted using the restricted-use data set that contains student, teacher, and school level variables.

Research Questions

The dissertation addressed the following research questions:

1. **What are the patterns of social studies teachers' content-area professional learning activities?**
    a. What types of professional learning activities do teachers engage in?
    b. How do patterns of professional learning activities vary by teacher and school characteristics?
2. **How is teacher engagement in overall content-area professional learning activities related to teachers' classroom practices and student achievement on the NAEP civics assessment?**

[3] All sample sizes rounded to the nearest tenth to comply with the reporting requirements of the restricted use data license

   a. Are teachers who engage in more overall content-area professional learning activities more likely to use interactive civics activities in the classroom controlling for school and teacher characteristics?
   b. Do students of teachers who engage in more overall content-area professional learning activities have higher achievement on the NAEP civics assessment, controlling for school, teacher, and school characteristics?

3. **How does the form of content-area professional development (traditional or communities of practice) relate to teachers' classroom practices and student achievement on the NAEP civics assessment?**
   a. How does the form of content-area professional development relate to teachers' use of interactive civics activities in the classroom?
   b. How does the form of content-area professional development (traditional or communities of practice) relate to students' civic achievement on NAEP, controlling for school, teacher, and student characteristics?
4. **How are teachers' classroom practices related to student outcomes and what role does the amount and type of content-area professional development play in that relationship?**
   a. Do students of teachers who use interactive civics activities in the classroom, have higher levels of civic achievement on NAEP, controlling for school, teacher, and school characteristics?
   b. Does teachers' engagement in more overall content-area professional learning activities *moderate* the relationship between teachers' use of interactive civics

activities in the classroom and student achievement on the NAEP civics assessment, controlling for school, teacher, and student characteristics?

c. Does the form of content-area professional learning *moderate* the relationship between teachers' use of interactive civics activities in the classroom and student achievement on the NAEP civics assessment, controlling for school, teacher, and student characteristics?

Significance of the Study

Many are concerned about the low levels of civic engagement among American youth. Advocates have called for using school-based civic education as a vehicle for improving students' civic skills and knowledge and ultimately fostering increased civic participation. Yet, learning civics in school does not reliably translate into increased civic engagement (Manning & Edwards, 2013) . This study focused on an area of civics instruction that has been under-examined in the literature: the role of teachers. In particular, it examined how teacher professional learning and instructional practices are related to student achievement in civics. Understanding what makes teachers effective civics instructors will help policymakers and practitioners design curriculum and professional learning opportunities to better prepare the next generation of American citizens.

Chapter 2

The purpose of this dissertation is to examine the relationships between teacher professional learning, use of interactive civics activities, and student academic achievement on the NAEP civics assessment. This chapter will situate this question within the larger educational literature on teacher and student learning. First, it will explore the literature on teacher

professional learning, focusing on how the research has evolved over the last few decades and what current research says about what makes for effective professional development. Then, it will examine some of the existing research on the relationship between interactive civics activities and student civic outcomes and connect teacher professional learning back to civics instruction. The chapter will conclude by describing the conceptual framework used in this dissertation.

Teacher Instructional Practices and Teacher Effectiveness

Are Some Teachers More Effective Than Others?

There is a growing body of literature that argues that teachers can make a critical difference in student educational outcomes. Researchers have consistently found that more of the variation in student achievement is between teachers than between schools (Guskey, 2003; Spillane & Louis, 2002). One study estimated that, controlling for student background, between 52-72% of the variability in students' academic growth in mathematics could be attributed to differences between teachers (Rowan, Correnti, & Miller, 2002).

Differences in student achievement between teachers have been found in a variety of settings. In a meta-analysis of teacher quality studies, Hanushek and Rivkin (2010) found that the average teacher effect was 0.17 standard deviations for math and 0.13 for reading and that the range of estimates was fairly narrow across studies. The study authors explained that this meant that, on average, a student who had a teacher at the 75th percentile rather than the 25th would improve from the 50th to the 59th percentile in math achievement. These teacher effects are larger than the effect sizes for other educational interventions. For example, a one standard deviation change in teacher effectiveness has a larger effect size than reducing class sizes from 25 to 15 students (Nye, Konstantopoulos, & Hedges, 2004).

Compared to reading and math, teacher effectiveness in civics has not been extensively studied. This may be because there are far fewer assessments of students' civic knowledge and skills. Some evidence does suggest that some civic teachers may be more effective than others. For example, Wright, Horn, and Sanders (1997) found that the same differences in teacher effectiveness for reading or math were also found in social studies. Additionally, on the CIVED international assessment in civics, 30% of the variance in U.S student achievement was between teachers (Torney-Purta, Richardson, & Barber, 2005). These findings suggest that there are likely substantial differences between teachers in how effective they are in teaching civics.

However, not all of the variation among teachers represents differences in teacher quality. Teachers are not randomly assigned to classes and some of the differences between teachers may represent pre-existing differences between students (Braun, 2005; Desimone & Long, 2010; Kupermintz, 2003). In particular, students with higher levels of parental education and family income are more likely to be assigned to higher performing teachers (Desimone & Long, 2010; Kane, Taylor, Tyler, & Wooten, 2010).

Differences in student achievement between teachers may also reflect factors outside of a teacher's control such as home factors, the school culture, or the relationships between students (Baker, et al., 2010; Braun, 2005; Corcoran, 2010; Kupermintz, 2003). Teacher effects can also be confounded both horizontally, by other teachers who are in the same grade level, and vertically by the previous teacher experiences of the student (Baker, et al., 2010; Corcoran, 2010; H. Hill, Kapitula, & Umland, 2010; Rothstein, 2010). For example, if a student had a poor teacher in 5th grade, the experience may carry over into their academic achievement in the following year. These effects are not always accurately captured by the statistical models that

measure teacher effects (Rothstein, 2010). As a result, teacher effects may not represent a causal relationship between teacher quality and student achievement.

Yet there is evidence that at least some of the variation in student achievement between teachers reflects real differences in teacher quality. Teacher effects have been found in studies where students were randomly assigned to teachers (Kane & Steiger, 2008; Nye, Konstantopoulos, & Hedges, 2004). For example, a study where teachers were randomly assigned to classrooms found that differences between teachers within a school explained, on average, 12.9% of the variance in student math achievement (Nye, Konstantopoulos, & Hedges, 2004). By comparison, differences between schools only explained, on average, 6.1% of the variance in student math achievement (Nye, Konstantopoulos, & Hedges, 2004). In addition, some portion of a teacher's effectiveness persists over time (Goldhaber & Hansen, 2012) and across different school environments (Chetty, Friedman, & Rockoff, 2012). These findings suggest that at least some of the differences in student achievement between teachers represent real differences in teacher quality.

What Makes a Teacher Effective?

Teachers do not come into classroom as blank slates. The way they teach may be influenced by many factor including their past educational experiences, their teacher training, their philosophy of teaching and what has worked well for them in the past. Whether these characteristics make any difference in how effective teachers are in increasing student academic achievement is an area of robust debate within the literature. Some researchers have argued that differences between teachers cannot be explained by any observed characteristics such as their training or years of experience in classroom (Gordon, 2008; Hanushek & Rivkin, 2010). They contend that trying to identify the characteristics that make a teacher effective, and trying to

replicate those attributes, is likely to be unproductive (Hanushek, 2011). Instead, these researchers argue that value-added models should be used to identify individually effective and ineffective teachers in order to reward the former and dismiss the latter (Hanushek & Rivkin, 2010; Sanders & Horn, 1998; Staiger & Rockoff, 2010).

Other researchers have challenged the view that teacher characteristics are unrelated to student academic achievement. In particular, these researchers argue that there are certain instructional practices that are characteristic of effective teachers (Baker, et al., 2010; Ball & Cohen, 1999; Darling-Hammond, 1999; Spillane & Louis, 2002). This claim is generally supported by the research literature. One meta-analysis of instructional practices found that studies of different teaching practices had average effect sizes ranging from .59 to 1.16 standard deviations increases in student achievement (Marzano, 2000). Another study found that differences in teacher practices accounted for approximately one-third of the variation in student academic achievement between teachers in the sample (Kane, Taylor, Tyler, & Wooten, 2010). Moreover, many studies have found a positive relationship between certain instructional practices and student academic achievement (Cohen & Hill, 2000; Guarino, Hamilton, Lockwood, Rathbun, & Hausken, 2006; Stronge, Ward, Tucker, & Hindman, 2007; Wenglinsky, 2001).

Researchers have identified a number of instructional practices that are associated with student academic achievement. For example, some studies have found that teachers who address more complex thinking in their instruction contribute more to student academic growth than teachers who focus on basic skills (Kane et al., 2010; Sternberg, 2003). One study of third grade elementary school teachers found that teachers in top quartile of value-added scores were more likely to focus on meaning instead of memorization than teachers in the bottom quartile (Stronge

et al., 2007). Top quartile teachers were also more likely to ask questions that addressed higher-order thinking skills than bottom quartile teachers (Stronge et al., 2007). Another study found that teachers whom observer rated as strong in terms of their use of cognitively demanding activities had greater improvements in their students critical reasoning skills than lower-rated teachers (Grossman, Cohen, Ronfeldt, & Brown, 2014).

Modeling specific strategies is another instructional practice that is associated with greater student academic achievement (Grossman, et al., 2010; Marzano, 2000). For example, a teacher might model questions that students should ask as they are reading a text such as "What is the author trying to say?" in order to improve students' reading comprehension abilities (Beck & McKeown, 2002). One study found that English Language Arts teachers with value-added scores in the top quartile were more than three times as likely to teach with these strategies than teachers in the second quartile (Grossman, et al., 2010).

Classroom management skills are also associated with higher student academic achievement. Teachers who have stronger classroom management skills, as measured by independent observers, had larger student gains in math than teachers with weaker classroom management skills (Kane et al., 2010). Another study found that teachers in the bottom quartile of value-added scores had, on average, a disruptive incident in their classroom every 12 minutes compared to once every two hours for the teachers in the top quartile (Stronge et al., 2007).

Do Effective Teachers Use More Interactive Activities?

There is a consensus among many professional educators that students benefit from a more interactive pedagogical approach. Most teacher professional standards of practice, such as the National Council of Teachers of Mathematics (NCTM) and the National Council for the Social Studies (NCSS) explicitly encourage teachers to use interactive activities like group-work

and class discussion in order to promote student learning (NCTM, 2014; NCSS, 2002). Indeed, studies have found that individuals learn more effectively by actively engaging with problems than by passively acquiring information (Bransford, Brown, & Cocking, 2000). Additionally, many attribute the success of high achieving countries such as Finland, Singapore, and Japan to their emphasis on project-based learning, inquiry skills, and classroom discourse (Darling-Hammond, 2012; Lewis, Perry, Friedkin, & Roth, 2012).

The research evidence on the benefits of interactive activities for student learning is mixed. Although many studies have found a relationship between more interactive instruction and student achievement, the effect sizes are often very small. For example, Cohen and Hill (2000) conducted a study of the mathematics reforms in California in the early 1990s to determine whether teachers who used the classroom practices advocated by the reforms, such as classroom discussion and group work, had greater student achievement in math. The study found that reform-oriented practices were significantly related to student achievement but that the effect size was small: a one standard deviation increase in framework practices corresponded to a .12 standard deviation increase in student math scores. This is the equivalent of the difference between scoring in the 50^{th} percentile and 55^{th} percentile. Other observational studies have found similar effect sizes. For example, Le et al., (2009) found that a one standard deviation increase in a similar set of reform-oriented practices had effect sizes of less than .05 for mathematics and less than .10 for science. Similarly, Hamilton et al., (2003) found that reform-oriented instructional practices in math and science had effect sizes of between 0.028 and 0.054 standard deviations.

Given that interactive instructional practices are generally seen as good practice, why have effect sizes from observational studies been so small? One possibility is that the consensus

is wrong. Indeed, some researchers have argued that minimally guided activities such as group work and independent investigations put too much of the onus on learning on the student (Kirschner, Sweller, & Clark, 2006; R. Mayer, 2004). Without significant teacher support, students may not actually learn the intended content (R. Mayer, 2004) or may lack the infrastructure to integrate any the informaiton with their existing knowledge (Kirschner, Sweller, & Clark, 2006). Additionally, independent and group learning may take away time from direct instruction from the teacher, reducing the amount of content covered in class (Rowan et al., 2002).

Another possibility is that traditional standardized assessments, which are often multiple choice tests, are inadequate measures of student learning from interactive activities. Some have argued that multiple choice tests emphasize factual knowledge at the expense of higher-order skills such as critical thinking and problem-solving (D. Mayer, 1998; Hamilton & Martinex, 2007; Pederson & Liu, 2003) As a result, these tests may not adequately capture the skills and knowledge advanced by interactive activities. Studies that compare student performance on traditional asessements to more open-ended assessment have found that interactive instructional practices have a stronger effect size on the more open-ended assessment (Grossman et al., 2014; Hamilton, et al., 2003; Le et al., 2009).

An alternative possibility is that that the survey measures of interactive instruction are inaccruate representations of how teachers actually teach. Survey measures of interactive instruction often rely on self-reports that place the onus on teachers to rememember what activities they used and how often they used them. Teachers may not fully remember what activities they used, or may overreport use in order to make their responses socially desiriable (Howard, Schmek, & Bray, 1979).

Moreover, teachers may interpret survey items differently from how the developers intended them (Hamilton & Martinez, 2007). For example, Hiebert and Stiegler (2000) compared teachers' survey responses about instruction to a videotaped observations of those teachers teaching a class. The researchers found that although teachers reported using reform-oriented instructional practices, the videotaped lessons indicated that teachers did not fully change their pedagogical approach. The videos showed that although students were working in groups the teachers' goals remained narrowly focused on procedures rather than conceptual understanding. As a result, "these classroom changes became changes about form rather than substance" (p.7). Teacher reports of interactive instruction may reflect the "trappings" of student-centered pedagogy rather than the underlying pedagogical approach (Hamilton & Martinez, 2007). Teacher survey responses may therefore underestimate the effect sizes of interactive instructional approaches.

Summary. Some of the variation in student academic achievement may be explained by differences in teacher effectiveness and there is a contentious debate in the literature about whether these differences can be explained by observable characteristics. However, there is substantial empirical evidence that certain instructional practices may make some teachers more effective than others. In particular, teachers who use more complex thinking and modeling strategies and have stronger classroom management skills are more effective in promoting student academic achievement. Interactive instruction is associated with greater student learning but the effect sizes from observational studies have been small.

Teacher Professional Learning

What is Teacher Professional Learning?

Teaching is a learning profession. The act of teaching requires constant decisions, adjustments, and accommodations in order to meet the instructional needs of students (Ball & Cohen, 1999; Sanders & McCutcheon, 1986). Teacher professional learning therefore encompasses a wide range of activities that support teachers' professional growth (Desimone, 2009). Indeed, any activity that is intended to improve teachers' practices, from informal exchanges with another teacher during lunch to week-long summer institutes, may be considered a form of teacher professional learning (Borko, 2004; Desimone, 2009).

American teachers engage in a substantial amount of teacher professional learning. For example, the Organisation for the Economic Co-operation and Development (OECD) conducted a study in 2013 teacher professional learning at the lower secondary level using nationally representative samples in 34 countries **Invalid source specified.**. American teachers in that study reported higher levels of involvement in professional learning than their international peers. Compared to the international average, American teachers were more likely to participate in workshops; conferences; individual or collaborative research; mentoring, peer coaching, and observations, and teacher networks. Overall levels of participation were high: 84% of American teachers participated in courses or workshops, 49% attended educational conferences, and 48% participated in educator networks. All of this professional learning translates into big business, comprising an estimated 1% to 6% of all school district expenditures (H. Hill, 2009).

Formal professional development activities such as workshops or training courses have long been the dominant form of professional development for teachers (Ball & Cohen, 1999; Wilson & Wineburg, 1993). Critics of this approach argue that this method of professional

learning is often ineffective at improving overall teacher quality. Trainings and institutes are often conducted outside of school, often over school vacations, making it difficult for teachers to transfer new skills to their classroom practice (Grossman, Wineburg, & Woolworth, 2001). Information from these trainings is often presented in a standardized form that may not be applicable to the teachers' specific school context (J.W. Little., 1993; Wilson & Berne, 1999). Formal professional development is often sporadic, usually taking place a few times a year and often lasting only a few days at most (Grossman, Wineburg, & Woolworth, 2001; Wilson & Berne, 1999).

In the 1990s the professional development literature began to expand beyond discrete professional development activities to study how teachers engage in professional learning in their daily interactions with other teachers (Ball & Cohen, 1999; Lieberman & McLaughlin, 1992; Darling-Hammond & McLaughlin, 1995). This literature was heavily influenced by Lave and Wenger's (1991) theory of "communities of practice." A community of practice is a group of individuals aligned along a common set of interests and goals (Brown & Duguid, 1991; Lave & Wenger, 1991; Wenger, 1998).The insight made by Lave and Wenger and others is that communities of practice are also *learning* communities (Brown & Duguid, 1991; Wenger, 1998). Individuals within communities of practice developed a shared set of stories, experiences, and tools that guide them in their practice (Wenger, 1998). Furthermore, these communities of practice largely exist independent of the institutional structure of the organization (Brown & Duguid, 1991; Stamps, 1997). This allows them to be more flexible and adapt better to changing conditions (Stamps, 1997).

Inspired by the potential of communities of practice, many educational organizations began trying to create their own "communities of practice" for teachers. One popular approach to

creating communities of practice is the professional learning community model. In professional learning communities, groups of teachers work together to share resources, discuss curriculum, and plan assessments (Giles & Hargreaves, 2006; McLaughlin & Talbert, 2006). This allows teachers to collaborate more effectively with one another and learn from the collective experience of other teachers in the school (DuFour, 2004; Giles & Hargreaves, 2006; King, 2002). Research on teacher professional communities within schools has found that professional community is associated with improved teaching practice and greater student achievement (Kraft & Papay, 2014; Lee & Smith, 1996; Louis & Marks, 1998; Vescio, Ross, & Adams, 2008).

However, school-based professional communities are often limited by the skills and experience of the teachers within a particular school (O'Day, 2002) As a result, some researchers have proposed using teacher professional learning networks to supplement professional communities within schools (Lieberman, 2000). Teacher professional learning networks are loose, independently organized collections of individual educators who are interested in similar educational topics (Lieberman & McLaughlin, 1992). These networks provide teachers with the opportunity to learn from and collaborate with educators from other schools (Lieberman & McLaughlin, 1992; Darling-Hammond & McLaughlin, 1995). Advocates for teacher professional networks argue that networks provide more flexibility and opportunities for innovation and collaboration than professional communities within schools (Lieberman & McLaughlin, 1992; Lock, 2006).

The internet has blurred some of the lines between school-based teacher professional communities and teacher professional networks. Growing numbers of educators are using social media platforms like Twitter or Pinterest to connect with educators and form online communities around specific topics (Blitz, 2013; Cho, Ro, & Littenberg-Tobias, 2013; Davis, 2010; Forte,

Humphreys, & Park, 2012; Greenhow, Dexter, & Riedel, 2006). These networks can be accessed all over the world allowing teachers to expand their professional network beyond their school community and collaborate and share ideas with educators around the world (Berry, Norton, & Byrd, 2007; Krutka & Milton, 2013; Lock, 2006). Teachers can now communicate with other teachers in another state or country as easily as they can talk to the teacher in the classroom down the hallway. Although research on online teacher networks is still in its nascent stage, preliminary evidence suggests that they benefit teachers' instructional practices (Blitz, 2013; Carpenter & Krutka, 2014; Forte, Humphreys, & Park, 2012).

However, teacher communities of practice also face challenges. Schools cannot create communities of practice simply by declaring one into existence (Grossman, Wineburg, & Woolworth, 2001). Communities of practice require high levels of trust and commitment among teacher as well as engaged and facilitative leadership from school leaders in order to be successful (Bryk, Camburn, & Louis, 1999; Grossman, Wineburg, & Woolworth, 2001). Teacher communities of practice can also have negative effects. Teachers may feel social pressure to conform to their colleagues' ideas and expectations resulting in "groupthink" that impedes innovation and growth (Giles & Hargreaves, 2006; Guskey, 2003). Teachers may also be unwilling to disrupt the status quo by criticizing a colleague who they believe has been underperforming or to suggest a new instructional approach.

Additionally, the divide between traditional professional development and communities of practice, while still present, has shrunk dramatically. Many traditional professional development programs, such as workshops or summer institutes, incorporate active learning components such as coaching or discussion groups (Glazerman, et al., 2008; Garet, et al., 2011). Furthermore, some traditional professional development programs are now using social

networking platforms to develop online communities of practice (Arnold & Paulus, 2010; Russell, Carey, Kleiman, & Venable, 2009; Storandt, Dossin, & Lacher, 2012).

Communities of practice are also increasingly becoming part of the formal organization of schools, with many schools setting aside "professional learning community time" for teacher teams to work together (DuFour, 2004; McLaughlin & Talbert, 2006). Although this ensures that communities of practice are ingrained in school culture (DuFour, 2004) it also limits inhibits some of the flexibility that is present in more informal communities of practice (Koliba & Gajda, 2009). In this way, communities of practice in schools are increasingly beginning to resemble traditional professional development activities. As a result, in the near future the distinction between traditional professional development and communities of practice may no longer be as important.

Does Teacher Professional Learning Improve Instructional Practices and Student Achievement?

In order for teacher professional learning to be effective, it should improve both teaching and student learning (Desimone, 2009). Many studies have found a positive relationship between participating in professional learning activities, improved instructional practices, and increased student achievement (D. Cohen & Hill, 2000; Desimone et al., 2002; Desimone, Smith & Phillips, 2013; Supovitz & Turner, 2000; Wenglinsky, 2001). Yet much of the evidence on the effectiveness of teacher professional learning comes from observational studies. A 2007 meta-analysis of teacher professional development programs studies found that out of the 1,300 studies they examined only 9 used an experimental or strong quasi-experimental design (Yoon et al., 2007). In an observational study, unlike with an experimental design, there is no way to determine whether participating in professional learning *causes* changes in teaching practice or

student achievement (Borko, 2004; Yoon et al., 2007). For instance, teachers who engage in more professional learning activities may also be more likely to use effective teaching practices, teach higher achieving students, or simply be more effective teachers.

In their meta-analysis of experimental or quasi-experimental professional learning programs, Yoon et al. (2007) found that, on average, professional learning increased student achievement by about 21 percentile points. However, most of the studies in the meta-analysis were conducted with fewer than 20 teachers at only a handful of sites. This limits the generalizability of the findings of these studies (Borko, 2004; Wayne et al., 2008). In the years since the meta-analysis was published, there have been six large-scale multi-site experimental studies of teacher professional learning programs (Garet, et al., 2008; Garet, et al., 2011; Gersten, Dimino, Jayanthi, Kim, & Santoro, 2010; Glazerman, et al., 2008; Jacobs, Franke, Carpenter, & Levi, 2007; O'Dwyer, Dash, Kramer, Humez, & Russell, 2010). Five out of the six studies found significant differences in teacher instructional practices between the treatment and control groups. Yet, only one of the studies (Jacobs et al., 2007) found that professional development significantly improved students' academic achievement.

There are a number of different ways to interpret this result. One possibility is that professional development does not improve teacher effectiveness (Hanushek, 2011). According to this view, efforts to improve teaching quality through professional development are destined to fail because we do not yet know enough about what instructional practices are effective (Hanushek, 2011; Loveless, 2014). Consequently, professional development may be effective at changing teacher practices, but these changes will not result in improved student achievement.

Another possibility is that the design and measures used by these types of studies make it difficult for them to identify the effects of professional development. One common criticism of

these types of large-scale professional development studies is that they use standardized assessments and measures as outcome variables (Fishman, Marx, Best, & Tal, 2003; Wayne, Yoon, Zhu, Cronen, & Garet, 2008). Although using standardized measures increases the generalizability of research findings (Desimone, 2009), if the outcome variable is not closely aligned to the content of the professional learning program it may be difficult to observe the effects of the intervention (Wayne et al., 2008). Indeed, the only experimental study to show a significant effect on student achievement (Jacobs et al., 2007) was also the study that was most closely aligned to the focus of the intervention.

An additional problem with using experimental designs to study the effect of professional learning programs is that control teachers are generally free to engage in their own professional learning outside the context of the study (Wayne et al., 2008). For example, in Glazerman et al.'s (2008) study of a beginning teacher mentoring program, 83% of teachers in the control group reported that they also had a mentor. The existence of this "ambient professional development" may reduce the observed impact of the professional development program on teacher and student outcomes (Wayne et al., 2008, p.473).

What are the Characteristics of Effective Teacher Professional Learning?

There is general agreement in the literature about what characteristics make for effective professional development (Darling-Hammond et al., 2009; Desimone, 2009; J.W. Little 2006; Wayne et al, 2008; Wilson & Berne, 1999). Desimone (2009) summarizes these characteristics as duration, content focus, coherence, active learning, and collective participation. The following sections will examine the existing literature on each of these five characteristics of teaching professional learning programs.

Duration. It is commonly argued that professional development activities of limited duration, such as one day-workshops, are ineffective forms of professional development (Darling-Hammond et al., 2009; Desimone, 2009; Little 2006; Wayne et al, 2008; Wilson & Berne, 1999). This has led to calls for more intensive and on-going forms of teacher professional development (Darling-Hammond et al., 2009; Desimone 2009). However, research on the relationship between the duration of professional development and teacher and student outcomes has found mixed results. In their review of experimental and quasi-experimental studies of teacher professional learning, Yoon et al. (2007) found that only professional development programs that lasted longer than 14 hours had a positive and significant effect on achievement. Supovitz and Turner (2000) also found that science teachers who participated in less than 20 hours of professional development used significantly less investigative teaching practices and had a less investigative classroom culture than teachers who engaged in more than 40 hours of professional development. However, other studies have found no significant relationship between the time spent in professional development and student and teacher outcomes (Desimone et al., 2002; Penuel, Fishman, Yamaguchi, & Gallagher, 2007; Wenglinsky, 2001).

One of the limitations with arguments for more sustained forms of professional development is that they often ignore issues of context, intensity, and scale (Opfer & Pedder, 2011). Although it may seem intuitively true that more time spent in professional development is better than less time, this is not always the case. Short-term professional development programs can be highly effective for learning a small set of specific strategies that are directly applicable to teachers' classroom practices. For example, Fishman et al. (2003) describe an evaluation of a science teacher workshop that was specifically focused on teaching students to read watershed maps. In the workshop, teachers participated in the same activity that they would

later use with their students. Workshop leaders modeled effective teaching strategies. When the participating teachers taught the lesson, students showed significant improvements between the pre and post surveys. Moreover, the one-day workshop was more effective than a multi-day workshop offered the previous year that was more focused on improving teachers' knowledge. Other short-term programs that address a similarly small set of specific skills also have demonstrated positive results (Harris, et al., 2012; Walker, et al., 2011).

Content focus and coherence. Researchers frequently argue that effective professional development needs to be focused on practice and aligned with school goals (Ball & Cohen, 1999; J.W. Little, 2006). Professional development that is directly relevant to what teachers are doing in the classroom is more likely to be internalized than broad "laundry lists" of goals (J.W. Little, 2006, p. 3). Ball and Cohen (1999) compare the more general forms of teacher professional learning to "expecting someone to learn to swim on a sidewalk" (Ball & Cohen, 1999, p. 12).

Professional development programs that focus on teaching specific content have been found to have stronger outcomes (Garet et al., 2001; Desimone et al., 2002; Desimone, Smith, & Phillips, 2013; Huffman, Thomas, & Lawrenz, 2003) For example, Desimone, Smith, and Phillips (2012) found that teachers who participated in professional development with a mathematics focus were significantly more likely to focus on advanced topics in their classrooms. Another study discovered that math and science teachers who had participated in curriculum development and examining practice were more likely to use standards-based instructional practices than teachers who participated in more general forms of professional development (Huffman, Thomas, & Lawrenz, 2003). Alignment with school practices is also important. Participating in professional development programs that reflect school-wide practices is significantly related to improvements in teacher knowledge and practice (Garet et al., 2001).

However, all of the findings about the importance of content-focus and coherence are based on observational studies. As a result, we cannot exclude the possibility that teachers who seek out coherent content-based professional development are simply more effective teachers. There is not yet enough experimental studies of professional development programs to determine whether content-based professional development is more effective than more general forms of professional learning (Gersten, Taylor, Keys, Rolfhus, & Newman-Gonach, 2013; Yoon et al., 2007).

Active learning and collective participation. The final characteristics that researchers argue are necessary for effective professional development are active learning and collective participation (Darling-Hammond et al., 2009; Desimone, 2009; J.W. Little 2006). Active learning requires that teachers take an active part in the learning process through hands-on activities, group work, and discussions (Desimone, 2009; J.W. Little 2006). Collective participation implies that all teachers within a specific school, grade, or department are participating in the professional development program (Desimone, 2009; J.W. Little 2006). The assumption is that this will lead to greater collaboration and a stronger sense of professional community than individually targeted professional development opportunities (Grossman, Wineburg, & Woolworth, 2001).

Observational studies of professional learning programs have found that active learning and collective participation are both associated with more positive outcomes (Desimone et al., 2002; Garet et al., 2001; Penuel et al., 2007). Findings from experimental studies are more ambiguous. Gersten et al. (2010) found that participating in a teacher study group that emphasized active learning improved teachers' knowledge and instructional practices. However, the teacher study group did not have a significant effect on student achievement. In another

study, Garet et al. (2008) compared the relative effect of two professional learning interventions: a teacher summer institute and the same summer institute with an added 60 hours of in-school coaching. The researchers found that although both treatments improved teacher knowledge and instructional practices, there was no significant difference between the coaching and non-coaching interventions. Furthermore, neither treatment had any effect on student reading scores.

Summary. Teachers engage in many different formal and informal forms of professional learning. Researchers are increasingly calling for schools to harness communities of practice within schools to engage teachers in professional learning. These communities of practice forms of professional development are often described as more effective than traditional forms of professional learning such as workshops or conferences. As a result, many aspects of communities of practice, such as collaboration among teachers, have been integrated into traditional forms of professional development. However, the empirical evidence about what makes for effective teacher professional learning is ambiguous at best. Much of the teacher professional learning literature is based on observational studies that may be affected by selection bias. Characteristics of professional learning that seem to be effective in one context often fail to work in another (Opfer & Pedder, 2011). This suggests that what makes professional learning effective may be highly dependent on the context and the goals of the professional learning program.

Civic Education

What are the Civic Purposes of Education?

Preparing students for the rights and responsibilities of citizenship has long been a goal of the American education system. George Washington, in his farewell address as President, argued that the country should invest in education because “as the structure of a government gives voice

to public opinion, it is essential that the public opinion should be enlightened" (Carnegie Corportation of New York and CIRCLE, 2003, p. 11). Similarly Thomas Jefferson (1785) in his *Notes on the State of Virginia* called for a system of public schooling for the purpose of "rendering the people safe, as they are the ultimate guardians of their own liberty" (p. 148). Many state constitutions also specifically mention the importance of educating the citizenry for democracy (Tolo, The Civic Education of American Youth: From State Policies to School District Practices (Policy Research Project, Number 133), 1999). For example, in the Massachusetts state constitution it states that:

> Wisdom, and knowledge, as well as virtue, diffused generally among the body of the people, being necessary for the *preservation of their rights and liberties*; and as these depend on spreading the opportunities and advantages of education in the various parts of the country, and among the different orders of the people, it shall be the duty of legislatures and magistrates, in all future periods of this commonwealth, to cherish the interests of literature and the sciences [emphasis added] (Constitution of the Commonwealth of Massachusetts, 1780, Chapter V Section II).

Today, many educational leaders and policymakers cite preparing students for democratic citizenship as one of the most important purposes of the U.S. educational system. For example, Secretary of Education Arne Duncan (2012) argued in a speech that "Preparing all students for informed, engaged participation in civic and democratic life is not just essential—it is entirely consistent with the goals of increasing student achievement and closing achievement gaps" (para 7). This sentiment crosses partisan boundaries. Republican-appointed former Supreme Court Justice Sandra Day O'Connor has been a strong advocate for civic education and was one of the

major backers of a comprehensive civic education bill that was signed into law in Florida (CIRCLE Staff, 2014). O'Connor and co-author Lee Hamilton wrote in a 2008 op-ed that "A healthy democracy demands sustained citizen participation, and our schools must give students the knowledge and tools to participate" (para. 4).

Although most educators would likely agree that preparing students for citizenship is important, there are also profound disagreements among educators about how students should be educated for citizenship. Citizenship is not value-neutral. Defining what it means to be an educated citizen requires prioritizing certain values over others. Westheimer and Kahne (2004) argue that there are three different typologies of what it means to be a "good citizen": the personally responsible citizen, the participatory citizen, and the justice oriented citizen. Each of these citizenship typologies represents different sets of values about what should be prioritized in civic education. Educators who value personal responsibility aspire to teach students about the importance of following societal rules and treating others with respect and dignity. In contrast, those who value participatory citizenry seek to prepare students with the skills and knowledge to actively participate in collective efforts to solve community problems. Finally, those who value justice-oriented citizenship strive to develop students' understanding of the structural forces that perpetuate inequality and inspire in students a commitment to remedying social injustices.

The conflict between these three typologies of good citizenship often appears in debates about the civic goals of education. Those who see personal responsibility as the primary purpose of civic education tend to be more concerned about whether schools are imparting common civic values and appreciation for American democracy (Levinson, 2012). They are often troubled by reports that American students lack basic civic knowledge. For example, the conservative Pioneer Institute lamented that "the collective grasp of basic history and civics among American

students is alarmingly weak" and warned that "civic and historical illiteracy now presents a serious threat to our national survival" (Pondisco, Sewall, & Stotsky, 2013, pp. 1-2). Supporters of this view often contend that this deficit in knowledge stems from schools not spending enough time teaching civic content (Farkas & Duffet, 2010; Stauss, 2013).

In contrast, those who view participation as the primary civic purpose of education are more likely to argue that schools should promote civic participation. Those who value participatory citizenship often support "interactive" civic learning activities, such as service-learning or debates, which they argue develop students' civic participatory skills (Billig, Root, & Jesse, 2005; Quigley, 2007). For example, in Project Citizen, a curriculum for middle and high school students, students research a problem in their community; develop a public policy solution; and design and action plan to get their policy implemented by the government (Project Citizen, 2014).

Supporters of justice-oriented citizenship are also concerned about low levels of civic participation; however, they are more likely to view this problem within the context of broader structural inequality. These critics often find the traditional definitions of civic engagement problematic because they believe they marginalize experiences of low-income and minority youth (Levinson, 2012; Rubin, 2012). They prefer forms of civic education that speak to marginalized students' life experiences by raising students' critical consciousness and encouraging students to engage in social action through rallies or protests (Ginwright & Cammarota, 2002). For example, students in Sistas and Brothas United, a program for African-American and Latino youth in the Bronx, learn about educational inequity in the public school system and develop an agenda for taking their concerns to policymakers (Shiller, 2013). Unlike Project Citizen, where the focus of the organization is to "fix problems", students in Sistas and

Brothas United seek to combat structural forms racism and classism within society (Shiller, 2013).

The question is whether these perspectives are necessarily adversarial or are there ways that they can successfully complement one another? I think the answer to this question is a tentative yes. Calls for inculcating a common set of civic values provide a basis for establishing the shared norms and beliefs that are necessary for building a civic society. Participatory perspectives bring an emphasis on developing practical skills and knowledge that students' will need to be active citizens. Justice oriented perspectives push back against efforts to smooth over racial and class differences and force students to think about inequality in a structural way. All of these perspectives represent necessary components of a comprehensive citizenship education. That is not to say that these perspectives can or should be reconciled. Arguing over what it means to be a democratic citizen is in many ways the definition of what it means to be a democracy.

Does Learning Civics in School Improve Student Civic Outcomes?

Regardless of how one conceptualizes the civic purpose of education, there is a widespread consensus that schools should be preparing students for citizenship (Carnegie Corportation of New York and CIRCLE, 2003; Campaign for the Civic Mission of Schools, 2011; Torney-Purta, 2002). However, the research on education and civic outcomes is mixed. On one hand, individuals with more education are more likely to have higher degrees of civic participation, even when other socio-economic factors are taken into account (Hillygus, 2005; Plutzer, 2002; Verba, Schlozman, & Brady, Voice and Equality: Civic Voluntarism in American Politics, 1995). Millennials with college degrees are four times more likely to vote in local elections than those without high school diplomas and five times more likely to have contacted a

public official (CIRCLE, Harvard Institute of Politics, and Mobilize.org, 2012). Some scholars have argued that education causally increases civic participation by improving students' civic knowledge and skills (Dee, 2004; Sondheimer & Green, 2010). This claim is bolstered by research that has shown a positive relationship between civic knowledge and participation (Delli Carpini & Keeter, 1996; Popkin & Dimock, 1999).

Yet, learning civics in school has a tenuous relationship with civic participation. Using the National Educational Longitudinal Studies of 1988 and 2002, Bachner (2010) found that taking a full-year civics course in high school increased students' likelihood of voting in future elections. However, the effect size was relatively small; depending on the election year, taking a full year of civics increased a subject's probability of voting by 2.5-5.5%. Other studies have found a weaker relationship between learning civics in school and civic participation. For example, Kawashima-Ginsberg and Levine (2013) did not find a statistically significant relationship between learning about voting in high school and students' electoral engagement. A systematic review of studies on the relationship between civic education and civic participation generally found that learning civics in school did not increase civic participation (Manning & Edwards, 2013).

Additionally, efforts to make civics a mandatory course have not produced increases in civic engagement. The number of states requiring students to complete a civics course has more than doubled, rising from 21 in 1988 to 45 in 2014 (Bachner, 2010; CIRCLE 2014). More students are also taking civics courses, increasing from 62% in 1982 to 86% in 2009 (Bachner, 2010; U.S. Department of Education, 2009). However, youth voting participation has remained relatively flat since the 1980s, and other indicators of civic participation have declined (CIRCLE, Harvard Institute of Politics, and Mobilize.org, 2012). Additionally, the disparity in participation

between youth of different social classes has increased dramatically. For example, in 1982 there was a 10 point difference in voter participation between youth whose parents did not graduate high school and those whose parents had graduate degrees. By, 2006 that difference had increased to 23 points (Ingels, Glennie, Lauff, & Wirt, 2012).

There are a number of factors that might explain the modest relationship between learning civics in school and civic participation. Many civics classes emphasize factual knowledge rather than developing students' civic skills (Kahne, Chi, & Middaugh, 2006). For example, a study of six high school civics teachers found that teachers almost never mentioned non-voting civics activities such as boycotting, petitioning, joining advocacy groups, or demonstrating (Niemi & Niemi, 2007).

Emphasizing factual knowledge over participatory skills may mean that students are not developing skills that might lead to greater civic participation. A national study of high school civics teachers found that although large numbers of teachers were confident that graduating students would be able to identify protections guaranteed in the Bill of Rights (79%), far fewer teachers in the study felt that students were prepared to actively challenge the political status quo (37%) (Farkas & Duffet, 2010). This suggests that civics courses, as currently implemented, may not be adequately preparing students in the skills they need in order to be actively engaged citizens.

Another reason learning civics in school may not promote civic participation is that the civic content students learn in school is often disconnected from their life experiences. Civics courses rarely examine social problems or discuss ways for students to become involved in addressing those issues (Kahne, Chi, & Middaugh, 2006). This is particularly problematic for students of color, for whom there is often a disjuncture between the civic ideals conveyed in

schools and the everyday reality of their lives (Levinson, 2012; Rubin, 2012). For instance, Godsay, Kawashima-Ginsberg, Kiesa, and Levine (2012) found that many of the low-income youth who they interviewed for a study about civic engagement among that population had negative memories of learning civics in school. As one subject in the study recalled: "We have a lot of issues now, we have issues in 2010…we need to be discussing those issues, not what happened back when George Washington was President" (p. 36).

As a result, civic participation may be more strongly related to the quality of civic learning. For example, Kahne and Sporte (2008) found that students who experienced more high quality civic learning opportunities in school, such as experiencing a climate of open classroom discussion and working on community projects, were more likely to express interest in civic participation. Numerous intervention studies have found that students who are exposed to interactive civics activities have more positive civic outcomes (Kahne, Chi, & Middaugh, 2006; McDevitt & Kiousis, 2006; Syvertesen, et al., 2009). This suggests that interactive forms of civic education may help students develop the knowledge and skills they need to become actively engaged in citizens.

What are Different Interactive Approaches to Civic Education?

In the past thirty years, there have been a number of pedagogical movements in civic education to make civics more engaging and relevant to students' lives. Some of the movements, such as service-learning, are now established parts of the curriculum in many schools. Other efforts, such as youth participatory action research, are still in the process of being developed and refined. This section will review the research on different interactive forms of civics instruction.

Service-learning. Service-learning emerged in the 1980s as a way to include more experiential forms of instruction within the school curriculum (Billing, 2000; Furco, 1996). Service-learning is an instructional method where students participate in community service that is integrated into the academic content that students are learning in the classroom (Celio, Durlak, & Dymnicki, 2011; Furco, 1996). For example, students might volunteer at a local homeless shelter a few days a week, while in class students are learning political and social issues related to poverty (Yates & Youniss, 1998). Service-learning is not limited to social studies classes. In science class, student might collect soil samples to look for environmental toxins (Gardella, Mililo, Sinha, Oh, & Manns, 2009) or plant trees in order to prevent soil erosion (Boss, 1999). These activities both help students learn scientific content and engage them in helping solve problems within their community.

Proponents of service-learning argue that it develops students' civic skills, helps them learn more about their local community, and increases their awareness about political and social issues (Billig, Root, & Jesse, 2005; Eyler & Giles, 1999; Furco, 1996; Myers-Lipton, 1998). Studies of service-learning have found that it generally promotes positive civic outcomes. Students who volunteered during high school are much more likely to vote than students who did not volunteer (Callahan, Muller, & Schiller, 2010). Additionally, students who participate in service-learning have greater civic responsibility, increased civic engagement, and stronger problem-solving skills (Billing, 2000; Eyler, Giles, & Braxton, 1997; Moely & Ilustre, 2013). A meta-analysis of service-learning programs found that, on average, programs had effect sizes of between .28 and .43 standard deviations on students' civic attitudes, engagement, and knowledge (Celio, Durlak, & Dymnicki, 2011).

Service-learning also has many critics. Some have argued that service-learning projects often do not provide positive benefits to the community (Eby, 1998). Service projects often put great demands on the staff of the community organizations (Blouin & Perry, 2009; Eby, 1998). Additionally, service projects are often designed more to fit the needs of the student than to address community needs (Eby, 1998). This can lead to unintended consequences. For example, Biddle (2014) describes a school service-learning program in Africa where students spent the day helping to build an orphanage. The author recalls how at the end of the day, when students were asleep, local community members undid all of the students' shoddy work and rebuilt the structure.

Service-learning can also lead to negative student civic outcomes. Participating in service projects may divert students' focus from the political origins of many social problems and foster apathy toward political actions (Flanagan, 2004; Levinson, 2012). Others have argued that service reinforces the racial and class privilege of White middle and upper class students, especially when they are serving low-income communities (Flanagan, 2004; Hollis, 2004). Service-learning projects may reinforce these students' perceptions that they are "fixing" deficiencies within these communities (Eby, 1998; Hollis, 2004; Marullo & Edwards, 2000). For low-income students, service projects may feel like exploitation. Many of the activities that youth perform as part of a service, such as picking up garbage or packing boxes, are what the parents of low-income youth normally get paid to do as part of their low-wage jobs (Levinson, 2012). As one youth noted in Godsay et al.'s (2012) study of low-income youth about his service-learning experiences, "The janitors get paid for it! I just feel like they was using us. I mean what's picking up trash showing us, besides being clean?...It didn't teach us nothing" (p.38).

Civic simulations. Civic simulations are activities that emulate the civic experiences that students might engage in as adults through games, competitions, and dramatizations (Campaign for the Civic Mission of Schools, 2011). For example, the Discovering Justice program engages middle school students in examining and analyzing court cases with the assistance of attorney volunteer who serve as coaches. At the end of the semester, students act as lawyers in a mock trial tournament before a real judge in an actual courtroom (Discovering Justice, 2014). The theory behind these activities is that by providing students with first-hand exposure to civic activities, students will both improve their civic skills and be more likely to participate in civic activities later in life (McDevitt & Kiousis, 2006; Levinson, 2012).

Participation in civic simulations is positively associated with certain civic outcomes. Civic simulations are correlated with greater civic knowledge and skills; students who report participating in civic simulations in schools have higher scores on the NAEP civics assessment (Kawashima-Ginsberg, 2013; Niemi & Junn, 1998). Civic simulation interventions have also demonstrated positive outcomes. For example, a randomized control trial of iCivics, a computer game that teaches students to develop arguments on civic topics, found that the program significantly improved students' persuasive writing skills (Kawashima-Ginsberg, 2012). A randomized control trial of Student Voices, a mock election curriculum, found that the program positively affected students' voting confidence, attention to news, and political discussion in school (Syvertesen et al., 2009). Similarly, a quasi-experimental study of Kids Voting USA, another mock election curriculum, found similar positive findings in these areas (McDevitt & Kiousis, 2006).

However simulations are time-consuming and often require many additional hours of preparation and planning from the teacher (Levinson, 2012). It may be difficult for some teachers

to justify using instructional time to conduct simulations, especially when it may take time away from standards-based instruction. This tension may explain why civic simulations are so rarely used in schools. For example, less than half of 8th graders (48%) reported having ever participated in a civic simulation such as a mock trial (Kawashima-Ginsberg, 2013).

Another problem with civic simulations is that they can potentially hurt civic efficacy, especially for students from marginalized populations. Within a simulation, leadership roles may be assumed by students from privileged backgrounds who are more used to seeing their ideas and opinions be put into action. As a result, simulations may reinforce societal power dynamics where privileged students are permitted to make decisions and the ideas and opinions of marginalized are silenced (Bernstein, 2007). Additionally, simulations can sometimes perpetuate trauma for students from marginalized backgrounds. For example, African-American students may feel stigmatized by engaging in historical simulations that "relive" the traumas of slavery or segregation. Similarly, "poverty simulations" may be alienating to low-income students who may feel that their life experiences are being in turned into a game. Teachers using simulations must therefore be conscious of issues of power and representation when preparing in-class simulations.

Discussions and debates. Classroom discussions can take many forms ranging from discussing current events on a regular basis (Kawashima-Ginsberg, 2013) to debating controversial issues such as abortion and gun control (Hess, 2009). Advocates of classroom discussion argue that it promotes student dialogue, teaches students about diverse viewpoints, and forces them to consider alternative points of view (Rubin, 2012). Discussing political issues in civics classes also provides opportunities to expose students to views and positions that they might not be exposed to elsewhere (Hess, 2009). Research on political discussion has found that

when individuals with different viewpoints discuss issues, they each develop a more complex understanding of political issues than if they discussed the issue only with likeminded individuals (Mutz, 2002).

There is some evidence that classroom discussion is positively associated with student civic outcomes. For example, an analysis of CIVED (1999), an international assessment of civic learning, found that students, who felt that they could engage in open discussion with their teacher and peers, had greater levels of civic knowledge and civic engagement (Torney-Purta, 2002). Using the same data, Campbell (2008) found that classroom discussion climate moderated the negative effect of low socio-economic status on students' civic attitudes. Other studies have also found similar positive relationships between classroom discussion and student civic outcomes (Andolina, Jenkins, Zukin, & Keeter, 2003; Kawashima-Ginsberg, 2013; McDevitt & Kiousis, 2006; T. Zhang, Torney-Purta, & Barber, 2012).

Yet some critics have questioned the validity of measures of classroom discussion climate. Some researchers have argued that students and teachers tend to conflate any student talk with classroom discussion (Hess, 2009). For example, a validity study of the CIVED discussion measures found that students did not distinguish between in-class and out-of-class discussion with teachers when describing the state of their classroom discussion climate (Richardson, 2006). As a result, the observed effects of classroom discussion may reflect more the contribution of the overall school climate than the effects of the specific instructional practice.

Another limitation of classroom discussion is that political discussion may actual reduce students' interest in taking political action. For example, Mutz (2002) found that individuals who engaged in discussions with people they political disagreed with were less likely to participate in

politics. Mutz explained these findings by arguing that individuals who are exposed to alternative viewpoints may become more ambivalent about their own views and thus feel less confident in taking action. Additionally, people who have political diverse social and political networks may feel pressure to avoid taking actions that might be seen as controversial by those with whom they interact.

Mutz's findings may have some important implications for the relationships between discussion and political action in schools. In schools with politically diverse populations, teachers may be reluctant to raise controversial issues in the classroom out of concern for offending students or their parents. Moreover, if they do raise controversial issues they may risk reducing students' interest in taking political action. By contrast, in political homogenous schools, discussion may increase interest in civic action but may also result in greater political polarization because students are only exposed to one side of the issue.

Youth participatory action research. One still growing area of research in classroom-based civic practices is youth participatory action research. Participatory action research provides students with the opportunity work in groups to research and advocate on behalf of an issue that they care about (Foster-Fishman, Law, Lichy, & Aoun, 2010; Ozer, Ritterman, & Wanis, 2010; Rubin, 2012). As in other forms of interactive civics learning, participatory action research allows students to engage with real world problems. However, in youth participatory action research, students also take action to address the problem (Cammarota & Fine, 2008; Foster-Fishman et al, 2010). For example, students, at Central High School in Providence, explored the issue of gang violence through researching past interventions, meeting with representatives from law enforcements and non-profit organizations, and developed a documentary film to raise awareness about the issue (Pope, Stolte, & Cohen, 2011).

Youth participatory action research also differs from other forms of interactive civics learning in that it is explicitly guided by critical theory (Cammarota & Fine, 2008). Advocates for youth participatory action research argue that expertise lies within oppressed population and that only by galvanizing individuals within those oppressed communities can social change occur (Fine, 2009). Through the research process, students are encouraged through and dialogue and reflection "to develop their critical consciousness about their lives and broader community conditions" (Foster-Fishman et al., p.68). As a result, youth participatory action research programs tend to be most prevalent in school serving low-income, urban, students of color (Levinson, 2012).

Proponents of youth participatory action research argue that it develops students' civic skills by teaching them students how to advocate for an issue, and provide them with an opportunity to feel civically empowered (Cammarota & Fine, 2008; CIRCLE, 2013; Levinson, 2012). Research on youth participatory action research is still emerging, but case studies of existing programs suggest that it increases students' interest in civic participation and improves their research and communication skills (Pope, Stolte, & Cohen, 2011).

Nonetheless, youth participatory action research also faces challenges. One of the assumptions of youth participatory action research is that the research process is student-driven. But this desire can be undermined by the constraints of the existing school climate that may prioritize conformity over student initiative (Ozer et al., 2011). Students may also raise contentious issues and concerns that those in power may not be receptive to hearing (Levinson, 2012). This may limit the impact that students have in the short-term; structural problems like gang violence and school overcrowding are much difficult to address than cleaning up a park. As a result, students may become dispirited in the short-term. For example, Kahne and Westhimer

(2006) found that students taking part in a youth participatory action research program actually decreased in their belief that they can make a difference in society. The authors found similar sentiments in their interviews with students where many expressed frustration that they were "not taken seriously" (p.291).

Does Teacher Professional Development Improve Civics Instruction and Student Civic Outcomes?

Advocates for increasing teacher professional learning in civics argue that increasing teachers access to high quality professional development will improve the civics instruction that students receive in the classroom and thereby increase student civic outcomes (CIRCLE, 2013; Hess & Zola, 2012). There is some empirical evidence to support this claim. Researchers using the *CIVED* assessment found that students of teachers who had participated in a civics-related professional learning opportunity had higher levels of civic skills and knowledge (Torney-Purta, Richardson, & Barber, 2005). A national survey of civics teacher found that teachers who engaged in multi-day professional learning experiences were more likely to encourage political discussions among their students (CIRCLE, 2013). However, the study also found that teachers who participated in these professional development activities were also more likely to work in affluent communities and teach college-bound students.

The most extensive research on teacher professional development in civics has been conducted by Facing History and Ourselves, an organization that provides educational resources and training to teachers about the Holocaust and other examples of genocide and human rights abuses (Facing History and Ourselves, 2014). One of the goals of the organization is to teach students about the importance of civic participation (Barr, 2010). Abt Associates conducted a study using a randomized design to investigate the effects of Facing History's professional

development program on teacher and student outcomes (Boulay, et al., 2010). The study found that teachers in the treatment group were significantly more likely to feel that they had the knowledge to develop character in their students, create student-centered learning environments, and promote the development of historical understanding, tolerance, psycho-social development, deliberation skills, and civic learning. Students in the treatment group also had significantly higher levels of historical understanding, tolerance, and civic self-efficacy. But they did not differ from the control group in terms of their civic responsibility or civic engagement.

The positive outcomes from the Facing History study suggest that professional development in civics can improve civics instruction and student outcomes. However, Facing History also provides teachers with books, study guides, curriculum outlines, lesson plans, and other resources in addition to the professional development (Boulay, et al., 2010). As a result, it is difficult to disentangle the effect of professional development from the other resources that the program provides. Treatment teachers also taught different content than control teachers so some of the effects of the program may be attributable to teachers' teaching students about the Holocaust and genocide (Boulay, et al., 2010). These limitations notwithstanding, the positive outcomes from the program suggest that civics-content related professional development can produce positive changes in teacher instruction and student civic outcomes.

Summary. Citizenship has long been a part of the mission of public schools. However, civics instruction alone does not ensure that students will become active citizens. Rather, the quality of civics instruction is also important. Research on interactive civics activities has found that they generally improve students' civic outcomes. However, few experimental studies have been conducted on these types of activities and most studies have focused on specific programs. Additionally, many interactive civics learning programs are time-intensive and require

substantial resources. This may make it challenging for even the most dedicated teacher to implement them effectively in their classroom. Research on teacher professional development in civics is still emerging, but the existing evidence suggests that teacher professional learning might improve civics instruction and thereby increase students' civic learning.

Conceptual Framework

The preceding sections have presented evidence that some of the variation in student academic achievement may be explained by differences between teachers. They have also argued that professional learning is related to instructional practices and that more interactive civics instruction is associated with increased student civic outcomes. Desimone (2009) presents one conceptual framework for understanding how all these different components might relate to one another. The framework proposes a "core theory of action" (p. 184) for how professional development might influence student learning outcomes. The "theory of action" is as follows:

1. Teachers experience effective professional development
2. The professional development *increases* teachers' knowledge and skills and/or changes their attitudes and beliefs
3. Teacher use their knowledge and skills, attitudes and beliefs to *improve* the content of their instruction or their approach to pedagogy, or both
4. The instructional changes *foster* increased student learning [emphasis added] (p. 184)

Desimone's conceptual model is not intended as a series of causal links but rather as a framework for researching and evaluating teacher professional development. The connections between each of these elements are presented as non-recursive and interactive. Depending on the context some of the elements may have different degrees of emphasis or order. For example, improvements in student achievement may also be related to teacher attitudes and beliefs.

Nonetheless, all of these elements are necessary in order to "understand how professional development works to influence teacher and student outcomes" (p. 184).

In this study, Desimone's framework was modified to reflect the variables measured within the NAEP civics study. Teachers in the NAEP study were not surveyed about their attitudes and beliefs or their knowledge and skills related to teaching civics. Instead, they were only asked about types of subject-matter related professional development activities that they participated in over the last two years. The teacher survey did not provide any information about the quality of the professional development activity. Nevertheless, as described in Chapter 1, we can distinguish between "traditional" and "communities of practice" forms of professional learning. Therefore, based on Desimone's original, the conceptual framework of this study will be as follows:

1. Teachers participate in professional learning (traditional or communities of practice)
2. Teachers who engage in more professional development are more likely to use interactive civics activities in their classes
3. Students who participate in interactive civics activities in their classrooms are more likely to have higher achievement on the NAEP civics assessment (Figure 1)

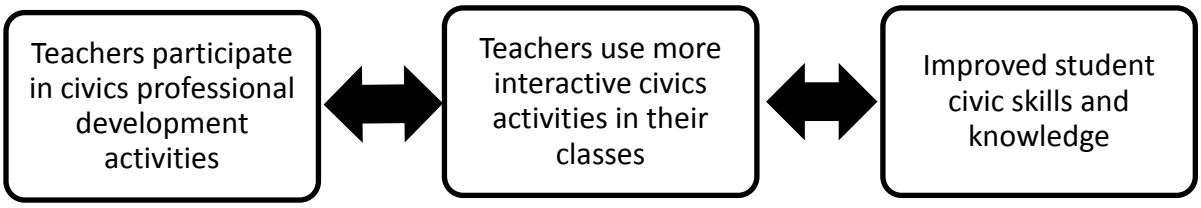

Figure 1. Conceptual Framework of the Study. Adapted from "Improving Impact Studies of Teacher' Professional Development: Toward Better Conceptualizations and Measures" by L. Desimone, Educational Researcher, 38,3, p.185

This conceptual framework poses certain hypotheses about the relationship between teachers activity and student achievement. It hypothesizes that there are differences between teachers in student achievement and that teacher instructional practices explain some of these differences. It also hypothesizes that teacher professional learning is positively related to the use of interactive activities. Finally, it hypothesizes that interactive civics activities are associated with greater civic skills and knowledge as measured by NAEP civics.

The research questions of this dissertation flow naturally from the hypotheses of this conceptual framework. The first research question examines the variation in teachers' content-area professional learning. The second and third research questions explore how professional learning is related to teachers' instructional practices and student learning. The fourth research question examines the relationship between instructional practices and student learning and how professional learning might moderate this relationship. The methods used to address these research questions is described in greater detail in the next chapter.

Chapter 3

The purpose of this dissertation was to describe the relationships between teacher professional learning, interactive civics instruction, and students' civic skills and knowledge as measured by the NAEP civics assessment. Specifically, it addressed the following research questions.

1. **What are the patterns of social studies teachers' content-area professional learning activities?**
    a. What types of professional learning activities do teachers engage in?
    b. How do these patterns of professional learning activities vary by teacher and school characteristics?
2. **How is teacher engagement in overall content-area professional learning activities related to teachers' classroom practices and student achievement on the NAEP civics assessment?**
    a. Are teachers who engage in more overall content-area professional learning activities more likely to use interactive civics activities in the classroom, controlling for school and teacher characteristics?
    b. Do students of teachers who engage in more overall content-area professional learning activities have higher achievement on the NAEP civics assessment, controlling for school, teacher, and school characteristics?
3. **How does the form of content-area professional development (traditional or communities of practice) relate to teachers' classroom practices and student achievement on the NAEP civics assessment?**

   a. How does the form of content-area professional development relate to teachers' use of interactive civics activities in the classroom?
   b. How does the form of content-area professional development (traditional or communities of practice) relate to students' civic achievement on NAEP, controlling for school, teacher, and student characteristics?

4. **How are teachers' classroom practices related to student outcomes and what role does the amount and type of content-area professional development play in that relationship?**
   a. Do students of teachers who use interactive civics activities in the classroom have higher levels of civic achievement on NAEP, controlling for school, teacher, and school characteristics?
   b. Does teachers' engagement in more overall content-area professional learning activities *moderate* the relationship between teachers' use of interactive civics activities in the classroom and student achievement on the NAEP civics assessment, controlling for school, teacher, and student characteristics?
   c. Does the form of content-area professional learning *moderate* the relationship between teachers' use of interactive civics activities in the classroom and student achievement on the NAEP civics assessment, controlling for school, teacher, and student characteristics?

This chapter describes the methods that were used to answer these research questions. First it will describe the data source including the population of interest and the sampling procedures used to generate the analysis sample. Then it will list the measures and variables used in this

study. It will conclude by presenting the analysis procedures, including a description of the statistical models.

Data Sources

This dissertation is a secondary data analysis of the 2010 NAEP civics assessment of 8th graders restricted-use data file. NAEP is a national assessment that assesses American students in a variety of subjects ranging from math to visual arts in 4th, 8th, and 12th grades. Instead of administering the exam to the entire U.S. student population, NAEP selects a random sample of students that is representative of the entire population of American public and private school students (NAEP, 2011). NAEP describes itself as the country's "National Report Card" because it is intended to be used as measure of the overall academic achievement of American elementary and secondary students at each of those grade levels (National Center for Education Statistics, 2010). NAEP also provides information about the academic achievement of different demographic subpopulations such as African-American or low-income students (National Center for Education Statistics, 2010). NAEP is not designed to provide information about a particular student's or school's performance; student and school identities are not disclosed to the public and NAEP results are not intended to be used for accountability purposes (Ho, 2012).

The NAEP civics assessment was first administered in 1969-70 as part of the initial administration of NAEP (Jones, 1996). The current version of the NAEP civics assessment frameworks was last revised for the 1998 administration of the test (National Assessment Governing Board, 2010). The NAEP civics assessment is designed to assess students' civic knowledge, for example, whether they know the purpose of each branch of government or can correctly identify the rights of a citizen (National Assessment Governing Board, 2010). It also is intended to assess students' civic intellectual and participatory skills (National Assessment

Governing Board, 2010). For example, students may be asked to take or defend a position on a public issue or develop a solution to a civic problem (National Assessment Governing Board, 2010).

NAEP civics is assessed in 4th, 8th, and 12th grades. However, only the 8th grade data was used to address the research questions. The 12th grade civics assessment was not compatible with the research questions because surveys were not collected from teachers. The 4th grade student data is linked to teacher surveys, but was not appropriate for this research question because of the age of the students. Social studies in elementary school is not a distinct subject and it is often deemphasized in favor of subjects that have high-stakes tests such as English Language Arts or Mathematics, (Burroughs, Groce, & Webekj, 2005; Fitchett & Heafner, 2010). For example, one national survey of elementary school teachers found that in the 3rd, 4th, and 5th grades, teachers spent an average of only two and a half hours of instructional time per week on social studies (Heafner & Fitchett, 2012). Consequently, it would have been difficult to estimate the relationship between instructional activities and students' civic knowledge and skills in elementary school.

There are also several advantages to examining 8th grade students in particular. Adolescence is an important period of political socialization, when students develop the knowledge, skills, and attitudes that shape their civic participation throughout their lives (Flanagan, 2004; Flanagan & Levine, 2010). Furthermore, many studies have found a relationship between civic experiences during adolescence and civic involvement later in life (Callahan, Muller, & Schiller, 2010; Hart, Donnelly, Younss, & Atkins, 2007). Understanding what factors are related to civic skills and knowledge at the beginning of adolescence may help

researchers and policymakers identify what school experiences might result in increased civic engagement through students' teenage years and into adulthood.

Population and Sample

Study Population

The population to which this study aims to generalize is 8th grade students who attended public and private schools in the United States during the 2009-2010 academic year (Rogers, Stoeckel, & Sikali, 2012). NAEP used a multi-stage stratified sampling procedure to select a sample of schools and students that is nationally representative of American public and private schools (National Center for Education Statistics, 2010).

Sampling Procedure

In the first stage of the sampling procedure, NAEP selected a sample of 50 to 100 geographic units known as primary sampling units (PSU) that represent one or more counties (NAEP, 2011). Within each PSU, NAEP compiled a comprehensive list of all the public and private schools in that area using information provided by the National Center for Education Statistics (NCES) (NAEP, 2011). Public schools were then grouped into strata based on shared characteristics such as minority student enrollment, achievement on state tests, and median income in the area of the school (NAEP, 2011). Private schools were stratified into groups based on degree of urbanization, minority enrollment, and type (e.g., Catholic, Lutheran). Schools were then sampled within each stratum. NAEP oversampled schools with certain racial/ethnic characteristics and undersamples private schools that have very low student enrollment (National Center for Education Statistics, 2010).

Sampled schools were then invited to participate in NAEP. The response rate was high: weighted by the number of students in the school, 97% of eligible public schools and 80% of

eligible private schools agreed to participate in NAEP civics (National Center for Education Statistics, 2010). Students were then randomly sampled from within the target grade in each school. The target sample size for students in each school is 110 students. Schools with more than 110 students selected a sample of 100 students and schools with fewer than 110 students administer NAEP to every student in the 8th grade (NAEP, 2009).

Data Collection

Data collection for the 2010 NAEP civics assessment was conducted between January and March of 2010 (Rogers, Stoeckel, & Sikali, 2012). In 2010, NAEP also assessed students in geography and U.S. history in addition to civics (Mullis, et al., 2012). Additionally, a small number of 8th grade students participated in a pilot study of an accessible booklet in mathematics (Mullis, et al., 2012). Students were randomly assigned to one of the tested subjects. As a result, only about 20 8th grade students within each school completed the civics assessment (National Center for Education Statistics, 2010). In total 9,630 8th grade students in 470 schools participated in the assessment. The final response rate for students in participating schools was 93% (National Center for Education Statistics, 2010).

Exclusions

Some students could not be assessed by NAEP because of limited English proficiency or a disability. NAEP allows assessment accommodations such as large print or one-on-one administration for students with a disability and bilingual dictionaries for English Language Learners. However, certain students cannot be assessed even with accommodations. For a student with disabilities, the student's Individual Education Plan (IEP) team was authorized to decide whether or not the student can participate in NAEP (Rogers, Stoeckel, & Sikali, 2012). Additionally, a student with an IEP was excluded from NAEP if their cognitive functioning was

severely impaired or if their IEP required an assessment accommodation that NAEP did not permit (Rogers, Stoeckel, & Sikali, 2012). English Language Learners were excluded from participating in NAEP if the student received instruction primarily in English for less than three years and was unable to take the assessment even with accommodations (Rogers, Stoeckel, & Sikali, 2012) . In total, 2% ($n = 160$) of students were excluded from the sample by NAEP because of a disability or because of limited English proficiency.

Additionally, some students and schools included in the NAEP sample were excluded from the analyses in this dissertation. In particular, 9% of students ($n = 910$)in the NAEP sample could be not matched with a teacher and 4% of schools ($k = 20$) did not have any survey data from teachers. Since the purpose of this study is to study the relationship between teacher characteristics and students' civic skills and knowledge, students who were not linked to a specific teacher were excluded from the analysis sample.

Measures

NAEP 8th Grade Civics Assessment

One of the stated goals in the frameworks of 2010 NAEP civics assessment was to "show how well American students are being prepared for citizenship in our constitutional democracy" (National Assessment Governing Board, 2010, p. vi). The frameworks divide citizenship into two components: civic knowledge and civic skills. Civic knowledge was described as representing "enduring questions" about American citizenship (p.15). The frameworks posed five questions that it claims form the core of civic knowledge:

1. What are civic life, politics, and government?
2. What are the foundations of the American political system?

3. How does the government established by the Constitution embody the purposes, values, and principles of American democracy?
4. What is the relationship of the United States to other nations and to world affairs
5. What are the roles of citizens in American democracy?

(p.x)

Civic skills were described as cognitive skills that are specifically related to the field of civics and government (p.22). These skills include identifying and describing concepts (p.23); analyzing events and arguments (p.24); evaluating, taking, and defining positions (p.25); and working effectively with others (p.27). Additionally, the NAEP civics frameworks defined the ability to deliberate about democratic issues as a civic skill: "To interact is to question, to answer, to deliberate with civility, as well as to build coalitions and to manage conflict in a fair, peaceful manner" (p.27). Students should also be able to "analyze the reasons or motivations for the use of emotional language" (p.25) and be able to identify and challenge "name calling, personal attacks, insinuation and innuendo" (p.26). The theory driving the civic skills component of the framework is that citizens should be able to deliberate with others in the public sphere (National Assessment Governing Board, 2010).

Instead of assessing civic knowledge and civic skills with separate items, the civic skills portion of the framework is integrated into questions about different areas of civic knowledge. This can best be demonstrated by examining publically-released items from the NAEP civics assessment[4]. For example, one item asked students to compare two diagrams representing two different democratic systems of government (Figure 2). This item required students to know the difference between parliamentary and presidential systems of democracy and have the

[4] Since there are no publically released from the 2010 8th grade assessment, examples will be used from publically released from the 2006 assessment. However, both assessments use the same frameworks so similar types of items should be found on the 2010 assessment.

information literacy skills to correctly interpret a diagram. It therefore simultaneously assessed students' civic knowledge and civic skills.

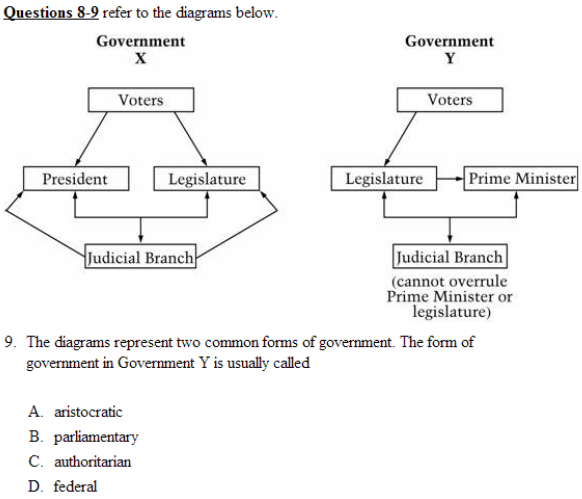

Figure 2. Example 8th Grade NAEP Civics Publically-Released Multiple Choice Item. Source: NAEP Questions Tool.

Another item presents students with the following scenario (Figure 3):

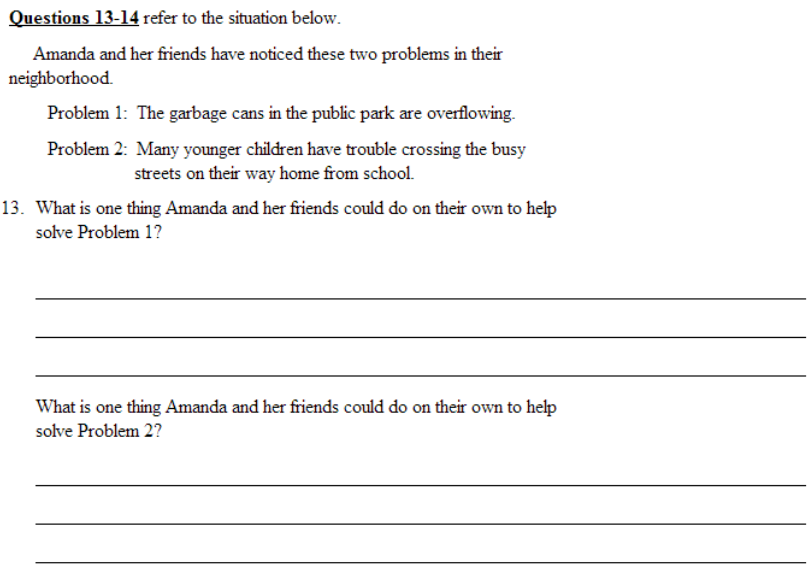

Figure 3. Example 8th Grade NAEP Civics Publically-Released Constructed Response Item. Source: NAEP Questions Tool.

This item assessed students' civic skills in being able to develop a solution to a problem and successfully defend their position. It also assessed students' civic knowledge about how society operates. For example, students would need to know what level of government (e.g., local, state, federal) was responsible for maintaining public parks in order to develop an effective solution to Problem 1.

The 2010 NAEP 8th grade civics assessment contained 166 multiple choice question and constructed response questions (Rogers, Stoeckel, & Sikali, 2012). Students were not given complete versions of the assessment. Instead, students were randomly assigned an assessment booklet that contained a portion of the items on the assessment (Rogers, Stoeckel, & Sikali, 2012). This allowed NAEP to assess more content areas without increasing the demand on students (Rogers, Stoeckel, & Sikali, 2012).

NAEP calculated student ability estimates using an item-response theory model (Rogers, Stoeckel, & Sikali, 2012). Item-response theory uses statistical modeling to estimate a value for student ability that takes into account the students' response pattern (Hambleton, Swaminathan, & Rogers, 1991). Student ability scores were calibrated so they are on a scale from 0-300 (Rogers, Stoeckel, & Sikali, 2012). Because students do not receive all of the items on the assessment, individual scores for each student cannot be calculated (Wu, 2005). To allow for secondary data analysis, NAEP computes five potential scores for each individual, known as plausible values. Plausible values are calculated using random draws from the posterior distribution of estimated abilities, given a particular response pattern (Rogers, Stoeckel, & Sikali, 2012; Wu, 2005). Combining these potential scores provides an estimate of a student's ability and a measure of the uncertainty associated with that estimate (Wu, 2005). All five plausible values were used in this study to estimate student achievement on the NAEP civics assessment.

Teacher Professional Learning

Teachers experience a wide range of interactions and experiences that are related to improving their instructional knowledge and skills (Borko, 2004; Desimone, 2009). These can range from informal conversations with another teacher in the hallway to attending a formal professional development workshop (Borko, 2004). The teachers' survey for NAEP civics measured discrete professional development opportunities that were specifically related to pedagogical content knowledge in social studies. Teachers were asked about their participation in the last two years in 12 different professional development activities that focused on teaching social studies and were only given the option of responding "Yes" or "No" to the questions. Additionally, teachers were asked about whether they currently held a leadership position in social studies, such as serving as a mentor teacher or department chair. Supporting beginning teachers and contributing to efforts to improve instruction in the school are often described in the literature as important components of teachers' professional development (Desimone, 2009; J.W. Little, 1993).

These 13 different forms of professional development were divided into two categories: traditional and communities of practice. Traditional professional development activities are experiences such as workshops or conferences, where the primary goal is to transfer information from the presenter to the participants (Desimone, Porter, Garet, Yoon, & Birman, 2002; J.W. Little, 1993; Wilson & Berne, 1999). In contrast, for communities of practice activities, knowledge is generated through interactions with other educators (R.T. Putnam & Borko, 2000). Some examples of communities of practice professional activities include coaching, team-teaching, and participation in discussion groups. These types of activities often involve more

teacher interaction and collaboration than traditional forms of teacher professional development (Desimone et al., 2002; DuFour, 2004; Penuel, Fishman, Yamaguchi, & Gallagher, 2007).

Professional learning experiences were assigned to categories based on previous studies that have compared traditional and communities of practice forms of professional development (Desimone et al., 2002; Penuel et al., 2007). For cases where the literature did not provide clear guidance, activities were assigned to categories based on the expected pedagogical approach of that activity. Professional activities that would likely use a social approach to educating teachers (R.T. Putnam & Borko, 2000) were classified as "communities of practice" whereas those that use a transfer model (Anderson, Reder, & Simon, 1996) were classified as "traditional."

Traditional professional learning consisted of 5 different types of activities: conferences, workshops, courses, consultation with experts, and independent reading. Communities of practice professional learning included 8 types of activities: observational visits, mentoring/coaching, committees, discussion groups, educator networks, collaborative research, co-teaching, and serving in a leadership position. Separate sum-total composite variables were created for both communities of practice and traditional professional learning activities. In addition, a composite variable was created representing the total number of professional learning activities. Since these three variables were completely collinear, they were examined separately in the analyses used to answer the research questions.

Interactive Civics

Many studies have observed a strong relationship between participating in interactive civics experiences and increased civic engagement (Billing, 2000; Hess, 2009; Levinson, 2012; McDevitt & Kiousis, 2006; Shiller, 2013). For the purposes of this study, interactive civics activities will be defined as instructional activities that are (1) student-centered, (2) require

students to interact with one another, and (3) engage students in authentic forms of civic participation. Some examples of interactive civics activities include discussing current events, debating controversial issues, working on group projects, and participating in civic simulations (Hess, 2009; Kahne & Sporte, 2008; Kawashima-Ginsberg, 2013; McDevitt & Kiousis, 2006).

Information on teachers' use of interactive civics activities was drawn from their survey response on the teacher background questionnaire. Teachers were asked about how often they used 11 different instructional practices and 8 different assessment strategies. Teachers responded on a scale of 1="Never or hardly ever", 2="Once or twice a month", 3="Once or twice a week", and 4="Almost every day."

Summing or averaging these values, as is often done in research on the relationship between teaching practices and student achievement (Connor, Son, Hindman, & Morrison, 2005; Lubienski, Lubienski, & Crane, 2008; Wenglinsky, 2001) posed two potential threats to validity. First, the response scale assigned a value to each activity based on the frequency it was used. However, some activities require greater levels of teacher and student preparation. For example, in-class debates generally require several days of student research and preparation so that students can familiarize themselves with the issues and practice their presentations. In contrast, a discussion of current events might occur informally without minimal preparation from the teacher and the student. Since these activities have different levels of intensity, assigning them the same value, based on how often they were used, would have posed a potential threat to the validity of the measure.

Second, summing or averaging scores across activities would have given the highest scores to teachers who used every activity on almost daily basis. However, this would not have reflected the consensus in the literature about what constitutes appropriate use of interactive

civics (Hess, 2002; McDevitt & Kiousis, 2006; Levinson, 2012). Although some activities, such as discussing current events or group work, could conceivably be integrated into every class, most interactive civics activities require substantial teacher and student preparation. If teachers conducted these activities almost every day, students and teachers might feel overwhelmed and the qualities of the activities might suffer. Additionally, interactive civics activities are generally described in the literature as a complement, rather than a complete replacement, for other forms of instruction (Levinson, 2012; Silva & Mason, 2003). Teachers who use multiple interactive civics activities every day may have little to cover content which may actually negatively influence student achievement. As a result, creating an index of interactive use by summing or averaging across activities may not accurately reflect the relationship between interactive civics use and student achievement on NAEP civics.

To address these issues, teachers' use of interactive civics was measured using an index that subjectively weighted each activity by intensity. The weight assigned to each activity was determined through a set of cognitive interviews were conducted with current and former middle school social studies teachers. The following section describes the cognitive interview procedures, the results of the interviews, and how the findings were incorporated into the weighting of the interactive civics index.

Cognitive interviews procedures. Cognitive interviews are a commonly used method in survey research for assessing the validity of survey instruments (D. Collins, 2003; Desimone & Le Floch, 2004). During cognitive interviews, subjects are given the survey instrument and asked to describe their internal thought process in answering the survey questions. The interviewer might then probe subjects about how they interpreted a question or why they decided on a

particular response category. In this study, the cognitive interviews were guided by two main research questions:

1. What types of instructional activities in social studies do teachers perceive as "interactive" according to the definition used by this study?
2. How would teachers "weight" the relative intensity of different types of interactive activities?

An initial set of interviews was conducted with a purposive sample of current and former middle school social studies teachers ($n = 10$). Teachers were recruited through professional networks, online listservs, and social media. Efforts were made to recruit teachers who work in both public and private schools, serve different types of student populations, teach in different parts of the country, and have different amounts of teaching experience. The recruited sample consisted of two private school teachers, three charter school teachers, and five public school teachers. Teachers with one to five years of teaching experience, teachers in urban areas, and teachers' of minority and low-income students were overrepresented in the cognitive interview sample compared with the NAEP civics teacher sample. Table 1 summarizes the characteristics of the teachers who participated in the cognitive interviews.

Table 1

Cognitive Interview Sample

Teacher ID	Years of Experience	School Type	School Sector	Urbanicity	School Size	School % Free or Reduced Priced Lunch	School % Minority
Teacher 1	1-5 years	Middle School	Public School	Small City	800-1000	12%	31%
Teacher 2	6-10 years	K-8 School	Public School	Large City	400-600	88%	97%
Teacher 3	10+ years	K-12 School	Private School	Large Suburb	600-800	N/A	1%
Teacher 4	1-5 years	Middle School	Charter School	Large City	200-400	71%	84%
Teacher 5	1-5 years	Middle School	Public School	Large City	200-400	37%	51%
Teacher 6	1-5 years	4th-8th	Private School	Midsize City	Less than 200	N/A	27%
Teacher 7	1-5 years	K-12 School	Charter School	Large City	800-1000	76%	100%
Teacher 8	10+ years	Middle School	Public School	Large Suburb	400-600	88%	95%
Teacher 9	10+ years	Middle School	Public School	Large Suburb	400-600	88%	95%
Teacher 10	1-5 years	K-12 School	Charter School	Midsize City	More than 1000	61%	87%

Note. Teacher 8 and Teacher 9 taught at the same school

During the interview, teachers were provided with the definition of interactive civics used in this study. They were then asked if they had clarifying questions about the definition. Subjects were also asked to provide examples of interactive civics activities from their own classroom experience to verify that they understood the definitions. They were then provided with a list of all of the instructional activities included on the NAEP teacher survey and asked to identify the activities that they believed matched the definition. Subjects were then asked follow-up questions designed to informally assess their perceptions about the relative impact of each of the interactive activities they selected. Then subjects were asked to rate each activity they selected on a scale of 0-3 where 0 indicated “No impact on student learning” and 3 indicated “High impact on student learning.” Subjects were first asked about the impact of activities when

conducted “once or twice a month” then “once or twice a week” and then “almost every day.” After each section, subjects were asked to explain the ratings they gave to each activity. The full interview protocol is available in Appendix A. The cognitive interviews were recorded and transcribed for analysis.

Cognitive interview findings. This section presents the findings from the cognitive interviews. First, it will describe the activities selected by subjects as examples of interactive civics activities. Then, it will present how subjects weighted the activities by the perceived impact on student learning. Finally, it will explain how these results were incorporated into the creation of an index of teacher use interactive civics activities.

Types of activities selected. Subjects generally agreed about what types of activities constituted interactive civics activities. Nine activities were selected by more than 70% of the sample as a form interactive civics (Table 2). The inter-rater agreement was fairly high with a kappa coefficient of $\kappa = .57$.

Table 2

Interactive Activities Selected

Activity	% Identified as Interactive Civics Activity
Debates or panel discussions	100%
Mock trials, role-playing, or dramatizations	100%
Community volunteer projects or services	90%
Group projects	90%
Write letters to state an opinion or solve a community problem	80%
Visitors from your community meet with the class to discuss important events and ideas	80%
Discuss current events	80%
Group presentations	80%
Use student government	70%
Individual projects	50%
Individual presentations	40%
Have students access information through the internet for use in the classroom	20%
Extended essays/papers on assigned topics	20%

Weighting of activities. Interview respondents were asked to weight activities differently depending on how they perceived that the activity would impact student learning. To illustrate these differences, the median weight given by subjects was calculated for each activity and frequency category (Table 3). Subjects assigned some activities large weights for conducting them "once or twice a month" because they viewed them as particularly beneficial for student learning. For example, the median rater weight for conducting debates and mock trials once or twice a month was a "3". As one teacher in the study noted, "any time you do debate, or role playing, or dramatization, those are the things that stand out in the mind. The kids get so much out of that!" However, other activities, such as using student government, only received a median weight of "1" for "once or twice" a month. Teachers felt that the activity would not have as much impact on student learning other activities on the list: "in general I think that student government doesn't involve everyone so I think it would be low impact."

Table 3

Median Weight by Activity and Frequency

Activity	Once or twice a month	Once a twice a week	Almost every day
Debates or panel discussions	3	2.5	2
Mock trials, role-playing, or dramatizations	3	2	1
Students write a letter to state an opinion or solve a community problem	2	2	1
Community volunteer projects or services	3	3	3
Discuss current events	1.5	2	2.5
Use student government	1	1.5	1
Group projects	2	1	1
Group presentations	2	1	1
Visitors from your community meet with the class to discuss important events and ideas	3	3	2

Note. Only includes activities selected by 70% or more respondents as an example of interactive civics

Another pattern that was consistent across interviews was that respondents felt that the impact of certain activities might decline if they were used too frequently. For example, the median weight given to mock trials, declined from "3" for once or twice a month, to "2" for once or twice a week, to "1" for almost every day. Respondents offered a number of different reasons why the impact of an activity might decrease in impact the more often teachers used it. Some felt that high intensity activities like mock trials and debates required too much preparation to be conducted effectively on a weekly or daily basis. One teacher noted in her class, "Any time [students] are presenting, talking, role-playing, doing any of these things, they have to write it first. They can't just have a few bullets and get up and be like 'here is my thing' in order for it to be good and for the whole class to get to get something out of it they have to write it first."

Subjects were also concerned that students would lose interest in an activity it was used too often: "you lose the novelty factor. You lose the panache of 'oh we have a big project.'". Another teacher observed that "I would say the more you can spice it up and maybe have something different every one to two weeks that would get them really interested." Finally, subjects felt that using the activities alone, without also building up content knowledge, might not benefit student learning. For example, discussing the impact of discussing current events one teacher commented that "I think for some kids who talk about things at home it can lead to a lot of learning but I think for a lot of kids, they don't have the background knowledge to easily access it."

Subjects' weights followed a fairly consistent pattern across activities. Subjects gave higher weights to activities conducted once or twice a month and gradually decreased the weight as the frequency of the activity intensified. The only exception to this pattern was "discussing current events" which raters assigned successively higher weights for more frequent use.

Constructing the index. The nine activities that were selected by at least two-thirds of responds as "interactive civics" were incorporated into the scale. The weights of these activities were determined based on the median weight given by subjects in the cognitive interview sample. The median was used instead of the mean in order to prevent outlier responses from skewing the weights.

However, the median weights were not directly used in the construction of the index. First, weights were trimmed so they were whole numbers. Additionally, weights were not allowed to decrease with increased frequency of use the activity. For example, the maximum weight of 3 was assigned to debates whether teachers did them once or twice, once or twice a week, or almost every day. This was done in order to make the index conceptually consistent; that a higher score on the index consistently indicated a greater use of interactive civics than a lower score. In order to account for the possibility that use of interactive civics, beyond a certain amount, may be detrimental to student learning a quadratic term was added to the analysis. Including a quadratic term in the models, allowed the relationship between the predictor and the outcome to have a non-linear slope. The final weights set of weights used to construct the interactive civics index are presented in Table 4.

Table 4

Weights Used to Construct Interactive Civics Index

Activity	Once or twice a month	Once a twice a week	Almost every day
Debates or panel discussions	3	3	3
Mock trials, role-playing, or dramatizations	3	3	3
Students write a letter to state an opinion or solve a community problem	2	2	2
Community volunteer projects or services	3	3	3
Discuss current events	1	2	3
Use student government	1	1	1
Group projects	2	2	2
Group presentations	2	2	2
Visitors from your community meet with the class to discuss important events and ideas	3	3	3

Validating the weighted interactive civics index. The weights used to construct the interactive civics index were then presented to a small validation sample of $n = 3$ middle school social studies teachers (Table 5). These teachers were asked to evaluate the selected activities and the relative intensity weights assigned to the response categories for each of the items. The validation sample largely concurred with the classifications of the original cognitive interview sample. Respondents, on average, agreed with 74% of the activities defined by the cognitive interview sample as interactive civics activities. Agreement with the weight depended on the frequency of the activity. Validation respondents agreed with, on average, 85% of the weights for "once or twice a month." Agreement with the higher frequency categories was much lower. Validation sample respondents agreed, on average, with only 52% of the weights for once or twice a week and 26% of the weights for almost every day. Like the original interview sample, teachers in the validation sample observed that conducting interactive civics activities too

frequently might negatively impact student learning. This potential non-linearity was addressed through including a quadratic term in the analysis.

Table 5

Validation Sample

Teacher ID	Years of Experience	School Type	School Sector	Urbanicity	School Size	School % Free or Reduced Priced Lunch	School % Minority
Teacher 1	1-5 years	K-8 School	Charter School	Large City	200-400	64%	88%
Teacher 2	1-5 years	7-8 (Junior High)	Public School	Large Suburb	800-1000	4%	36%
Teacher 3	1-5 years	6-9	Charter School	Large City	400-600	94%	100%

Student-Level Covariates

Student achievement is not only related to experiences inside the classroom, but it is also associated with non-academic factors as well. Many studies have found a strong association between student demographics, such as race or socio-economic status, and academic achievement (Coleman, 1968; Lee & Smith, 1997; Rumberger & Palardy, 2005). Additionally, student demographics are also related to opportunities to learn content and to be taught by high-quality teachers (Oakes, Ormseth, Bell, & Camp, 1990; Wang, 1998). As a result, the following student-level covariates were included in the analysis to control for student demographic factors that may be related the predictors and outcomes of interest.

Gender. Gender may be related to students' performance in a number of ways. Some studies have found that women tend to be less interested in politics than men (Delli Carpini & Keeter, 1996; Verba, Burns, & Schlozman, 1997). Interactive civics activities, such as debates trials or group projects may replicate the gender hierarchies the outside world with women feeling reluctant to speak out and take on leadership roles (Bernstein, 2007). Teachers may

exacerbate these hierarchies by rewarding boys for their assertiveness while critiquing girls for the same behaviors (Beane, Kawashima-Ginsberg, Kiesa, & St.Rose, 2014). As a result, gender may be an important factor in the relationship between interactive civics activities and civic achievement. Information in the NAEP database about gender was drawn from school records by NAEP and coded as a dichotomous variable (Rogers, Stoeckel, & Sikali, 2012).

Race/Ethnicity. Researchers have found that there are significant gaps between White and Black and Hispanic youth in civic knowledge and civic participation (A. Cohen & Chafee, 2012; Hart & Atkins, 2002; Kawashima-Ginsberg, 2013; Niemi & Junn, 1998). Black and Hispanic youth also are less likely to take classes specifically focused on civics topic and participate in out-of-school time activities that foster civic skills and knowledge (Kahne & Middaugh, 2008; Kawashima-Ginsberg, 2013). Information in the NAEP database about students' race was obtained from school records by NAEP and was coded as a categorical variable (Rogers, Stoeckel, & Sikali, 2012).

English Language Learner status. Students who are classified as English Language Learners may have lower scores on the NAEP civics assessment. Although not all English Language Learners are themselves immigrants, many are immigrants or are the children of immigrants (Garcia, Jensen, & Scribner, 2009). As a result, they may be less familiar with the institution and process of the American government because their parents are more likely to be immigrants and thus may be less familiar with U.S. politics (McDevitt & Chaffee, 1998; Niemi & Junn, 1998). Additionally, English Language Learners may struggle with the language complexity of the assessment, even with accommodations (Abedi, 2004). Information in the NAEP database about students' English language status was drawn from school records by NAEP and is coded as a dichotomous variable (Rogers, Stoeckel, & Sikali, 2012).

Disability status. Even with accommodations, students with disabilities often perform worse than students without disabilities on standardized assessments (Sireci, Scarpati, & Li, 2005). Students who receive special education services are required to have an Individualized Education Plan (IEP), which provides information about what services and accommodations the child needs (National Center for Learning Disabilities, 2014). As a result, IEP status serves as a proxy for which students are likely to have disabilities. Information in the NAEP database about students' disability status was obtained by NAEP from school records and was coded as a dichotomous variable (Rogers, Stoeckel, & Sikali, 2012).

Days absent from school. Student absenteeism is associated with higher rates of risk behavior and lower levels of academic achievement (Eaton, Brener, & Kann, 2008; Ginsburg & Chudowsky, 2012). As a result, students who have higher levels of absenteeism are more likely to have lower scores on NAEP. Students reported the number of days they were absent from school on the NAEP student background questionnaire. Days absent from school was measured as an ordinal variable where 0="No days absent", 1="1 or 2 days", 2="3 or 4 days", 4= "5 to 10 days", and 5="More than 10 days".

Socioeconomic status measures. Socioeconomic status reflects a students' ability to access cultural, social, and economic resources (Harwell & LeBeau, 2010). Socioeconomic status is associated with civic knowledge and interest in civic participation (Kahne & Sporte, 2008; National Center for Education Statistics, 2010; Niemi & Junn, 1998); students from lower socioeconomic backgrounds often have fewer opportunities to learn civics in school or participate in out-of-school time activities that promote civic participation (Flanagan, 2004; Kahne & Middaugh, 2008).

Single measures of socioeconomic status can have validity problems because they do not account for different sources of socioeconomic status (Harwell & LeBeau, 2010). To address this issue, socioeconomic status was measured with three different variables (Harwell & LeBeau, 2010). The first measure was a student's eligibility for a free or reduced priced lunch. Students are eligible for a free lunch if they have family incomes at or below 130% of the poverty level and a reduced priced lunch if their family's income is between 130% to 185% of the poverty level (United States Department of Agriculture Food and Nutritition Services, 2014). Information in the NAEP database about free or reduced priced lunch status was taken from school records by NAEP and was coded using three categories: eligible, not eligible, and information not available (Rogers, Stoeckel, & Sikali, 2012). In order to create a dichotomous indicator of free or reduced priced lunch status, students who did not have information about free lunch eligibility (because their school did not participate in the program or the school did not provide the information) were recoded as not eligible.

Parent education was also examined as a measure of socio-economic status. Many studies suggest that parents play an important role in young people's political socialization (Carnegie Corportation of New York and CIRCLE, 2003; Flanagan, 2004; Plutzer, 2002). Students whose parents have higher levels of education may have more exposure to political discussions and thus may be more interested in political topics (Niemi & Junn, 1998). Students were asked on the background questionnaire about their mother and father's highest level of education. Education was measured on a 1-4 scale where 1="Did not graduate from high school" 2= "Graduated from high school", 3= "Some education after high school", and 4="Graduated from college". A variable in the NAEP database, indicating the highest level of education achieved by either parent, was used to measure parental educational attainment.

The final indicator used to estimate socio-economic status was a home educational resource index. The number of educational resources available at home is positively related to students' civic knowledge (Homana, 2009; Niemi & Junn, 1998; Torney-Purta, Richardson, & Barber, 2005). This may be because students in homes with more educational resources have more opportunities to learn about political issues or because educational resources serve as a proxy for how engaged the students' parents are in following current events and political issues (Niemi & Junn, 1998; Plutzer, 2002). Home educational resources was measured using four dichotomous variables (access to newspaper, magazines, computer, and an encyclopedia) and one ordinal variable (number of books in the household) on the NAEP student background questionnaire. These variables were summed together to create a home educational resource index composite variable.

Teacher-Level Covariates

Teacher characteristics such as educational background and teaching experience are associated with instructional practices and student achievement (Darling-Hammond, 1999; Spillane & Louis, 2002). The following variables were included in the analysis to statistically control for other factors that may be related to the predictors and outcomes of interest.

Teacher race/ethnicity. Teachers' racial and ethnic background may be related to professional development opportunities, use of interactive civics activities, and student achievement on NAEP civics. White teachers are more likely to work in suburban schools which may have more resources to support interactive civics and teachers' professional development (Strizek, Pittsonberger, Riordan, Lyter, & Orlofsky, 2006). However, African-American teachers often report higher levels of professional community within their schools (Bryk, Camburn, & Louis, 1999) and may be more likely to participate in communities of practice. Information in

the NAEP database about teacher race/ethnicity was drawn from the NAEP teacher background survey and was coded using a categorical variable.

Years' experience teaching social studies. Teaching experience is positively associated with student academic achievement; however almost all of the benefits associated with experience are realized in the first few years of teaching (Hanushek, 2011; Rockoff, 2004). Years of experience may also be related to professional development opportunities. Newer teachers are more likely to receive mentoring (Glazerman, et al., 2008) and spend more time on professional development activities (Garet et al., 2001). However, experienced teachers may be more likely to attend professional conferences or develop professional networks around teaching. Teachers were asked about the numbers of years they had taught social studies on the NAEP teacher background survey.

Licensure status. Teachers with regular certification to teach in their subject-area are more likely to have students with higher levels of academic achievement (Darling-Hammond, Chung, & Frelow, 2002; Goldhaber & Brewer, 2000; Wayne & Youngs, 2003). Teachers with a temporary or emergency certification are also more likely to be teaching out-of-subject area, returning to teaching after a hiatus, or lack sufficient educational or subject-matter training (Darling-Hammond, Berry, & Thoreson, 2001). These teachers are also more likely to have students with lower levels of family income and parental education (Goldhaber & Brewer, 2000). Licensure status was measured using teacher self-reports on the teacher background survey.

National Board Certification. The national board certification is a voluntary board certification assessment for teachers administered by the National Board for Professional Teaching Standards (NBPTS). Teachers who receive National Board certification undergo an intensive assessment process that includes evaluation of student work samples, videotaped

lessons, and teachers' reflection on their teaching practices (Sato, Wei, & Darling-Hammond, 2009). Candidates must also demonstrate content and pedagogical content knowledge through performance assessment tasks such as evaluating curriculum materials and analyzing teaching scenarios (Sato, Wei, & Darling-Hammond, 2009). Students of National Board certified teachers generally have higher levels of academic achievement than non-board certified teachers (Goldhaber & Anthony, 2007; Strategic Data Project, 2012). As a result, students of National Board certified teachers may have higher achievement on the NAEP civics assessment. National Board certification was measured using a dichotomous variable on the teacher background survey.

Social science major. Studies of teachers' majors have found that students who are taught by a teacher who majored in the content area they teach have higher levels of student achievement (Boyd et al., 2009; Goldhaber & Brewer, 2000; Kukla-Acavedo, 2009; Monk, 1994). Students of teachers who majored in a social science may therefore have greater achievement in civics. Teachers were asked to indicate their undergraduate and, if applicable, graduate major on the teacher background survey. Social science majors included history, political science, geography, social science education, or any other social science field such as psychology, anthropology, or economics.

Subject-areas taught. Teachers who only teach social studies may be more likely to participate in professional learning on social studies content and develop interactive civics activities for their students. Additionally, teachers who only teach social studies may teach a systematically different population of students than teachers who also teach other topics. This is because teachers who teach multiple topics may be more likely to teach in smaller and under-resourced schools. Furthermore, teachers who also teach non-social studies topics may also be

more likely to have certification for teaching English Language Learners or students with disabilities, which may be related to student performance on NAEP civics. Teachers were asked about whether they taught only social studies on the teacher background survey.

Number of students linked to each teacher. The method NAEP used to sample and exclude student lead to disproportionate cluster sizes between teachers. For example, a teacher who teaches English Language Learners likely taught fewer students and was also more likely to have students excluded from the NAEP sample for lack of English proficiency. Cluster sizes may vary systematically by student characteristics which may affect both the types of instructional activities used in the classroom and student achievement on NAEP civics. Additionally, the reliability of a teacher's estimated mean student achievement decreases when there are fewer students linked to that teacher (Kane & Steiger, 2008; Chetty, Friedman, & Rockoff, 2012).[5] In order to control for possible bias introduced by the size of the cluster, the number of students attached to each teacher was included as a teacher-level covariate

School-Level Covariates

There is substantial empirical evidence in the literature that where a student attends school can have a profound influence on their academic achievement (Carbonaro & Covay, 2010; Greenwald, Hedges, & Laine, 1996). Students who attend schools with supportive academic cultures have higher academic achievement than their peers, even when controlling for students' socioeconomic background (Louis & Marks, 1998; MacNeil, Prater, & Busch, 2009). Additionally, schools have widely different access to resources. Schools in richer areas are generally able to invest more time and money into their students and teachers than schools in poorer areas (Greenwald, Hedges, & Laine, 1996). As a result, school-level covariates were

[5] See RQ2 for an explanation of how reliability is calculated with multilevel models

examined to account for variation in the outcome variables that may be attributable to school characteristics,

School context. School context was examined in order to account for contextual differences between schools that may be related to student achievement. Information about region, urbanicity, and school affiliation (e.g., public, private, or charter) were provided by NAEP. Additionally, the principal was asked about the percentage of students who are absent on an average day on the school background survey. The percentage of students absent on an average day was measured with an ordinal variable where 1="0-2%", 2="3-5%", 3="6-10%", and 4="More than 10".

Student demographics. Information about the percent of students in public schools who are eligible for free or reduced priced lunch and who are identified as students of color were obtained from the NCES Common Core of Data. This file was then merged with the NAEP civics school data file. A small number of public schools ($n < 10$) were missing information in the NCES Common Core of Data on the percent of students who were eligible for free or reduced priced lunch. In these cases, information about free or reduced priced lunch eligibility was obtained from state Department of Education websites. Additionally, information from the NCES Private School Universe database was used to determine the percent of students of color within private schools. Since the Private School Universe survey does not ask about free lunch eligibility, the percent of students eligible for free lunch within private schools was calculated using the NAEP civics school background questionnaire. On the questionnaire, this question was asked as a categorical variable with response options representing different percent ranges of students eligible for free or reduced priced lunch (e.g., 11-25%). In order to make the private

school data consistent with the public school information, the midpoint of the category was used to estimate private school free lunch eligibility.

Number of teachers in each school. The NAEP sampling procedures also lead to different numbers of teachers being sampled within each school. Smaller schools likely have fewer social studies teachers and this may affect teachers' opportunities to participate in professional learning within the school. Additionally, with smaller schools, the estimate of the school-level mean is likely to be less reliable. To adjust for potential bias, the number of sampled teachers within the school was included in the analysis as a school-level covariate.

Data

Data Structure

When observations are nested within units, such as classrooms or schools, traditional single-level statistical models, such as ordinary least squares regression, may not be appropriate (Cheong, Fotiu, & Raudenbusch, 2001). These types of models assume that observations are statistically independent (Raudenbush & Bryk, 2002; Thomas & Heck, 2001). However when observations are clustered within units this assumption is often violated. This is because observations within the same cluster are more likely to be similar to one another than observations from a different cluster (Thomas & Heck, 2001). When observations are not independent, using the usual method of calculating the standard error will result in standard errors that are systematically underestimated (Cheong, Fotiu, & Raudenbusch, 2001; Raudenbush & Bryk, 2002). Underestimated standard errors increase the likelihood of Type I error; incorrectly rejecting a true null hypothesis.

Another issue with using a standard ordinary least squares (OLS) regression model is that it can be difficult to model and interpret contextual effects.[6] A contextual effect is when a cluster-level characteristic, such as teacher background or school culture, is thought to be related to outcomes at the individual level (Bickel, 2007; Raudenbush & Bryk, 2002). In order to study contextual effects with a standard OLS regression model, observations either need to be *aggregated* to the cluster level or cluster variables need to be *disaggregated* to the individual level. Both of these options pose potential estimation and interpretation problems. When observations are aggregated to the cluster level the estimated parameters will likely be different from the individual level, unless the numbers of units in each cluster are the same, because smaller clusters will have the same "weight" as larger clusters (Raudenbush & Bryk, 2002). This can lead to estimated coefficients at the cluster-level that may be of a different magnitude, or even a different direction, than at the individual level (Gelman, Shor, Bafumu, & Park, 2007). Disaggregating cluster level information to the individual level violates the assumption that each observation is statistically independent leading to standard errors that are systematically too small (Raudenbush & Bryk, 2002).

To account for the nested nature of the data (i.e., students clustered within teachers within schools), the analyses conducted as part of this research study used hierarchical linear models (HLM). Hierarchical linear models are a special class of regression models that are designed to be used with clustered data (Bickel, 2007; Raudenbush & Bryk, 2002). These multilevel models allow for the calculation of the correct standard errors (Raudenbush & Bryk, 2002) . Additionally, HLM makes it possible to correctly model the relationship between cluster-level variable and individual outcomes (Raudenbush & Bryk, 2002).

[6] The term "contextual effects" in this study are used to describe a correlational relationship between two variables, where one variable is at a higher-level than the other. It should not be interpreted as indicating a causal relationship between two variables.

Cluster Sizes

NAEP does not sample intact classes when it samples students within schools. Rather, NAEP samples students within the school and then links their data with a teacher through school records (Rogers, Stoeckel, & Sikali, 2012). As a result, teachers who teach smaller classes were less likely to have their students sampled by NAEP. Furthermore, teachers who teach special populations (e.g., ELL students or students with disabilities) were less likely to be included in NAEP, because they generally had fewer students who are eligible for the study. This led to some teachers having very small clusters of students, as few as a single student, in the sample. Additionally, some smaller schools only had a single teacher in the sample.

One often cited guideline for multilevel modeling is that researchers should have at least 30 observations at each level of analysis (Bell, Morgan, Kromrey, & Ferron, 2010; Hox, 1998). However, there are many situations where it is difficult to control the number of observations per level-2 unit. For example, researchers studying neighborhood effects using large nationally representative samples often observe considerable sparseness in some clusters due to unequal distribution of individuals across neighborhoods (Bell, Morgan, Kromrey, & Ferron, 2010; Clarke & Wheaton, 2007). In these situations, researchers have little control over the number of units per cluster. As a result, a number of researchers have conducted Monte Carlo simulation studies to learn how multilevel models perform when cluster sizes are small. In these studies, researchers manipulate the number of clusters, the intra-class correlation, and the complexity of the model to determine how these factors might interact with cluster size (Bell et al., 2010; Clark & Wheaton, 2007; Maas & Hox, 2005; Theall, et al., 2011).

These studies have found that small cluster sizes can, in some cases, negatively affect various dimensions of model estimation and performance. Small cluster sizes can result in an

overestimate of the between-cluster variability and an underestimate of the within-cluster variability (Bell et al., 2010; Clarke & Wheaton, 2007; Maas & Hox, 2005; Theall, et al., 2011). These problems are exacerbated when group sizes are unbalanced, meaning that some clusters are much larger than other (Clarke & Wheaton, 2007). The problems are also more common in estimating random intercept-and-slopes models as opposed to random intercept-only models.

Underestimating the between-cluster variability can have a number of negative consequences for the model. When between-cluster variability is underestimated the standard errors for the level-2 fixed effects become too small (Bell et al., 2010; Maas & Hox, 2005). This leads to confidence intervals that are too small and an increased rate of Type I error for the level-2 fixed effect (Bell et al., 2010; Maas & Hox, 2005).Additionally, when cluster sizes are small the standard error of the between-cluster variability also increases (Bell et al., 2010; Theall, et al., 2011). Increased standard error leads to confidence intervals that are too wide and an increased rate of Type II error for the level-2 random effect (Bell et al., 2010; Theall, et al., 2011). This may lead some analysts to incorrectly conclude that clustering does not exist when it is present (Theall, et al., 2011). Finally, small cluster sizes reduce the statistical power of the model to estimate level-2 fixed effects (Bell, Morgan, Schoenberger, Kromkey, & Ferron, 2014). For example Bell et al., (2014) found that statistical models were generally underpowered $(1 - \beta < .8)$ with 30 level 2 units and clusters sizes between 5-10 observations.

However, most of these problems only occurred in extreme cases of data sparseness where the number of level-2 units was very small $(j < 50)$ and the proportion of single-unit clusters was greater than 50%. These problems can be almost entirely eliminated by increasing the number of level-2 units (Bell et al., 2010; Clarke & Wheaton, 2007; Maas & Hox, 2005). For example, when Bell et al., (2010) increased the number of level-2 units to 500, the standard error

of the level-2 fixed effects was no longer too small and the rate of Type I error fell to the nominal .05 level. Similarly, when the number of level-2 units increases, the upward bias in the between-cluster variability also disappears (Clarke & Wheaton, 2007). Clarke and Wheaton (2007) found that there was almost no bias when the number of level-2 units was at least 200 and there were less than 10% of clusters with only a single unit. Increasing the number of level-2 units also effectively increases the statistical power of the model (de Jong, Moerbeek, & Rien, 2010; Raudenbush, 1997).

In addition, the simulation studies generally found that small cluster sizes did not bias estimates of the level-1 and level-2 fixed effects (Bell et al., 2010; Clarke & Wheaton, 2007; Maas & Hox, 2005; Theall, et al., 2011). Convergence problems were not an issue in random-intercept models and only became a minor issue in random-intercept and slopes models (Clarke & Wheaton, 2007). As a result, many of the simulation studies concluded that small cluster sizes should not be an issue when the number of level-2 units is large and the proportion of very small clusters is small (Bell et al., 2010; Clarke & Wheaton, 2007; Theall, et al., 2011).

Based on the characteristics of the current sample and proposed model the best option was to retain the small clusters in the analysis. The NAEP civics assessment sample had large numbers of teachers ($j = 1120$) and schools ($k = 470$). In addition, the proportion of teachers with only one student in the sample was less than 10% and the number of schools with only one teacher was less than 20%. Furthermore, the goal of this analysis was to estimate fixed effects which are not sensitive to cluster size. Excluding small clusters from the analysis would have produced biased fixed and random effects because small clusters are likely systematically different from larger clusters on a number of dimensions. Removing smaller clusters would therefore have reduced some of the generalizability of the estimates from the model to the

American 8th grade public and private school student population. However, the presence of these small clusters may have slightly affected the estimates of the random components in the models.

Software

The study used HLM6 (Raudenbush, Bryk, & Congdon, 2004) to conduct the multilevel analyses. HLM6 is a statistical program that is specifically designed for conducting multilevel analysis. Additionally, the software is able to correctly apply sampling weights and handle plausible values (Raudenbush, Bryk, & Congdon, 2004). These characteristics make it an ideal program for the types of analyses conducted to address the research questions in this study.

Estimation Procedures

Full maximum likelihood estimation was used to estimate the parameters in the model (Raudenbush & Bryk, 2002). Maximum likelihood estimation identifies the parameters that maximize the likelihood function: the probability of observing a given set of data assuming a certain distribution and an unknown parameter produced the data (Greene, 2011). Maximum likelihood estimates have a number of desirable properties for large samples. Maximum likelihood estimates are consistent, meaning that there is a high probability they will be near the true parameter if the sample size is large enough (Greene, 2011; Raudenbush & Bryk, 2002). Additionally, maximum likelihood estimates are asymptotically efficient, meaning that if the sample size is large enough the maximum likelihood estimates will be approximately unbiased, with a minimal amount of variance (Greene, 2011; Raudenbush & Bryk, 2002).

Full maximum likelihood is the default estimation method for a three-level model in HLM6 (Raudenbush, Bryk, & Congdon, 2004, p. 66). HLM6 allows two-level models to be estimated using full maximum likelihood or restricted maximum likelihood (Raudenbush, Bryk, & Congdon, 2004). Restricted maximum likelihood is often used when the number of level-2

clusters is small, because it corrects the bias in the estimates of the variance components that can emerge in those types of situations (Raudenbush & Bryk, 2002). However, these estimates were unlikely to be biased in this analysis, because there are many level-2 clusters ($j = 1120$). Additionally, when the number of level-2 unit is large the differences between full maximum likelihood and restricted maximum likelihood are likely to be minimal (Raudenbush & Bryk, 2002). In order to be consistent across analyses, full maximum likelihood estimation was used to estimate parameters for all multilevel models in this study.

Data Preparation

A number of steps were taken to prepare the data for statistical analysis. These steps included addressing the missing data, recoding variables for analysis, ands applying sampling weights.

Addressing missing data. Missing data is a common problem in educational research. However, the best procedure for addressing missing data depends on the amount and pattern of missing data. If the data are missing completely at random (MCAR), meaning that the missingness is unrelated to any other variables, observed or unobserved, then observations with missing data can been removed through listwise deletion without biasing the estimates (Horton & Kleinman, 2007; R.J. Little & Rubin, 2002; Schafer, 1999). One method for testing this assumption is to use Little's (1988) MCAR statistical test. Little's MCAR tests the null hypothesis that the data are missing completely at random. If the null hypothesis is rejected it suggests that the pattern of missing data is not plausibly missing completely at random.

Even if the data are not MCAR, when the amount of missing data is less than 5%, it may be possible to exclude observations with missing data without introducing substantial bias (Schafer, 1999). However, since the proposed models had large number of predictors, even a

negligible amount of missing cases on each item may result in a substantial number of cases being removed from the analyses. Additionally, even a small amount of missing data can introduce substantial bias if the missing observations are particularly influential (Schafer & Graham, 2002). Listwise deletion was therefore not an appropriate method of addressing missing data in this study.

The data may be assumed to be missing at random (MAR) meaning that the pattern of missing data was not related to any *unobserved* characteristics, but may be related to observed characteristics (Horton & Kleinman, 2007; R.J. Little & Rubin, 2002). Alternatively, the data may be missing not at random (MNAR) meaning that the pattern of missing data is related to both observed and unobserved characteristics (Horton & Kleinman, 2007; Schafer & Graham, 2002). There is no formal test for whether the data are MNAR instead of MAR or MCAR. Instead, the plausibility of MNAR is assessed by examining the pattern of missing data (Horton & Kleinman, 2007; Schlomer, Bauman, & Card, 2010).

However, it is likely that even when the data are MNAR, at least some of the missingness may be accounted for by other observed characteristics (Schafer & Graham, 2002). When variables with missing data are correlated with observed variables then these observed variables may be able to account for the pattern in missing data, making the assumption of MAR plausible (Shin, 2013). As a result, imputation can be a useful tool for reducing bias when data are not strictly MAR (Schafer & Graham, 2002).

There are a number of different ways to approach imputing missing data. One option is to use conditional mean imputation, where a missing value is imputed based on the other observed variables in the data (Schafer & Graham, 2002). However, this option is problematic because it overstates the relationship between the observed and missing data patterns—the observed

variables are unlikely to perfectly predict the missing value (Schafer & Graham, 2002). As a result, the standard errors for the imputed values are artificially too small.

Alternatively, one can use a stochastic regression model to impute the missing value (Puma, Olsen, Bell, & Price, 2009). In a stochastic model, a residual error term representing a random draw from a standard normal distribution is included in the model (Schafer & Graham, 2002). This adjusts the standard error of the imputed value so it reflects the variance of the observed data (Puma et al., 2009). There are a number of different approaches to using stochastic model to impute missing data. One approach is to use a single imputation model. Single imputation models are fairly accurate amount of missing data is sparse (<5%) (Schafer, 1999). Another option is to use multiple imputation procedures which replace the missing values with several potential values that are predicted from different stochastic models (Gelman & Hill, 2007; Horton & Kleinman, 2007). Simulation studies have found that multiple imputation produces minimal bias for common amounts of missing data (e.g., 10-20%) when data are MAR (Schlomer, Bauman, & Card, 2010).

Strategies for imputing data for multi-level data are still being developed (Hox, 2013; Van Buuren, 2011). Some research suggests that imputation strategies that work well in conventional cases do not perform adequately when used with multilevel models. Gibson and Olejnik (2003) analyzed different types of missing data procedures using simulated data and found that multiple imputation performed particularly poorly for estimating level-2 fixed effects. Similarly, Cai (2008) found that multiple imputation produced more biased level-2 fixed effects than listwise deletion when used with multilevel data. Both studies found that maximum likelihood procedures using the E-M algorithm performed much better than multiple imputations when used with multilevel data (Cai, 2008; Gibson & Olejnik, 2003). In contrast, Puma et al.,

(2009) found that multiple imputation procedures produced much less bias than alternative methods when used with simulated data from a hypothetical educational randomized controlled trial.

There is also considerable debate about at what level the imputation should occur with multilevel data. Some authors have argued that you should not impute across multiple levels of data (Gibson & Olejnik, 2003; Van Buuren, 2011), However other researchers have found that imputing across multiple levels does not bias estimates, and that it is not necessary to take the hierarchical structure into account when the amount of missing data is less than 30% (D. Zhang, 2005). Some simulation studies have found that including multiple levels of data improved the estimation of fixed and random effects (L. Collins, Schafter, & Kam, 2001; Cai, 2008).

Using the scholarly literature as a guide, missing data in this study were addressed through the following procedure: First, the number of cases for each variable was calculated to determine the extent of missing data at each level of analysis. Second, Little's MCAR statistical test was performed to determine if the data were plausibly missing completely at random. Third, the correlations between the observed variables and the missing indicator were examined to decide whether the missing data were more plausibly MAR or MNAR.

Based on these three factors, single imputation was selected that best fit the amount and pattern of missing data in the data file. Single imputation provides reasonably accurate estimates of missing values when the extent of missing data is small (<5%) as it was for most of the variables in this study (see Missing Data section in Chapter 4) (Schafer, Multiple imputation: a primer, 1999). Additionally, a simulation study of a hypothetical educational study found that single stochastic imputation did not bias the estimates or standard errors at either the student or school level when 5% or less of the data were missing (Puma et al., 2009). Since the amount of

missing data was minimal in this study, single stochastic imputation was determined to be the most reasonable approach to imputing missing data in this study.

Missing data was imputed seperately at each level of analysis. School level predictors were included in the imputation model at both the teacher and student levels in order to improve the precision of the model (Cai, 2008). Separate imputation models were estimated for each item with the exception of items that were part of larger scales which were imputed using the same model.The type of imputation function was determined based on the charachteristics of the imputed variable. Scale variables were imputed using ordinary least squares regression and categorical variables were imputed using a logistic regression model. In a few cases, the logistic model produced unreasonable imputed values for categorical variables. In these situations an ordinary least squares regression model was used to calculate imputed values and the predicted value was rounded to the nearest whole number.

Recoding variables in SPSS. All analysis variables were recoded using SPSS syntax. Frequencies for these analysis variables were then computed. These frequencies were then thoroughly reviewed for potential coding errors.

Computing sampling weights. The complex sampling procedure that NAEP used to sample schools and students means that individual units have unequal probabilities of selection (Rogers, Stoeckel, & Sikali, 2012). This means that certain students and schools were more likely to be included in the sample than others. For example, students in schools with high concentrations of students from certain racial/ethnic groups were oversampled (Rogers, Stoeckel, & Sikali, 2012). Additionally, certain types of schools or students were more likely to participate in NAEP than others (Rogers, Stoeckel, & Sikali, 2012). These factors lead to certain types of schools or students being more likely to be included in the sample than others. When units have

unequal probabilities of being in the sample, unweighted estimates will likely be biased (Stapleton, 2009).

To address this issue, NAEP developed sampling weights that account for the unequal probability of selection for both schools and students (Rogers, Stoeckel, & Sikali, 2012). Sample weights are equal to the inverse probability of selection for the relevant unit (Rutkowski, Gonzalez, Joncas, & von Davier, 2010). For example, if a school has a probability of selection of .05, then the school's sample weight will be $(1/.05) = 20$. Since students are sampled within schools, student weights are equal to inverse joint probability that the school will be selected and that the student will be in the sample given that they are in a selected school ($P_k \cap P_{i|k}$) (Rogers, Stoeckel, & Sikali, 2012). This means if the school had a probability of selection of $P_k = .05$ and the student had a probability of $P_{i|k} = .40$ of being selected within their school, then the student's overall weight will be $[1/(.05 * .40)] = 50$.

One challenging aspect of weighting the proposed models is that NAEP does not include a weight for teachers. This is because the NAEP does not randomly sample teachers within schools. Instead, teachers were selected for the study if at least one of their students has been randomly selected (Rogers, Stoeckel, & Sikali, 2012). For schools with only one 8th grade social studies teacher this is not a problem; their probability of selection was the same as their school. ($P_j = P_k$). However, if there is more than one social studies teacher the probability of selection for a teacher is the joint probability that a teacher will be selected given that they their school was selected ($P_k \cap P_{j|k}$). The probability that a teacher was selected ($P_{j|k}$) is equal to $1 - \prod_{i=1}^{n_{jk}}(1 - P_{ijk})$ where P_{ijk} is the probability that a student in the teacher's class was sampled by NAEP and n_{jk} is the number of eligible students the teacher teaches. These weights cannot be calculated because there is no information in the data file about a) the probability of

selection for students who were not selected and b) the total number of eligible students that the teachers teaches. Without this information there is no way to determine the probability that a teacher would be included in the NAEP study. As a result, teacher weights could not be included in any of the models in this study.

The use of sample weights is further complicated by the fact that this study requires the use of multilevel models. There is considerable debate in the literature about how to correctly use sampling weights when using multilevel models (Braun, Jenkins, & Grigg, 2006; Pfeffermann, Skinner, Holmes, Goldstein, & Rabash, 1998; Rutkowski et al., 2010). Although many different weighting methods have been proposed, there is a lack of consensus in the literature about what is the best approach (Braun, Jenkins, & Grigg, 2006; Pfeffermann et al., 1998).

Additionally, raw sampling weights cannot be used directly with multilevel models (Carle, 2009). Raw sampling weights are calibrated so that the sum of all weights is equal to the number of units in the entire population (Pfeffermann et al., 1998; Thomas & Heck, 2001). This presents problems when calculating standard errors because the statistical software will assume that the effective sample size is equal to the entire population, deflating the estimates of the standard errors (Thomas & Heck, 2001). In order to address this issue, sample weights need to be "scaled" so that the sum of all weights is equal to the sample size rather than the population size (Carle, 2009; Thomas & Heck, 2001). HLM6 addresses this issue by scaling the sampling weights so the mean weight at each level of analysis is equal to 1 (Raudenbush, Bryk, & Congdon, 2004). This ensures that the sum of all sample weights is equal to the number of units at each level of analysis (Raudenbush, Bryk, & Congdon, 2004).

The original plan for weighting cases was to follow the recommendations of Rutkwoski et al., (2010) and apply separate sampling weights both the student and school levels. However,

when this was procedure was used to analyze the data in HLM6 the models failed to successfully converge. There are a number of reasons why this may have occurred. The recommendations for weighting multilevel models are based on two-level models (students within schools) rather than three-level models (students within teachers within schools) (Braun, Jenkins, & Grigg, 2006; Pfeffermann et al., 1998; Rutkowski et al., 2001). Three-level models may pose additional computational challenges to scaling weights that are not present in two-level models. Additionally, the absence of a teacher weight may have imposed even further complicated the estimation of the model.

Subsequently, the models were re-estimated using a student weight that was equal to inverse of the joint probability that the school and student were selected for the study ($P_k \cap P_{i|k}$). This weight encompassed both the probability that the student was selected and the probability their school was selected. It thus adjusted for the unequal probability of selection at the student and school level. This weighting scheme has been employed by other researchers who have analyzed three-level models with NAEP data and has been shown to be fairly robust (Cheong, Fotiu, & Raudenbusch, 2001).

For descriptive analysis, students and schools were weighted using their respective students and school level design weights. Descriptive statistics, such as percentages and standard errors, were calculated using the SPSS complex sampling procedure. The SPSS Complex Sampling accounts for the complex sampling procedure used to generate the NAEP sample (Rogers, Stoeckel, & Sikali, 2012), by correctly applying design weights and calculating standard errors using jackknife estimation.

Statistical Procedures

This section will describe the statistical procedures that were used to address each research question. Descriptions of the procedures will be organized by research question. Analysis procedures that were used in multiple research questions, such as multilevel models, will be described under the first research questions in which they were used. Subsequent research questions using the same procedures will refer back to previous research questions.

RQ1: What are the patterns of civics teachers' professional learning activities?

RQ1a. What types of professional learning activities do teachers engage in? The percentage of teachers who participated in each subject-matter related teacher professional learning activity was calculated for all 13 teacher professional learning variables. Additionally, a 95% confidence interval was calculated for each of the variables.

RQ1b. How do these patterns of professional learning activities vary by teacher and school characteristics? For nominal variables (e.g., teacher degree, school location) the numbers of teachers who participate in each type of professional learning were disaggregated by teacher and school characteristics and the marginal mean percentage was calculated for each cell. Chi-square tests were used to estimate whether the marginal mean percentage in each cell was significantly different than what would be expected if there was no relationship between participating in the professional learning activity and the teacher and school characteristics. For interval level variables (e.g., years of experience), the marginal mean value was calculated for each professional learning activity. Independent sample t-tests were then used to compare differences in means between teachers who did participated in the professional learning activity and those who did not.

When many independent statistical tests are performed on the same data the likelihood of finding a statistically significant relationship due to chance increases substantially (Ioannidis, 2005). For example, if one conducts 20 independent statistical tests on the same data the probability of finding a statistically significant result due to chance is 64% (Gelman, Hill, & Yajima, 2012). One solution to this problem is to lower the *p* value that indicates a statistically significant relationship. This study used a Benjamini-Hochberg correction to adjust for multiple comparisons (Benjamini & Hochberg, 1995) The Benjamini-Hochberg correction reduced the α level so the false discovery rate (FDR) on any set of analyses is no greater than .05, meaning that no more than 5% of all results will be statistically significant due to chance.

The Benjamini-Hochberg correction applies the following procedure to determine the correct level to set α so that the FDR is less than .05:

$$\text{let } k \text{ be the largest } i \text{ for which } P_{(i)} \leq \frac{i}{m} q^*;$$

$$\text{then reject all } H_{(i)} i = 1,2, \ldots, k$$

Where q^* is the desired FDR rate, *m* is the number of comparisons, and *i* is the number of *p* values that satisfies the inequality $i \leq q^* m^{-1}$ (Benjamini & Hochberg, 1995). In this analysis there were 351 comparisons and 24 *p*-values less than or equal to (.05)*(351)$^{-1}$. Solving for $P_{(i)}$

$$P_{(i)} \leq \frac{24}{351}(.05) \leq .00342.$$

Therefore, the α level for this analysis was set at .00342; meaning that all *p* values less than or equal to .00342 were rejected.[7]

[7] This is slightly different from the usual practice with null hypothesis significant testing of only rejecting hypotheses when the *p* values are less than the alpha level

RQ2. How is teacher engagement in overall professional learning activities related to teachers' classroom practices and student achievement on the NAEP civics assessment?

RQ2a. Are teachers who engage in more overall professional learning activities more likely to use interactive civics activities in the classroom, controlling for school and teacher characteristics?

Unconditional Model. The first model that was estimated (Model 1) was a two-level unconditional model with the interactive civics variable as the outcome variable. The unconditional model can be represented as

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$$

Where Y_{ij} is the interactive civics use of teacher *i* in school *j,* γ_{00} is the grand mean for interactive civics use, u_{0j} is the school-level random effect, and r_{ij} is the teacher-level random effect. Both u_{0j} and r_{ij} are assumed to be independently distributed, with u_{0j} having a distribution of $\sim N(0, \tau_{00})$ and r_{ij} having a distribution of $\sim N(0, \sigma^2)$ (Raudenbush & Bryk, 2002).

The unconditional model is used to calculate the intraclass correlation coefficient. The intraclass correlation coefficient indicates the proportion of variance in the outcome variables between level-2 groups (Raudenbush & Bryk, 2002). In this analysis, it represents the proportion of variance in teachers' interactive civics use between schools. The intraclass correlation coefficient is

$$\hat{\rho} = \frac{\hat{\tau}_{00}}{\hat{\tau}_{00} + \hat{\sigma}^2}$$

Where $\hat{\tau}_{00}$ is the estimated variance of the school random effect and $\hat{\sigma}^2$ is the estimated variance of the teacher random effect (Raudenbush & Bryk, 2002).

Another statistic that can be derived from the unconditional model is the reliability of each cluster's mean as an estimate of the population mean. The reliability of the intercept indicates to what extent the variance in the estimate of the sample mean in any school is a result of true variance in the population means (Raudenbush & Bryk, 2002). When the reliability is high it suggests that the observed group mean accounts for a large portion of the unknown variance of the true dependent variable mean (Raykov & Macroulides, 2006). In a two-level model, the reliability of the sample means for each cluster is

$$\hat{\lambda}_j = \frac{\hat{\tau}_{00}}{\hat{\tau}_{00} + \frac{\hat{\sigma}^2}{n_j}}$$

where n_j is the number of observations in each cluster *j* (Raudenbush & Bryk, 2002). The overall reliability of the sample mean can be calculated by averaging together the sample mean reliabilities of all the clusters in the sample (Raudenbush & Bryk, 2002). The reliability will be close to 1 when a) the cluster means vary substantially across level-2 units and b) when the cluster size for each level-2 unit is large (Raudenbush & Bryk, 2002). If the sample mean is not reliable the estimated sample mean for each cluster is "shrunk" toward the grand mean (Raudenbush & Bryk, 2002).

Conditional Models. Next, teacher-level covariates were added to the model (Model 2). These included teacher race/ethnicity, years of experience teaching social studies, licensure status, board certification, undergraduate and graduate major, subject areas taught, and the number of students in the sample taught by the teacher. In order to account for the non-linear relationship between years of teaching experience and teaching outcomes (Rockoff, 2004) a quadratic term was added for years of experience teaching social studies.

In this model and subsequent models, dichotomous variable were entered into the model as dummy variables. Additionally, categorical predictors were recoded as dummy variables and entered into the model with one category preserved as the reference group. Finally, all interval level variables were grand-mean centered, meaning the grand mean ($\bar{X}_{..}$) were subtracted from all X_{ij} observation (Raudenbush & Bryk, 2002). This was done to improve the interpretability of the intercept (β_{0j}). When variables are grand-mean centered, the intercept can be interpreted as the expected outcome for a subject whose value on the centered variable is equal to the grand mean (Raudenbush & Bryk, 2002).

All level-1 predictors were fixed so that their slopes were not allowed to vary randomly across level-2 units (Raudenbush & Bryk, 2002). The Level-1 model can be expressed as

$$Y_{ij} = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qi} + r_{ij}$$

Where β_{0j} is the school-level intercept and X_{qi} are the teacher covariates entered into the model. No predictors will be entered into the school-level intercept in Model 2 so β_{0j} will be

$$\beta_{0j} = \gamma_{00} + u_{0j}.$$

Non-significant predictors were retained in this and all subsequent models, because the covariates included in the models are all theoretically and substantively important for understanding the outcomes of interest. Omitting predictors from a statistical model may result in an underspecified model, which can bias the estimates of the fixed effects (Raudenbush & Bryk, 2002). Additionally, even if the omitted predictor were to have an average fixed effect of 0, removing it would have misspecified the model if the slope of the predictor varied significantly between level-2 groups (Raudenbush & Bryk, 2002). Although including non-significant predictors may result in overspecification, this was not a major concern for the analyses in this

study. Overspecification of the statistical model decreases the degrees of freedom and may inflate the standard error of the other predictors in the model (Wooldridge, 2009). However, as sample size increases the effect of including an unnecessary predictor on the standard errors of the other coefficients decreases to zero (Wooldridge, 2009). Given the large sample size of the data in this study, unnecessary predictors were unlikely to have a large impact on the statistical significance of the fixed effects. Additionally, unlike with an underspecified model, an overspecified model would not bias the estimates of the fixed effects (Wooldridge, 2009).

In Model 3, the slopes of the relationships between the teacher predictors and use interactive civics activities were allowed to vary randomly across schools by adding a random component, one at a time, to each level-1 slope. Random slopes were retained only if (a) the variance of the random slope was statistically significant, (b) the estimate of the school-level mean slope has an acceptable level of reliability ($\hat{\lambda} > .1$), and (c) the chi-square difference test for the deviance statistic indicates that the random coefficient significantly improves the fit of the model compared to the fixed coefficient (Li, Duncan, Harmer, Acock, & Stoolmiller, 1998).

Although it may be of research interest to use school-level characteristics to explain the variation in teacher slopes across schools, the structure of data made this unfeasible for this study. This was because the number of teachers in each school is relatively small and therefore adding predictors may have compromised the stability of the model. Additionally, when cluster sizes are very small, it is often not possible to add predictors to explain variation in the slopes (Raudenbush & Bryk, 2002). Finally, examining the school-level moderators of the relationship between teacher characteristics and interactive civics use was not a primary goal of this study, so including predictors to explain variation in the slope would not have substantively added to the analysis.

Next, school-level covariates were added to the model for the level-2 intercept in order to predict variation between schools in teachers' use of interactive civics activities (Model 4). The statistical model for the level-2 intercept was:

$$\beta_{0j} = \gamma_{00} + \sum_{q=1}^{Q} \gamma_{qj} W_{qj} + u_{0j}$$

Where W_{qj} are the school-level covariates that were entered into the model. School-level covariates included region, urbanicity, sector, the percentage of students absent on an average day, the percentage of students in the school eligible for free or reduced priced lunch, the percentage of minority students in the school, and the number of sampled teachers in each school. Since no predictors were added to the model for the level-2 slopes, their equations were:

$$\beta_{qj} = \gamma_{q0} + u_{qj} \text{ (for randomly varying slopes)}$$

$$\beta_{qj} = \gamma_{q0} \text{ (for non-randomly varying slopes)}$$

In the final model, Model 5, the total number of professional learning activities was added as a level-1 predictor:

$$Y_{ij} = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} + \sum_{p=q+1}^{P} \beta_{pj} (PDTOTAL)_{pij} + r_{ij}$$

Where β_{pj} is the fixed effect for the relationship between total number of professional learning activities and interactive civics activities controlling for teacher- and school-level covariates.

Effect Sizes. When sample sizes are large even very weak relationships will be statistically significant because the analyses have a lot of statistical power (J. Cohen., 1994). As a result, effect sizes were calculated in order to understand the magnitude of the relationship between the predictors and outcomes of interest (Shadish, Cook, & Campbell, 2002). Effect sizes were measured in terms of the standardized difference in the outcome variable of a one standard

deviation increase in the predictor of interest. In educational studies, experimental interventions generally have an average effect size of between .2 and .3 standard deviations, or an increase of 8 to 12 percentile points (C. Hill, Bloom, Black, & Lipsey, 2007). Policy interventions, where the relationships between the intervention and outcomes are more diffuse tend to have lower effect sizes, between .01 and .15 standard deviations (Yeh & Ritter, 2009) Effect sizes were also measured by comparing teachers in the top quartile (75th percentile) to teachers in the bottom quartile (25th percentile). This strategy is frequently employed in educational studies to contrast the difference in student experiences between having a high performing and low performing teachers (Hanushek & Rivkin, 2010; Stronge et al., 2007).

RQ2b. Do students of teachers who engage in more overall professional learning activities have higher achievement on the NAEP civics assessment, controlling for school, teacher, and school characteristics?

Unconditional Model. To analyze the relationship between teacher characteristics and student outcomes a three-level model were estimated with students at level-1, teachers at level-2, and schools at level-3. In a three-level model, the unconditional model (Model 1) is as follows:

$$Y_{ijk} = \gamma_{000} + u_{00k} + r_{0jk} + e_{ijk}$$

Where Y_{ijk} is the NAEP scaled score for student *i* with teacher *j* in school *k,* γ_{000} is the grand mean for student achievement on NAEP, u_{00k} is the school-level random effect, (Raudenbush & Bryk, 2002). As with the two-level models, all of the random effects were assumed to be independent and normally distributed with a mean of zero and a constant variance (Raudenbush & Bryk, 2002).

In a three-level model there are two intraclass correlation coefficients of interest: the variability in student achievement on NAEP civics among teachers within schools and the

variability in student achievement on NAEP civics between schools. For the variability among teachers within schools, the intraclass correlation coefficient is:

$$\hat{\rho} = \frac{\hat{\tau}_\pi}{\hat{\tau}_\beta + \hat{\tau}_\pi + \hat{\sigma}^2}$$

Where $\hat{\tau}_\pi$ is the variance of the teacher-level random effect, $\hat{\tau}_\beta$ is the variance of the school-level random effect, and $\hat{\sigma}^2$ is the student-level variance (Raudenbush & Bryk, 2002). Consequently, the intra-class correlation for the variability of student achievement on NAEP civics is:

$$\hat{\rho} = \frac{\hat{\tau}_\beta}{\hat{\tau}_\beta + \hat{\tau}_\pi + \hat{\sigma}^2}$$

There are also two reliability statistics in the three-level model. The reliability of the sample mean for each teacher is:

$$reliability(\hat{\pi}_{0jk}) = \hat{\tau}_\pi / [\hat{\tau}_\pi + \frac{\hat{\sigma}^2}{n_{jk}}]$$

Where n_{jk} is the number of students for teacher *j* in school *k* (Raudenbush & Bryk, 2002). The reliability of the school mean for each school is:

$$reliability(\hat{\beta}_{00k}) = \frac{\hat{\tau}_\beta}{\hat{\tau}_\beta + \{\sum[\hat{\tau}_\pi + \frac{\hat{\sigma}^2}{n_{jk}}]^{-1}\}^{-1}}$$

The overall teacher and school sample mean reliabilities were calculated by averaging the reliabilities for all teacher and schools (Raudenbush & Bryk, 2002).

Conditional Models. The first three-level conditional model (Model 2) included only student-level predictors. The level -1 model can be expressed as:

$$Y_{ijk} = \pi_{0jk} + \sum_{Q=1}^{Q} \pi_{qjk} a_{qijk} + e_{ijk}$$

Where a_{qijk} are the student-level covariates that were entered into the model. Student level covariates include gender, race, English Language Leaners status, disability status, days absent from school, eligibility for free or reduced priced lunch, parental education, and the home educational resource index.

At this stage, the level-2 and 3 models did contain any predictors but allowed the estimate of the intercept to vary randomly between schools and among teachers within schools:

Level 2

$$\pi_{0jk} = \beta_{00k} + r_{0jk}$$

Level 3

$$\beta_{00k} = \gamma_{000} + u_{00k.}$$

Next, in Model 3 student level predictors were, one at a time, allowed to vary randomly across teachers within schools. As with the two-level model, the random coefficient for each slope was only retained if the variance in the slope was statistically significant, the slope was reliable, and the random coefficient significantly improved the fit of the model. In Model 4, teacher-level covariates were added to model for the student-level intercept:

$$\pi_{0jk} = \beta_{00k} + \sum_{v=1}^{V} \beta_{0vk} X_{vjk} + r_{0jk}$$

Where X_{vjk} represents the teacher-level covariates. The models for the student-level predictor slopes were:

$$\pi_{qjk} = \beta_{q0k} + r_{qjk} \text{ (for randomly varying slopes)}$$

$$\pi_{qjk} = \beta_{q0k} \text{ (for non-randomly varying slopes)}$$

Next, school-level covariates were added to the level-3 intercept (Model 5):

$$\beta_{00k} = \gamma_{000} + \sum_{T=1}^{T} \gamma_{00t} W_{tk} + u_{00k}$$

Where W_{tk} represents the school-level covariates. In Model 6, the slopes for teacher-level predictor were allowed, one at a time, to vary randomly between schools. As in Model 3, the random coefficients was only retained if the variance in the slope was statistically significant, the slope was reliable, and the random coefficient significantly improved the fit of the model over the model with a fixed coefficient. The models for the teacher-level slopes in Model 6 were:

$$\beta_{v0k} = \gamma_{v00} + u_{v0k} \text{ (for randomly varying slopes)}$$

$$\beta_{pqk} = \gamma_{v00} \text{ (for non-randomly varying slopes)}$$

In the final model (Model 7), the total number of teacher professional learning activities was added to the level-2 intercept model:

$$\pi_{0jk} = \beta_{00k} + \sum_{V=1}^{V} \beta_{0vk} X_{qij} + \sum_{S=v+1}^{S} \beta_{0sk} (PDTOTAL)_j + r_{0jk}$$

Where β_{0sk} represents the fixed effect for the relationship between the number of teacher professional learning activities and student achievement on NAEP civics controlling for student, teacher, and school characteristics.

RQ3. How does the form of professional development (traditional or communities of practice) relate to teachers' classroom practices and student achievement on the NAEP civics assessment?

This research question was answered by adding indicators of each form of professional development separately to model and then comparing the difference between the fixed effects to determine if the difference between the coefficients is significantly different from zero. Each form of professional development was estimated separately because the parameter of interest in

this study is the independent contribution of each form of professional development, not the marginal contribution of each form when controlling for participation in the other form. Significant differences in the fixed effects would suggest that the relationship between professional learning with classroom practices and student achievement varies depending on the form of professional learning. Additionally, it would suggest that some forms of professional learning may be more effective than others.

RQ3a. How does the form of professional development relate to teachers' use of interactive civics activities in the classroom? To answer this question, each form of professional development was entered separately into the model to predict teachers' use of interactive civics activities. In Model 5, the total number of traditional teacher professional development activities $(TRADPD)_{ij}$ was added to the model:

$$Y_{ij} = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} + \sum_{P=q+1}^{P} \beta_{pj} (TRADPD)_{pij} + r_{ij}$$

Where β_{pj} is the fixed effect for the relationship between number of traditional professional development activities and use of interactive civic activities. In Model 6, the number of traditional professional development activities was removed from the model and the number of communities of practice professional development activities $(COPPD)_{ij}$ was added to the model:

$$Y_{ij} = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} + \sum_{S=q+2}^{S} \beta_{sj} (COPPD)_{sij} + r_{ij}$$

Where β_{sj} is the fixed effect for the relationship between number of communities of practice professional development activities and use of interactive civics in the classroom.

The difference between the coefficients for the two form of professional development ($\beta_{pj} - \beta_{sj}$) was then calculated to determine if the difference is significantly different from zero. The standard error for the difference between the coefficients was equal to:

$$SE_{\beta_{pj}-\beta_{sj}} = \sqrt{(SE_{\beta_{sj}})^2 + (SE_{\beta_{sj}})^2}$$

With the coefficients assumed to be statistically independent.[8] The statistical significance of the difference was assessed using a z-test:

$$z = \beta_{pj} - \beta_{sj}/(SE_{\beta_{pj}-\beta_{sj}})$$

The z statistic has a standard normal distribution under the null hypothesis that two coefficients are equal (Clogg, Petkova, & Haritou, 1995). Thus, if the area under the probability distribution is less than .05 we would reject the null hypothesis and conclude that the difference between the coefficients is not equal to zero. This would suggest that the two forms of professional learning have different relationships to teachers' use of interactive civics.

RQ3b. How does the form of professional development (traditional or communities of practice) relate to students' civic achievement on NAEP, controlling for school, teacher, and student characteristics? As in the previous research question, each form of professional development was entered separately into the model. First, the number of traditional professional development teachers participated in $(TRADPD)_{jk}$ was used to predict the variance in the teacher level intercept for student achievement on NAEP (Model 8). All other model properties (e.g., covariates, fixed effects) remained the same as in the previous research question:

[8] This assumption may not be correct; however, because HLM6 does not allow the computation of simultaneous models the covariance of the fixed effects cannot be calculated. This means that the standard error of the difference between the means is likely underestimated in this analysis. One consequence of this is that the tests of the difference between the coefficients may be overly conservative (see Clogg et al., 1995).

$$\pi_{0jk} = \beta_{00k} + \sum_{V=1}^{V} \beta_{0vk} X_{vjk} + \sum_{S=V+1}^{S} \beta_{0sk} (TRADPD)_{sjk} + r_{0jk}$$

Where β_{0sk} is the estimated fixed effect for the relationship between number of teacher traditional professional development activities and student achievement on NAEP civics. Next, the number of communities of practice professional development activities $(COPPD)_{jk}$ teachers participated was entered into a separate model to predict variation in the teacher level intercept (Model 9):

$$\pi_{0jk} = \beta_{00k} + \sum_{V=1}^{V} \beta_{0vk} X_{qij} + \sum_{Z=q+2}^{Z} \beta_{0zk} (COPPD)_{j} + r_{0jk}$$

Where β_{0zk} is the fixed effect for the relationship between number of communities of practice professional development activities and student achievement on NAEP civics. The difference between the fixed effects for traditional and communities of practice professional development was assessed using a z-test to determine if the two forms of professional differ significantly in their relationships to student achievement on NAEP civics.

RQ4. How are teachers' classroom practices related to student outcomes, and what role does the amount and type of professional development play in that relationship?

RQ4a. Do students of teachers who use interactive civics activities in the classroom, have higher levels of civic achievement on NAEP, controlling for school, teacher, and school characteristics? This research question used the same model-building procedures as the previous research questions. Teachers' use of interactive civics $(INTACTV)_{jk}$ was then entered into the model for the level-2 intercept (Model 9):

$$\pi_{0jk} = \beta_{00k} + \sum_{V=1}^{V} \beta_{0vk} X_{vjk} + \sum_{S=s+1}^{S} \beta_{0sk} (INTACTV)_{sjk} + r_{0jk}$$

Where β_{0sk} is the fixed effect for the relationship between interactive civics usage and student achievement on the NAEP civics assessment. In the next model (Model 10), a quadratic term was added to model test for a possible non-linear relationship between interactive civics use and achievement on NAEP civics:

$$\pi_{0jk} = \beta_{00k} + \sum_{V=1}^{V} \beta_{0vk} X_{vjk} + \sum_{S=s+1}^{S} \beta_{0sk} (INTACTV)_{sjk} + \sum_{Z=z+1}^{Z} \beta_{0sk} (INTACTV)^2{}_{sjk} + r_{0jk}$$

RQ4b. Does teachers' engagement in more overall professional learning activities moderate the relationship between teachers' use of interactive civics activities in the classroom and student achievement on the NAEP civics assessment, controlling for school, teacher, and student characteristics? In order to test whether teacher professional learning served as a moderator of interactive civics activities, an interaction term representing the product of number of professional learning activities and use of interactive civics $(PDTOTAL \times INTACTV)_{jk}$ was added to model. (Baron & Kenny, 1986). This term were added to Model 8 along with number of professional learning activities $(PDTOTAL)_{jk}$ and interactive civics activities$(INTACTV)_{jk}$:

$$\pi_{0jk} = \beta_{00k} + \sum_{V=1}^{V} \beta_{0vk} X_{vjk} + \sum_{S=V+1}^{S} \beta_{0sk} (INACTV)_{sjk} + \sum_{Z=v+2}^{Z} \beta_{0zk} (PDTOTAL)_{zjk} + \sum_{A=v+3}^{A} \beta_{0ak} (PDTOTAL \times INTACTV)_{ajk} + r_{0jk}$$

Where β_{0ak} represents the fixed effect for the interaction between overall professional learning activities and interactive civics usage.

RQ4c. Does the form of professional learning moderate the relationship between teachers' use of interactive civics activities in the classroom and student achievement on the

NAEP civics assessment, controlling for school, teacher, and student characteristics?

Interaction terms with interactive civics usage were created for both the indicators of traditional professional development $(TRADPD \times INTACTV)_{jk}$ and the indicators of communities of practice $(COPPD \times INTERACTV)_{jk}$. Subsequently, Model 8 included traditional professional development activities$(TRADPD)_{jk}$, interactive civics activities$(INTACTV)_{jk}$, and the interaction term $(TRADPD \times INTACTV)_{jk}$

$$\pi_{0jk} = \beta_{00k} + \sum_{v=1}^{V} \beta_{0vk} X_{vjk} + \sum_{S=q+1}^{S} \beta_{0sk} (INACTV)_{sjk} + \sum_{Z=q+2}^{Z} \beta_{0zk} (TRADPD)_{zjk} + \sum_{A=q+3}^{A} \beta_{0ak} (TRADPD \times INTACTV)_{ajk} + r_{0jk}$$

where β_{0ak} is the fixed effect for the interaction between traditional professional development activities and interactive civics usage. In Model 9, communities of practice professional development activities $(COPPD)_{jk}$ replaced traditional professional development in the model

$$\pi_{0jk} = \beta_{00k} + \sum_{q=1}^{V} \beta_{0vk} X_{qj} + \sum_{S=q+1}^{S} \beta_{0sk} (INTACTV)_{sjk} + \sum_{C=q+4}^{C} \beta_{0ck} (COPPD)_{cjk} + \sum_{D=q+5}^{D} \beta_{0dk} (COPPD \times INTACTV)_{sjk} + r_{0jk}$$

where β_{0dk} is the fixed effect for the interaction between communities of practice professional development and interactive civics usage. The difference between the interactions terms was then calculated and a z-test was used to assess whether the two forms of professional learning had significantly different moderating effects.

Chapter 4

The following chapter will describe the results of the analysis. It will begin with a description of the data: describing the extent of missing data and the procedures used to impute missing values and reporting descriptive statistics. It will then describe the analysis results organized by research question.

Missing Data

Student-level missing data was minimal (Table 6). None of the students were missing data on the outcome variables, the NAEP plausible values. There was also no missing data for any of the school-reported student characteristics such as race, gender, and IEP and English Language Learner status. The only variables missing values were those reported by the student: parent highest education, home educational resources, and days absent from school. Parent education had a particularly high number of missing cases with 11.5% ($n = 1020$)[9] of students missing information.

Little's MCAR test (1988) was conducted to test the null hypothesis that the pattern of missing data is missing completely at random; that the pattern of missingness was unrelated to any observed or unobserved variable. A rejection of the null hypothesis suggests that the pattern of missing is not likely to be missing completely at random. For student-level data, Little's MCAR test was statistically significant ($\chi^2(337) = 1232.11,, p < .001$) indicating that the null hypothesis that the data are missing completely at random should be rejected. Additionally, the pattern of missingness was significantly correlated with certain student characteristics (Table G1 in Appendix G). Missing data on the parent education variable was positively correlated with free or reduced priced lunch status ($r = .18, p < .001$) and being an English Language Learner

[9] All whole numbers are rounded to the nearest ten and all percentages rounded to the nearest tenth.

($r = .17, p < .001$). Missing data for parent highest education was also negatively correlated with students' achievement on the NAEP civics assessment ($r = -.23, p < .001$).[10] These patterns of missing data suggest that parental education was not missing completely at random.

Table 6

Student-Level Missing Data

		Missing	
Variable	N	Count	Percent
NAEP Civics Plausible Values			
Plausible Value 1	8880	0	0.00%
Plausible Value 2	8880	0	0.00%
Plausible Value 3	8880	0	0.00%
Plausible Value 4	8881	0	0.00%
Plausible Value 5	8880	0	0.00%
Parent Highest Education	7860	1020	11.50%
Free or Reduce Priced Lunch	8880	0	0.00%
Home Index Variables			
Newspaper	8860	20	0.20%
Magazine	8860	20	0.30%
Computer	8800	80	0.90%
Encyclopedia	8860	30	0.30%
Number of Books	8860	20	0.30%
Race	8880	0	0.00%
Gender	8880	0	0.00%
Days Absent From School	8850	30	0.40%
English Language Learner	8880	0	0.00%
Individualized Educational Plan (IEP)	8880	0	0.00%

Teachers had more variables with missing data than students. However, the extent of missing data was minimal, never exceeding more than 5% of all observations (Table 7). Little's MCAR test was not statistically significant, meaning that the null hypothesis that the data are missing completely at random could not be rejected ($\chi^2(2419) = 2489.40, p = .156$). This suggests that the pattern of missing data may be missing completely at random. Additionally,

[10] First plausible values used an example. Other plausible values had similar negative correlations.

examining the correlation between teacher characteristics and the pattern of missing data revealed very few significant and sizable relationships (Tables G2-G5 in Appendix G).

Table 7

Teacher-Level Missing Data

		Missing	
Variable	N	Count	Percent
Professional Development Activities			
College Course	1080	40	3.80%
Workshop or Training	1100	20	1.90%
Conference or Professional Association Meeting	1080	40	3.40%
Observational Visit to Another School	1070	50	4.20%
Mentoring/Coaching/Peer Observation	1090	30	2.90%
Committee or Task Force	1090	30	2.70%
Discussion or Study Group	1090	30	3.00%
Teacher Collaborative or Network	1070	50	4.30%
Individual or Collaborative Research	1080	40	3.30%
Independent Reading on a Regular Basis	1090	30	2.90%
Co-teaching/Team-teaching	1080	40	3.30%
Consultation with Subject Specialist	1080	40	3.50%
School Leadership position in Social Studies	1120	10	0.50%
Classroom Use of Interactive Civics Activities			
Debates or Panel Discussion	1110	10	1.10%
Mock Trials, Role-Playing, or Dramatizations	1110	10	1.10%
Write a Letter to State Opinion	1110	10	1.10%
Community Visitors	1110	10	1.00%
Community Volunteer Projects	1110	10	0.90%
Discuss Current Events	1110	10	0.70%
Student Government	1110	20	1.30%
Group Projects	1110	10	1.20%
Group Presentations	1110	10	1.10%
Teacher Race	1120	0	0.00%
Years Teaching 6-12 Social Studies	1120	0	0.30%
Regular/Standard Certification	1100	20	1.60%
National Board Certification	1100	20	2.10%
Undergraduate Major	1110	10	1.30%
Graduate Major	1120	10	0.50%
Teaches Only Social Studies	1100	20	2.10%
Number of Students in Sample	1120	0	0.00%

At the school level, missing data were only present in two variables: the percent of students who were eligible for free or reduced priced lunch ($n < 3$) and the number of students

absent on an average day ($n = 30$) (Table 8). Information about the number of students absent on an average day was taken from the principal survey. Schools were missing data on this variable because the principal either skipped the question or did not complete the survey. Complete data was available for all other variables. Little's MCAR test was not statistically significant, indicating that the null hypothesis that the items were missing completely at random could not be rejected ($\chi^2(25) = 23.853, df = 25, p = .528$). The correlation matrix for the pattern of missingness at the school-level (Table G6 in Appendix G) indicated that private schools and schools in the Northeast were more likely to be missing the free or reduced priced lunch indicator. There were no obvious patterns for which schools were more likely to be missing data about student absences.

Table 8

School-Level Missing Data

		Missing	
Variable	N	Count	Percent
Urbanicity	450	0	0.00%
Region	450	0	0.00%
School Affiliation	450	0	0.00%
Student Absences % on Average Day	420	30	6.90%
Student % Free or Reduced Priced Lunch	450	<3	0.50%
Student % Minority	450	0	0.00%
Number of Teachers in Sample	450	0	0.00%

Note. Cell sizes less than three not reported in order to protect confidentiality of schools

Although the amount of missing data was not extensive, the pattern of missing data at the student level suggested that the data were not plausibly missing completely at random. This indicated that deleting observations with missing data could bias estimates. As a result, a stochastic imputation procedure was used to impute missing values. Stochastic imputation uses the observed variables in the dataset to predict a potential value for the missing information. In

order to adequately represent the uncertainty associated with this estimated an error term is included in the model (Schafer & Graham, 2002; D. Zhang, 2005). Since the amount of missing data was minimal, single imputation was selected as the procedure for imputing missing data (Schafer, 1999; Puma et al., 2009). The imputed values for student-, teacher-, and school-level variables are presented in Table 9[11].

Table 9

Comparing Imputed and Non-Imputed Variables Means and Standard Errors

	Imputed			Non Imputed		
Student-Level Variable	N	M	SE	N	M	SE
Newspaper in Home	8880	0.343	0.007	8860	0.343	0.007
Magazine in Home	8880	0.612	0.007	8860	0.613	0.007
Computer in Home	8880	0.931	0.004	<8800	0.931	0.004
Encyclopedia in Home	8880	0.719	0.005	8850	0.719	0.005
Numbers of books in home	8880	1.818	0.020	8860	1.818	0.020
Parent education variable (highest mother or father)	8880	2.152	0.024	7860	2.199	0.023
Days absent from school	8880	0.801	0.011	8850	0.802	0.012
Teacher-Level Variable	N	M	SE	N	M	SE
Years' experience teaching 6-12 social studies subjects	1120	10.259	0.259	<1120	10.250	0.260
Regular/Standard Certification	1120	0.903	0.009	1100	0.903	0.009
National Board Certification	1120	0.119	0.010	1100	0.118	0.010
Undergraduate major in non-social science	1120	0.387	0.015	1110	0.388	0.015
Undergraduate major in social science	1120	0.613	0.015	1110	0.613	0.015
Graduate major in non-social science	1120	0.363	0.014	<1120	0.363	0.014
Graduate major in social science	1120	0.176	0.011	<1120	0.175	0.011
No graduate degree	1120	0.461	0.015	<1120	0.462	0.015
College course taken after first certification	1120	0.277	0.013	1080	0.275	0.014
Workshop or training	1120	0.748	0.013	1100	0.747	0.013
Conference or professional association meeting	1120	0.533	0.015	1080	0.525	0.015
Observational visit to another school	1120	0.229	0.013	1070	0.226	0.013
Mentoring/Coaching/Peer observation	1120	0.487	0.015	1090	0.487	0.015
Committee or Task Force	1120	0.506	0.015	1090	0.504	0.015
Regularly scheduled discussion or study group	1120	0.481	0.015	1090	0.483	0.015
Teacher collaborative or network (including online networks)	1120	0.258	0.013	1070	0.259	0.013
Individual or collaborative research	1120	0.443	0.015	1080	0.443	0.015
Independent reading on a regular basis	1120	0.770	0.013	1090	0.767	0.013
Co-teaching/team-teaching	1120	0.425	0.015	1080	0.423	0.015

[11] Values rounded to the thousandths place in order to make meaningful comparisons between imputed and non-imputed values

Table 9 Continued

	N	M	SE	N	M	SE
Consultation with subject specialist	1120	0.364	0.014	1080	0.364	0.015
School leadership position in social studies	1120	0.345	0.014	<1120	0.344	0.014
Students participate in debates or panel discussion	1120	2.208	0.027	1110	2.208	0.027
Students participate in civic simulations	1120	1.699	0.019	1110	1.700	0.019
Students write a letter to state an opinion or solve a community problem	1120	1.594	0.021	1110	1.593	0.021
Students participated in community project	1120	1.182	0.014	1110	1.180	0.014
Students discuss current events	1120	3.013	0.023	1110	3.013	0.023
Students do student government	1120	1.406	0.022	1110	1.402	0.022
School-Level Variable	N	M	SE	N	M	SE
% Students absent on an average day	450	1.795	0.066	420	1.799	0.066
% Students Free or Reduced Priced Lunch	450	34.265	2.444	450	35.309	2.321

In general, the differences in the means between the imputed and non-imputed values were very small. Differences between standard error of the means were marginal and the imputed values often had larger standard errors than the non-imputed values. These results suggests that the imputation procedures did not substantially bias the estimates or artificially reduce the standard errors of the variables.

Descriptive Statistics

The weighted student level descriptive statistics are presented in Table 10. To summarize, 59% of students in the weighted sample were White, 20% were Hispanic, 14% were Black, 5% were Asian or Pacific Islander, 1% were Native American, and 1% were some other race. Girls made up approximately half (49%) of students in the sample. Additionally, nearly half (42%) of students in the weighted sample were eligible for free or reduced priced lunch and substantial portions of students had IEPs (10%) or were classified as English Language Learners (5%). The composition of the weighted sample was representative of the known 8th grade United States student population in 2009-2010 in terms of race, gender, and free or reduced priced lunch status (National Center for Education Statistics, 2010). However, students who were English Language

Learners or had IEPs were slightly underrepresented in the weighted sample (National Center for Education Statistics, 2010). These students were more likely to be excluded from the sample by NAEP because they would not have been able to complete the assessment, even with accommodations.

Table 10

Student Descriptive Statistics

Variable	N	Min	Max	M	SE	SD	.95 Confidence Interval	
							LO	HI
White Non-Hispanic only	8880	0.00	1.00	0.59	0.01	0.49	0.56	0.62
Black Non-Hispanic only	8880	0.00	1.00	0.14	0.01	0.35	0.13	0.16
Hispanic, any race	8880	0.00	1.00	0.20	0.01	0.40	0.17	0.23
Asian/Pacific Islander Non-Hispanic only	8880	0.00	1.00	0.05	<0.01[a]	0.21	0.04	0.05
Native American	8880	0.00	1.00	0.01	<0.01[b]	0.09	0.00	0.02
Other (School records)	8880	0.00	1.00	0.01	<0.01[c]	0.11	0.01	0.02
Eligible for Free or Reduced Priced Lunch	8880	0.00	1.00	0.42	0.01	0.49	0.39	0.45
Student has Individualized Education Plan (IEP)	8880	0.00	1.00	0.10	<0.01[d]	0.30	0.09	0.11
Student is an English Language Learner (ELL)	8880	0.00	1.00	0.05	<0.01[e]	0.23	0.04	0.06
Female	8880	0.00	1.00	0.49	<0.01[f]	0.50	0.48	0.50
Newspaper in home	8880	0.00	1.00	0.34	0.01	0.47	0.33	0.36
Magazine in home	8880	0.00	1.00	0.61	0.01	0.49	0.60	0.63
Computer in home	8880	0.00	1.00	0.93	<0.01[g]	0.25	0.92	0.94
Encyclopedia in home	8880	0.00	1.00	0.72	0.01	0.45	0.71	0.73
Numbers of books in home	8880	0.00	3.00	1.82	0.02	1.01	1.78	1.86
Home index variable	8880	0.00	7.00	4.42	0.03	1.64	4.36	4.49
Parent education	8880	0.00	3.00	2.15	0.02	1.03	2.10	2.20
Days absent from school	8880	0.00	4.00	0.80	0.01	0.93	0.78	0.82
Plausible NAEP civics value #1	8880	10.94	252.93	151.94	0.76	33.36	150.42	153.45
Plausible NAEP civics value #2	8880	0.00	242.87	151.98	0.76	33.42	150.47	153.49
Plausible NAEP civics value #3	8880	6.68	258.79	151.83	0.75	33.49	150.33	153.34
Plausible NAEP civics value #4	8880	0.00	249.00	151.99	0.75	33.48	150.49	153.49
Plausible NAEP civics value #5	8880	11.79	250.82	152.19	0.76	33.48	150.67	153.71

[a]Actual value is 0.0040 [b] Actual value is 0.0033 [c] Actual value is 0.0016 [d] Actual value is 0.0042 [e] Actual value is 0.0048 [f]Actual value is 0.0049 [g] Actual value is 0.0043

Descriptive statistics for the teacher variable are presented in Table 6. Teachers in the sample were much more likely than their students to be white (79%) and much less likely to be Hispanic (9%). Additionally, only 61% of teachers in the sample reported majoring in a social science (e.g., history, political science) as an undergraduate and only 18% had a graduate degree in a social science. Most of the teachers in the sample (74%) reported that the only subject they taught was social studies. The average number of years' of experience teaching social studies was 10.26 years.

Table 11

Teacher Descriptive Statistics

Variable	N	Min	Max	M	SE	SD	.95 Confidence Interval	
							LO	HI
Teacher is White, Non-Hispanic	1120	0	1	0.79	0.01	0.41	0.76	0.81
Teacher is Black, Non-Hispanic	1120	0	1	0.09	0.01	0.28	0.07	0.10
Teacher is Hispanic, any race	1120	0	1	0.09	0.01	0.29	0.08	0.11
Teacher is Asian/Pacific Islander, Non-Hispanic	1120	0	1	0.01	<0.01[a]	0.12	0.01	0.02
Teacher is Native American, Non-Hispanic	1120	0	1	0.01	<0.01[b]	0.08	<0.01[c]	0.01
Teacher is Other/Multiracial, Non-Hispanic	1120	0	1	0.01	<0.01[d]	0.11	0.01	0.02
Years' experience teaching 6-12 social studies subjects	1120	0	66	10.26	0.26	8.68	9.75	10.77
Years' experience teaching 6-12 social studies subjects (squared)	1120	0	4356	180.51	9.33	312.25	171.01	190.02
Regular/Standard Certification	1120	0	1	0.90	0.01	0.30	0.89	0.92
Certified by the National Board for Professional Teaching Standards	1120	0	1	0.12	0.01	0.32	0.10	0.14
Undergraduate major in non-social science	1120	0	1	0.39	0.01	0.49	0.36	0.42
Undergraduate major in social science	1120	0	1	0.61	0.01	0.49	0.58	0.64
Graduate major in non-social science	1120	0	1	0.36	0.01	0.48	0.33	0.39
Graduate major in social science	1120	0	1	0.18	0.01	0.38	0.15	0.20
No graduate degree	1120	0	1	0.46	0.01	0.50	0.43	0.49
Only teaches social studies	1120	0	1	0.74	0.01	0.44	0.71	0.76
Number of students per teacher in sample	1120	1	30	7.92	0.16	5.44	7.60	8.24

[a] Actual value is 0.0035 [b] Actual value is 0.0025 [c] Actual value is 0.0022 [d] Actual value 0.0033

The weighted school-level descriptive statistics are described in Table 12. Schools in rural areas comprised 50% of the sample with suburban (24%) and urban (26%) schools each representing about a quarter of all schools. Most of the schools in the sample (60%) were public non-charter schools, with the remaining schools either private (37%) or charter (3%). The average percentage of students who were eligible for free or reduced priced lunch in each school was 34%[12], and the average percentage of minority students was 32%.

Table 12

[12] The difference between the mean number of free and reduced priced lunch students and the average percentage of students who were eligible for free or reduced priced lunch in each school reflected the significant number of schools have no students who receive free or reduced priced lunch so this reduced the overall average.

School Descriptive Statistics

	N	Min	Max	M	SE	SD	.95 Confidence Intervals LO	HI
School in City	450	0	1	0.26	0.03	0.44	0.20	0.32
School in Suburb	450	0	1	0.24	0.03	0.43	0.17	0.31
School in Rural Area or Town	450	0	1	0.50	0.04	0.50	0.42	0.58
Northeast	450	0	1	0.22	0.03	0.42	0.16	0.29
Southeast	450	0	1	0.23	0.03	0.42	0.18	0.28
Central	450	0	1	0.29	0.03	0.45	0.22	0.35
West	450	0	1	0.26	0.03	0.44	0.20	0.31
Charter school	450	0	1	0.03	0.02	0.18	0.00	0.07
Private school	450	0	1	0.37	0.04	0.48	0.29	0.45
Public (non-Charter) school	450	0	1	0.60	0.04	0.49	0.52	0.68
Students Absent on an Average Day	450	1	4	1.79	0.07	0.70	1.66	1.93
% Students Free or Reduced Priced Lunch	450	0	100	34.26	2.44	32.16	29.38	39.15
Percent Minority Students	450	0	100	32.31	2.59	33.71	27.14	37.48
Number of Teachers in Sample	450	1	8	1.58	0.05	1.05	1.49	1.67

Constructed Indices

Teacher professional development. Teacher professional development was measured using three different indices: total professional development, traditional professional development, and communities of practice professional development. Total professional development represented teachers' overall participation in 13 different content-specific professional development activities over the last two years. On average teachers in the study participated in 5.86 types of professional learning activities related to teaching social studies (SD=3.21). The distribution of these activities in described in further detail in the Analysis section under Research Question 1.

Traditional professional development included only those forms of professional learning where the primary goal is to transfer information from the presenter to the participants. Some examples of traditional professional learning include courses, workshops, or conferences. Communities of practice professional development consisted of activities that involved teacher interaction and collaboration such as co-teaching, coaching and mentoring, and participating in

teacher networks. On average, teachers in the study participated in 2.69 traditional activities (SD=1.45) and 3.17 communities of practice activities (SD=2.11).

Interactive civics. Interactive civics use was measured based on teachers' self-reported use of nine different types of instructional activities (Table 13). Teachers' use of interactive civics varied considerably depending on the activity. For example, 97% of teacher reported discussing current events, but only 16% had students do a community project and 11% of teachers brought in visitors from the community. Few teachers reported conducting any interactive civics activity, other than discussing current events, more than once or twice a week.

Table 13

Frequencies of Use of Interactive Civics Activities

Interactive Civics Activity	Never	Once or Twice a Month	Once or Twice a Week	Almost Every Day
Students participate in debates or panel discussion	22.2%	44.0%	24.6%	9.2%
Students participate in civic simulations	39.1%	52.6%	7.6%	0.7%
Students write a letter to state an opinion or solve a community problem	51.7%	38.0%	9.6%	0.7%
Students participated in community project	84.1%	14.0%	1.4%	0.4%
Students discuss current events	2.7%	21.9%	46.9%	28.5%
Students do student government	71.2%	19.1%	7.7%	2.1%
Student work on group projects	25.3%	62.4%	11.2%	1.0%
Students make group presentations	33.6%	57.9%	8.0%	0.4%
Students interact with visitors from the community	88.7%	11.0%	0.4%	0.0%

As described in Chapter 3, cognitive interviews with current social studies teachers were used to weight the frequency of each activity based on their intensity level. Activities were designated as either 0 "No impact" to 3 "High impact" on student learning (Table 14). Teachers in the study had a mean of 11.07 (SD=4.82) on the interactive civics use index. This is the approximately equivalent to conducting three high impact interactive civics activities and one medium impact activity over the course of the year.

Table 14

Frequencies of Use of Interactive Civics Activities by Perceived Impact on Student Learning

Interactive Civics Activity	(0) No Impact	(1) Low impact	(2) Medium impact	(3) High Impact
Students participate in debates or panel discussion	22.2%	--	--	77.8%
Students participate in civic simulations	39.1%	--	--	60.9%
Students write a letter to state an opinion or solve a community problem	51.7%	--	48.3%	--
Students participated in community project	84.1%	--	--	15.9%
Students discuss current events	2.7%	21.9%	46.9%	28.5%
Students do student government	71.2%	28.8%	--	--
Student work on group projects	25.3%	--	74.7%	--
Students make group presentations	33.6%	--	66.4%	--
Students interact with visitors from the community	88.7%	--	--	11.3%

Internal consistency. The internal consistency of the constructed indices was measured using the Cronbach's alpha coefficient and the average inter-item correlation (Table 10). The total number of teacher professional learning activities had a Cronbach's alpha coefficient of $\alpha = .78$ which is considered to be an acceptable level of reliability (George & Mallery, 2003; Peterson, 1994). However, the other indices had poor reliability. Traditional professional learning had a Cronbach's alpha coefficient of $\alpha = .63$, communities of practice professional learning $\alpha = .67$, and interactive civics $\alpha = .68$. The average inter-item correlation for all indices were between $r = .2$ and $r = .3$. This suggests that although the items that were used to form the composite were somewhat related to one another, the correlations between the items were not particularly high. This may be because the underlying construct is multidimensional or because the way the items were worded or scaled introduced measurement error. The implications of the low reliability of the indices will be discussed further in the Limitations section of Chapter 5.

Research Questions

This section will examine the findings of the study by research question. The first research question examined the patterns of teacher participation in professional development. This question was explored through descriptive statistics and univariate analyses. The second question investigated how teacher professional learning was related to use of interactive civics activities and students' civic knowledge and skills. These relationships were estimated using multilevel models. The third research question focused on the relationship between the form of professional learning, traditional or communities of practice, and teachers' use of interactive civics and students' civic skills and knowledge. As with the previous research question, this question was addressed through a multilevel analysis. The final research questions examined how the use of interactive civics related to students' civic knowledge and skills and whether teacher professional learning moderated this relationship. This question was explored through the introduction of interaction terms into the multilevel analysis.

RQ1: What are the patterns of civics teachers' professional learning activities?

RQ1a: What types of professional learning activities do teachers engage in?

Teachers in the study reported participating in many content-related professional learning activities. Three-quarters of teachers in the study (74.8%) reported that they participated in a workshop on teaching social studies in the past two years and a similar percentage (77.0%) engaged in independent reading on social studies instruction on a regular basis. Communites of practice forms of professional development, though less prevalent than traditional forms, were still reported by a substantial portion of teachers. For example, 42.5% of teachers co-taught or participated in instructional teams, 44.3% engaged in collaborative research, and 48.7% received mentoring or coaching. Finally, more than a quarter of teachers (25.8%) participated in a teacher

network devoted to social studies teaching. Table 15 presents the frequencies for each form of professional development examined in this study.

Table 15

Teacher Participation in Professional Learning

Professional Learning Activity	% Participated	.95 confidence interval	
		LO	HI
Observational visit to another school	22.9%	20.5%	25.4%
Teacher collaborative or network (including online networks)	25.8%	23.2%	28.3%
College course taken after first certification	27.7%	25.0%	30.3%
School leadership position in social studies	34.5%	31.7%	37.3%
Consultation with subject specialist	36.4%	33.6%	39.2%
Co-teaching/team-teaching	42.5%	39.6%	45.4%
Individual or collaborative research	44.3%	41.3%	47.2%
Regularly scheduled discussion or study group	48.1%	45.2%	51.0%
Mentoring/Coaching/Peer observation	48.7%	45.8%	51.6%
Committee or Task Force	50.6%	47.7%	53.5%
Conference or professional association meeting	53.3%	50.3%	56.2%
Workshop or training	74.8%	72.2%	77.3%
Independent reading on a regular basis	77.0%	74.5%	79.4%

RQ1b: How do these patterns of professional learning activities vary by teacher and school characteristics? Teacher patterns of professional learning were disaggregated by teacher and school characteristics. Chi-square tests, for nominal variables, and independent sample t-tests, for interval variables, were used to determine what characteristics were significantly related to participation in each activity. A Benjamini-Hochberg correction was use to adjust for multiple comparisons. The Benjamini-Hochberg correction adjusts the α level so that false discovery is less than .05. In this analysis, the Benjamini-Hochberg correction resulted in a α level of .00342. This indicates that in order to be a statistically significant predictor, the test statistics for a teacher or school characteristic needed to have at p value less than or equal to .00342.

Certain teacher characteristics significantly predicted whether a teacher was likely to participate in professional learning (Table 16). One common predictor was teacher educational background. Teachers with undergraduate or graduate degree in the social sciences were more

likely to participate in 10 out of the 13 professional learning activities examined in this study. In contrast, teachers with graduate majors in non-social studies were less likely to attend conferences or participate in workshops on social studies instruction. Other teacher qualities that were positively associated with participating in professional learning included teaching social studies exclusively and years of experience teaching social studies. Teachers who were National Board certified were more likely to report participating in peer observations. This may be because receiving feedback on videotaped lessons is part of the certification process to become National Board certified (Sato, Wei, & Darling-Hammond, 2009).

School characteristics also predicted teacher participation in certain forms of professional development. Public school teachers were more likely than charter or private school teachers to enroll in a college course, team-teach, or consult with a subject specialist. Private school teachers were less likely than their peers to take college courses, serve on a committee, participate in a discussion group, team-teach, or consult with a subject specialist. There were also some regional effects. Teachers in the Northeast were more likely to report team-teaching, consulting with a subject specialist, and attending professional conferences while those in the West were more likely to participate in observational visits but were less likely to team-teach. Midwestern teachers and teachers in rural areas were less likely to participate in discussion groups. School student demographics were not related to participation in teacher professional learning with the sole exception of independent reading on a regular basis.

Table 16

Teacher and School Significant Predictors of Engaging in Professional Learning Activities

Professional Learning Activity	Characteristics of Teachers	
	More likely to participate in activity…	Less likely to participate in activity…
College course taken after first certification	Undergrad major in social science Graduate major in social science Works at a public school	Works at a private school
Workshop or training	Undergrad major in social science Graduate major in social science Only teaches social studies Years of experience teaching social studies	Graduate major in non-social science
Conference or professional association meeting	Undergrad major in social science Graduate major in social science Only teaches social studies Years of experience teaching social studies Lives in the Northeast	Graduate major in non-social science
Observational visit to another school	Undergrad major in social science Graduate major in social science Lives in the West	--
Mentoring/Coaching/Peer observation	Undergrad major in social science Graduate major in social science National Board Certification Only Teaches Social Studies	--
Committee or Task Force	Undergrad major in social science Only teaches social studies Years of experience teaching social studies	Works at a private school
Regularly scheduled discussion or study group	Only teaches social studies Works at a public school	Works at a private school Lives in a rural area Lives in the Midwest
Teacher collaborative or network	--	--
Individual or collaborative research	Graduate major in social science	--
Independent reading on a regular basis	Undergrad major in social science Only teaches social studies Years of experience teaching social studies	Percent of students in school eligible for free or reduced priced lunch Percent of students in school who are racial minorities
Co-teaching/team-teaching	Undergrad major in social science Graduate major in social science Works at a public school Lives in the Northeast	Works at a private school Lives in the West
Consultation with subject specialist	Works in a public school Lives in the Northeast	Works at a private school
School Leadership Position	Undergrad major in social studies Only teaches social studies Years of experiences teaching social studies	--

Note. Significant predictors after Benjamini-Hochberg FDR correction $p < .00342$

RQ2. How is teacher engagement in overall professional learning activities related to teachers' classroom practices and student achievement on the NAEP civics assessment?

RQ2a. Are teachers who engage in more overall professional learning activities more likely to use interactive civics activities in the classroom, controlling for school and

teacher characteristics? This question was examined by modeling the relationship between overall teacher professional learning and interactive civics with a two-level hierarchical linear model with teachers at level one and schools at level two. The results are described in Table 17.

Model building. The first stage of this analysis was to fit an unconditional model (Model 1). The unconditional model contained no predictors but included a random effect for the school-level intercept, allowing the intercept for interactive civics to vary across schools. In the unconditional model the variance of the random effect for the school-level intercept was statistically significant ($\chi^2(448) = 714.19, p < .001$). Additionally, the intraclass correlation coefficient was .19 indicating that nearly one-fifth of the variation in teachers' use of interactive civics was between schools rather than among teachers within schools. This suggests that there was significant variance between schools that could be explained by the model.

Teacher-level covariates were then added to the model (Model 2). The estimated coefficients were then examined to determine if they varied significantly across schools. Only the coefficient for "only teaches social studies" varied significantly across schools ($\chi^2(133) = 174.61,\ p < .01$) and had a reliable slope ($\lambda = .22$). Additionally, the chi-square difference test indicated that allowing the slope for "only teaches social studies" to vary randomly significantly improved the fit of the model ($\chi^2(2), p < .01$). However, the random effect for the slope for only teaching social studies was highly negatively correlated with the intercept ($r = -.79$). When the random slope was introduced the model the variance of the intercept increased from $\tau_{00} = 4.20$ to $\tau_{00} = 9.48$. Allowing the slope for "only teaches social studies" to vary randomly therefore may have artificially inflated the variance in the intercept. Since the variance in the intercept was the primary focus of this investigation, it was determined that slope for "only teaches social studies", as well all other teacher covariates, should be fixed in all future models

meaning that they were not allowed to vary between schools. Once the random slopes for teacher-level covariates were fixed, school-level covariates were added to the model to predict variation in the school-level intercept (Model 3). In the final stage (Model 4), the total number of professional learning activities that teachers' participated was added in the model.

Estimated fixed effects. The fixed effect for teacher participation in professional learning was statistically significant $\left(\hat{\beta} = .47, t = 11.05, p < .001\right)$. For each additional professional learning activity that a teacher participated in their estimated use of interactive civics increased by about one-half a unit. In other words, for each two additional professional development activities teachers participated in, the model predicted that they would conduct one additional low intensity interactive civics activity, such as discussing current events on a monthly basis. Teacher professional learning explained an additional 9.6% of the variation in interactive civics use among teachers within schools. The standardized effect size was .32 standard deviations which is considered a small to medium effect size (J. Cohen, 1992; C. Hill et al., 2007).

Another way to interpret this relationship is by comparing a teacher from the bottom quartile (25^{th} percentile) in terms of their professional development participation to a teacher in the highest quartile (75^{th} percentile). Teachers in the top quartile for professional development would, on average, be expected to have an interactive civics index score 1.86 units higher than teacher in the bottom quartile. This means that compared with bottom quartile teachers, top quartile teachers used the equivalent of one more medium impact activity, such as having students work on a group project or write a letter about a community problem.

Table 17

HLM Results: Professional Learning and Use of Interactive Civics

Parameter	Model 1		Model 2		Model 3		Model 4	
	Coeff	(SE)	Coeff	(SE)	Coeff	(SE)	Coeff	(SE)
Intercept	11.14***	(0.17)	10.25***	(0.64)	10.29***	(0.76)	10.82***	(0.71)
Teacher Variables								
Teacher Non-White			0.11	(0.35)	0.19	(0.37)	-0.06	(0.36)
Years Teaching ‡			-0.05*	(0.02)	-0.05*	(0.02)	-0.08*	(0.02)
Years Teaching Squared‡			<0.01[a]	(<0.01)[a]	<0.01[b]	(<0.01)[b]	<0.01[d]	(<0.01)[d]
Regular/Standard Certification			-0.63	(0.51)	-0.56	(0.52)	-0.56	(0.50)
National Board Certification			1.70***	(0.43)	1.71***	(0.44)	1.41***	(0.44)
Undergraduate Social Science			0.86	(0.30)	0.88	(0.30)	0.47	(0.29)
Graduate Non Social Science			0.44	(0.30)	0.38	(0.31)	0.45	(0.29)
[Ref=No Graduate Degree]								
Graduate Social Science			1.44***	(0.40)	1.36***	(0.40)	0.98**	(0.37)
Only Teaches Social Studies			0.38	(0.36)	0.24	(0.36)	-0.17	(0.34)
Number Students in Sample			0.04	(0.03)	0.02	(0.03)	-0.02	(0.03)
School Variables								
City [Ref=Suburb]					0.54	(0.43)	0.52	(0.41)
Rural					-0.86	(0.45)	-0.51	(0.42)
Northeast [Ref=West]					0.73	(0.47)	0.62	(0.46)
Southeast					0.94*	(0.41)	0.79*	(0.39)
Central					0.04	(0.54)	0.23	(0.50)
Charter [Ref=Public]					-0.74	(1.50)	0.34	(1.56)
Private					-1.15	(0.77)	-0.47	(0.71)
Student Absences‡					-0.47	(0.25)	-0.42	(0.24)
% Free Reduced Priced Lunch‡					-0.01	(0.01)	-0.01	(0.01)
% Minority‡					0.00[c]	(0.01)	0.00	(0.01)
Number Teachers in Sample‡					-0.27*	(0.13)	-0.30*	(0.12)
Total Professional Learning‡							0.47***	(0.04)
Variance Components								
Level-1 variance	18.73		17.86		17.98		16.25	
Level-2 variance	4.48***		4.20***		3.47***		3.12***	
Intra-class correlation coefficient	0.19		0.19		0.16		0.16	
Level-2 Reliability	0.35		0.34		0.30		0.30	
% Additional Variance Explained								
% Variance Explained Level-1	--		4.6%		-0.7%		9.6%	
% Variance Explained Level-2	--		6.1%		17.6%		9.8%	

Note. % Variance explained calculated in comparison to previous model
[a] actual values are 0.0001 (0.0013) [b] actual values are -0.0001 (0.0013) [c] actual value is -0.002 [d] actual values are 0.0007 (0.0009)
‡ Grand Mean Centered **p<.05 **p<.01 ***p<.001*

In addition to teacher professional learning, Model 4 contained other significant predictors of interest. Years teaching social studies had a negative relationship with interactive civics use $\left(\hat{\beta} = -.08, t = -3.62, p < .001\right)$. For each additional 10 years of experience a teachers expected use of interactive civics decreased by .16 standard deviations. By contrast, teachers with graduate degrees in a social science were predicted to use more interactive civics activities $\left(\hat{\beta} = .99\ t = 2.65, p < .01\right)$ as were National Board certified teachers $\left(\hat{\beta} = 1.37\ t = 3.13, p < .01\right)$. Teachers with graduate degrees used .093 standard deviations more interactive

civics activities than teachers with no graduate degrees and National Board certified teachers implemented .29 standard deviations more interactive civics activities.

There were also significant school-level effects. After controlling for other variables in the model, schools with more teachers used significantly fewer interactive civics activities ($\hat{\gamma} = -.27, t = -2.25, p < .05$). For each additional teacher in the sample teachers expected use of interactive civics activities decreased by .063 standard deviations. Additionally, schools in the Southeast reported greater use of interactive civics activities compared to schools in the West ($\hat{\gamma} = .73\ , t = 1.89, p < .05$). On average, teachers at schools in the Southeast used .16 standard deviations more activities than teachers in the West. There was no significant relationship between the percent of students in the school on free or reduced priced lunch and teachers' use of interactive civics activities($\hat{\gamma} = -.007\ , t = -.84.\ , p = .401$) or the percent of minority students in the school and use of interactive civics activities($\hat{\gamma} = -.001\ , t = -.16.\ , p = .874$).

RQ2b. Do students of teachers who engage in more overall professional learning activities have higher achievement on the NAEP civics assessment, controlling for school, teacher, and school characteristics? The relationship between teacher professional learning and student achievement on the NAEP civics assessment was modeled using a three-level hierarchical linear model with students at level one, teachers at level two, and schools at level three. Tables 18 and 19 summarize the results of the models used to address this research question.

Table 18

HLM Results: Professional Learning and Student Achievement on NAEP Civics (Models 1-4)

Parameter	Model 1		Model 2		Model 3		Model 4	
	Coeff	(SE)	Coeff	(SE)	Coeff	(SE)	Coeff	(SE)
Intercept	151.65***	(0.93)	163.15***	(0.76)	163.20***	(0.74)	163.31***	(1.91)
Student Variables								
Black [Ref=White]			-13.64***	(1.13)	-13.51***	(1.13)	-12.85***	(1.10)
Hispanic			-4.27***	(0.99)	-4.32***	(0.98)	-3.81***	(0.98)
Asian			0.01	(1.70)	0.23	(1.70)	0.42	(1.68)
Native-American			-4.18	(3.51)	-4.69	(3.70)	-3.99	(3.63)
Other			-2.63	(3.59)	-2.19	(3.46)	-1.99	(3.49)
FRPL			-8.00***	(0.83)	-8.15***	(0.82)	-7.88***	(0.82)
IEP			-30.69***	(1.41)	-30.61***	(1.50)	-30.32***	(1.49)
ELL			-33.15**	(1.94)	-33.50***	(2.01)	-33.13***	(2.02)
Female			-1.21	(0.65)	-1.04	(0.63)	-1.05	(0.63)
Home Index‡			3.99***	(0.22)	3.98***	(0.21)	3.97***	(0.21)
Parent Education‡			2.32***	(0.35)	2.40***	(0.34)	2.41***	(0.34)
Days absent‡			-2.96***	(0.32)	-2.98***	(0.33)	-2.95***	(0.33)
Teacher Variables								
Teacher Non-White							-5.10***	(1.32)
Years Teaching‡							0.12	(0.07)
Years Teaching Squared‡							-0.01	(<0.01[b])
Regular/Standard Certification							-2.67	(1.57)
Nationally Board Certified							1.05	(1.48)
Undergraduate Social Science							1.89	(0.97)
Graduate Non Social Science							-0.21	(1.07)
[Ref=No Graduate Degree]								
Graduate Social Science							-0.09	(1.22)
Only Teaches Social Studies							2.67*	(1.21)
Number Students in Sample							-0.02	(0.08)
Variance Components								
Level-1 variance	797.06		579.35		548.53		547.38	
Level-2 variance	125.95***		33.74***		29.90[a]		30.32[a]	
Level-3 variance	228.38***		58.56***		55.84***		50.33***	
FRLP	--		--		52.95		51.49	
IEP	--		--		220.85***		216.57***	
ELL	--		--		131.52		131.50	
Intraclass correlation coefficient								
Level-2	0.11		0.05		0.05		0.05	
Level-3	0.20		0.09		0.09		0.08	
Level-2 Reliability	0.49		0.28		0.18		0.18	
Level 3 Reliability	0.67		0.54		0.54		0.51	
% Additional Variance Explained								
% Variance Explained Level-1	--		27.3%		5.3%		0.2%	
% Variance Explained Level-2	--		73.2%		11.4%		-1.4%	
% Variance Explained Level-3	--		74.4%		4.6%		9.9%	

Note. Standard errors are in parentheses, % Variance explained calculated in comparison to previous model. FRPL=Student eligible for free or reduced priced lunch, IEP=Student has an IEP, ELL=Student is an English Language Learner

[a] Significance of Level-2 Variance could not be computed because there was not sufficient degrees of freedom

[b] Actual values are (.0043)

‡ Grand Mean Centered **p<.05 **p<.01 ***p<.001*

Table 19

HLM Results: Professional Learning and Student Achievement on NAEP Civics (Models 5-6)

	Model 5		Model 6	
Parameter	Coeff	(SE)	Coeff	(SE)
Intercept	160.18***	(2.13)	160.61***	(2.12)
Student Variables				
Black [Ref=White]	-11.87**	(1.13)	-11.95**	(1.13)
Hispanic	-3.37**	(1.09)	-3.40**	(1.09)
Asian	0.25	(1.70)	0.32	(1.69)
Native-American	-1.52	(4.03)	-1.36	(4.04)
Other	-2.05	(3.48)	-2.12	(3.47)
FRPL	-6.25***	(0.82)	-6.20***	(0.82)
IEP	-30.39***	(1.46)	-30.35***	(1.46)
ELL	-32.50***	(2.06)	-32.43***	(2.06)
Female	-1.07	(0.63)	-1.10	(0.63)
Home Index‡	3.83***	(0.21)	3.82***	(0.21)
Parent Education‡	2.16***	(0.33)	2.14***	(0.33)
Days absent‡	-2.96***	(0.33)	-2.96***	(0.33)
Teacher Variables				
Teacher Non-White	-3.01*	(1.32)	-3.39*	(1.34)
Years Teaching‡	0.09	(0.07)	0.06	(0.07)
Year Teaching Squared‡	-0.01	(<0.01[b])	<0.01[c]	(<0.01[c])
Regular/Standard Certification	-1.76	(1.60)	-1.73	(1.62)
Nationally Board Certified	0.61	(1.43)	0.20	(1.39)
Undergraduate Social Science	1.55	(0.93)	1.20	(0.93)
Graduate Non Social Science	-0.64	(1.02)	-0.64	(1.00)
[Ref=No Graduate Degree]				
Graduate Social Science	-0.33	(1.17)	-0.71	(1.20)
Only Teaches Social Studies	2.53*	(1.15)	2.18	(1.16)
Number Students in Sample	0.10	(0.09)	0.08	(0.09)
School Variables				
City [Ref=Suburb]	1.97	(1.19)	1.94	(1.18)
Rural	-1.61	(1.15)	-1.15	(1.16)
Northeast [Ref=West]	0.66	(1.46)	0.60	(1.45)
Southeast	-0.68	(1.29)	-0.86	(1.29)
Central	1.47	(1.43)	1.59	(1.44)
Charter [Ref=Public]	12.62***	(3.58)	13.17***	(3.63)
Private	-0.17	(2.16)	0.61	(2.16)
Student Absences‡	-2.11**	(0.67)	-2.07**	(0.67)
% Free Reduced Priced Lunch‡	-0.15***	(0.03)	-0.14***	(0.03)
% Minority‡	-0.03	(0.03)	-0.03	(0.03)
Number Teachers in Sample‡	1.24**	(0.39)	1.19**	(0.39)
Total Professional Learning‡			0.47**	(0.15)
Variance Components				
Level-1 variance	546.73		546.50	
Level-2 variance	32.07a		31.07a	
Level-3 variance	31.35***		31.30***	
FRLP	47.49		46.75	
IEP	214.13***		210.74***	
ELL	131.86		132.58	
Intraclass correlation coefficient				
Level-2	0.05		0.05	
Level-3	0.05		0.05	
Level-2 Reliability	0.19		0.18	
Level 3 Reliability	0.40		0.40	
% Additional Variance Explained				
% Variance Explained Level-1	0.1%		0.0%	
% Variance Explained Level-2	-5.8%		3.1%	
% Variance Explained Level-3	37.7%		0.2%	

Note. Standard errors are in parentheses, % Variance explained calculated in comparison to previous model. FRPL=Student eligible for free or reduced priced lunch, IEP=Student has an IEP, ELL=Student is an English Language Learner

[a] Significance of Level-2 Variance could not be computed because there was not sufficient degrees of freedom

[b] Actual values are (.0037) [c] Actual values are -.0045 (.0036)

‡ Grand Mean Centered **p<.05 **p<.01 ***p<.001*

Model building. The initial model (Model 1) included no predictors but did include random effects for the teacher-level intercept and the school-level intercept. This allowed the estimated mean student score on the NAEP civics assessment to vary across teacher within schools and across schools. There was significant variance between schools ($\chi^2(627) = 1436.72, p < .001$) and between teachers within schools ($\chi^2(448) = 1456.30,\ p < .001$). The intraclass correlation coefficient was .20 for schools and .11 for teachers within schools. This means that approximately 20% of the variation in student achievement was between schools and about 10% of the variance was among teachers within the same schools.

Student demographics were then entered as a block into the model (Model 2). In Model 3, the relationships between achievement and student demographics were allowed to vary randomly across teachers. The variance of the random slopes was significant for three student demographic variables: eligibility for free or reduced priced lunch ($\chi^2(666) = 787.73, p < .001$), ELL status ($\chi^2(252) = 323.41, p < .01$), and IEP status ($\chi^2(433) = 667.89, p < .001$). Additionally, the reliability of the random slopes were all greater than $\lambda > .1$ which is considered adequate. This suggests that the relationships between these student demographic variables and student achievement varied across teachers within the same school. The chi-square difference tests indicated that including level-2 random effects for these three student demographics significantly improved the fit of the model ($\chi^2(9) = 68.18, p < .001$)[13]. As a result, the random-effects for these level-2 slopes were retained in all subsequent models.

Some of the random slopes and intercepts were moderately correlated with one another. The random slope for free lunch status was moderately negative correlated ($r = -.36$) with the teacher-level intercept. This indicated that there were larger gaps in achievement between free

[13] Chi-Square difference statistic for first plausible values provided as an example. The chi-square difference was significant at $p<.001$ for all five plausible values

lunch and non-free lunch students for teachers who had lower overall student achievement on NAEP. Additionally, the random slopes for ELL and free lunch status were also moderately negatively correlated ($r = -.31$). This meant that teachers who had greater differences in achievement between ELL and non-ELL students tended to have smaller differences in achievement between free lunch and non-free lunch students.

Teacher and school level predictors were added in Models 4 and 5. The teacher-level predictors were also examined to determine if any of the slopes varied randomly across schools. However, none of the slopes for the teacher-level predictors varied significantly across schools, so all teacher-level slopes were fixed. In Model 6, teacher professional learning was added to the model.

Estimated fixed effects. Teacher professional learning was significantly related to student achievement on NAEP civics $\left(\hat{\beta} = .47, t = 3.23\ p < .01\right)$. For each additional professional learning activity teachers participated in over the grand mean, their students' average score on NAEP increased by .47 points. Additionally, teacher professional learning explained an additional 3.1% of the residual variation in the teacher-level intercept. However, the size of this relationship was small. The standardized effect size was .045 standard deviations, which is generally considered a very small effect size in education (J. Cohen, 1992; C. Hill et al., 2007). In percentile terms, a .045 standard deviation change is the difference between scoring in the 50th percentile and scoring in the 52nd percentile. Comparing a bottom quartile participant in professional learning to a top quartile teacher only slightly increases the size of the difference. The difference in average student NAEP scale scores between a bottom quartile and a top quartile teacher was 2.37 points. This is roughly the difference between scoring in the 50th and 53rd percentile on NAEP civics.

Compared to teacher professional learning, student demographics were much stronger predictors of student achievement on NAEP civics. Student demographics explained 27.3% of the student-level variation in NAEP civics achievement. There were significant differences in student achievement by race. Even after controlling for socio-economic factors, African-American ($\hat{\pi} = -11.95, t = -10.58\ p < .001$) and Latino students ($\hat{\pi} = -3.40, t = 3.14\ p < .01$) scored significantly lower than White students. African-American students scored .35 standard deviations lower than Whites and Latino students scored .10 standard deviations lower than Whites. There was no significant difference between White students and Asian and Native-American students.

Socio-economic factors were also related to student achievement on NAEP civics. Students who were eligible for free or reduce priced lunch scored significantly lower on the test ($\hat{\pi} = -6.20, t = -7.57\ p < .001$). The difference between students based on free lunch status was .19 standard deviations. In addition, parental education ($\hat{\pi} = 2.14, t = 6.43, p < .001$) and home educational resources ($\hat{\pi} = 3.82, t = 17.83\ , p < .001$) were both significantly related to student achievement on NAEP. A one unit increase in parental education, such as increasing from high school degree to some college, increased student scored by .064 standard deviations and a one unit increase in home educational resources was related to a .11 standard deviation increase in student scores.

Other student demographic variables were significant predictors. English Language Learners ($\hat{\pi} = -32.43, t = -15.77\ , p < .001$), and students with an IEP ($\hat{\pi} = -30.35, t = -20.78\ , p < .001$) also had significantly lower scores on the NAEP civics assessment. These differences were very large: ELL students scored .96 standard deviations and students with an IEP scored .91 standard deviations lower. The numbers of days students were absent from school

was significantly negatively related to student achievement on NAEP ($\hat{\pi} = -2.96, t = -9.07, p < .001$). A one-unit increase in days absent was associated with a .088 standard deviation drop in student scores. Girls scored slightly lower than boys ($\hat{\pi} = -1.10, t = -1.10, p = .08$) though the difference was only marginally significant and the effect size, at .033 standard deviations, was small.

At the teacher-level, after controlling for other student and school characteristics, the only significant predictor aside from professional learning was teacher minority status, which was negatively related to student achievement $\left(\hat{\beta} = -3.39, t = -2.52, p < .05\right)$. Students of minority teachers scored .10 standard deviations lower on NAEP civics than students of White teachers. No other teacher characteristics, such as years of experience or having a Master's degree, were significant predictors of student achievement on NAEP civics. Teacher predictors did not explain any additional variances in the level-2 intercept; the variance actually slightly increased by 1.4% (Model 4). This may because most of the teacher predictors were not statistically significant. Introducing non-significant predictors into the model can cause the residual variance to slightly increase (Raudenbush & Bryk, 2002). In this situation, adding mostly non-significant predictor may have increased the residual variance in the level-2 intercept. However, these teacher characteristics were also of theoretical importance and so removing would have affected the interpretation of the model.

School sector, urbanicity, poverty level, and attendance rate were all significant predictors of student achievement after controlling for other characteristics. Students attending charter schools performed significantly higher on NAEP civics than students attending public schools $\left(\hat{\beta} = 13.17, t = 3.63, p < .001\right)$. On average students in charter schools scored .39 standard deviations higher than students in public schools. However, there were fewer than 10

charter schools represented in the sample so this finding may not be robust. The percentage of students who were eligible for free or reduced priced lunch in the school was negatively related to student achievement: A 10% increase in the number of students eligible for free or reduced priced lunch was associated with a .047 standard deviations decrease in student achievement ($\hat{\gamma} = -.14\ t = -4.68, p < .001$). Additionally, the percent of students absent on an average day ($\hat{\gamma} = -2.07\ t = -3.10, p = .002$) was significantly negatively associated with student achievement on the NAEP civics assessment. A one unit increase in the student absences was associated with a .062 standard deviation decrease in student achievement. Finally, the number of teachers in the sample was positively related to student achievement on NAEP, although the effect size was small (.036 standard deviations). In total, school level characteristics explained an additional 37.7% of the residual variance at the school level (Model 5).

RQ3. How does the form of professional development (traditional or communities of practice) relate to teachers' classroom practices and student achievement on the NAEP civics assessment?

RQ3a. How does the form of professional development relate to teachers' use of interactive civics activities in the classroom? Participation in both traditional and communities of practice professional learning activities was positively related to a teachers' use of interactive civics activities (Table 20). For each traditional professional development activity teachers participated in their use of interactive civics activities increased by slightly less than one unit $\left(\hat{\beta} = .84, t = 8.70, p < .001\right)$. Each additional communities of practice activity a teacher participated in increased interactive civics use by a slightly smaller amount $\left(\hat{\beta} = .65, t = 10.46, p < .001\right)$. The standardized effect sizes were similar across the two forms of professional development. A one standard deviation increase in traditional professional

development had an effect size of .26 standard deviations, whereas the same change in communities of practice increased interactive civics use by .29 standard deviations.

Table 20

HLM Results: Form of Professional Learning and Use of Interactive Civics

Parameter	Model 5		Model 6	
	Coeff	(SE)	Coeff	(SE)
Intercept	10.93***	(0.73)	10.55***	(0.72)
Teacher Variables				
Teacher Non-White	<0.01[a]	(0.36)	-0.02	(0.36)
Year Teaching‡	-0.06***	(0.02)	-0.08**	(0.02)
Year Teaching Squared‡	<0.01[b]	(<0.01[b])	<0.01[d]	(<0.01[d])
Regular/Standard Certification	-0.70	(0.51)	-0.46	(0.50)
Nationally Board Certified	1.53**	(0.45)	1.42**	(0.43)
Undergraduate Social Science	0.62*	(0.29)	0.50	(0.29)
Graduate Non Social Science	0.51	(0.30)	0.38	(0.29)
[Ref=No Graduate Degree]				
Graduate Social Science	1.07**	(0.39)	1.05**	(0.37)
Only Teaches Social Studies	-0.12	(0.35)	-0.06	(0.34)
Number Students in Sample	-0.01	(0.03)	-0.01	(0.03)
School Variables				
City [Ref=Suburb]	0.51	(0.42)	0.53	(0.41)
Rural	-0.65	(0.43)	-0.53	(0.42)
Northeast [Ref=West]	0.46	(0.47)	0.78	(0.46)
Southeast	0.74	(0.40)	0.88*	(0.39)
Central	0.05	(0.52)	0.30	(0.50)
Charter [Ref=Public]	0.35	(1.54)	-0.07	(1.56)
Private	-0.78	(0.74)	-0.48	(0.72)
Student Absences‡	-0.39	(0.25)	-0.46	(0.24)
% Free Reduced Priced Lunch‡	-0.01	(0.01)	-0.01	(0.01)
% Minority‡	<0.01[c]	(0.01)	<0.01[e]	(0.01)
Number Teachers in Sample‡	-0.33*	(0.13)	-0.27*	(0.12)
Professional Learning				
Traditional	0.86***	(0.10)		
Communities of Practice			0.67***	(0.06)
Variance Components				
Level-1 variance	16.68		16.51	
Level-2 variance	3.41***		3.09***	
Intraclass correlation coefficient	0.17		0.16	
Level-2 Reliability	0.31		0.30	
% Additional Variance Explained				
% Variance Explained Level-1	7.2%		8.1%	
% Variance Explained Level-2	1.7%		10.9%	

Note. % Variance explained calculated in comparison to Model 4

[a] actual value is 0.00041 [b] actual values are 0.00042 (0.0097) [c] actual value is -0.0013 [d] actual values are 0.00068 (0.00098)

[e] actual value is -0.00049

‡ Grand Mean Centered *1 *p<.05 **p<.01 ***p<.001*

There was no significant difference between the two forms of professional learning in terms of their relationship with interactive civics use. A z-test was used to assess whether the mean difference between the estimated coefficients for traditional and communities of practice

professional learning was statistically significant different from zero. It found that the mean difference was not statistically significant ($z = 1.64, p = .104$). This indicates that there was no significant difference between the predictors in their relationship with teachers' use of interactive civics activities.

RQ3b. How does the form of content-area professional development (traditional or communities of practice) relate to students' civic achievement on NAEP, controlling for school, teacher, and student characteristics? Both traditional and communities of practice forms of teacher professional development were associated with significant increases in students' scores on the NAEP civics assessment (Table 21). For each additional traditional professional development activity that teachers engaged in their students' scores increased by nine-tenths of a point $(\hat{\beta} = .90, t = 2.88,\ p < .01)$. Each additional communities of practice professional development activity improved student achievement slightly less, increasing student scores by about two-thirds of a point $(\hat{\beta} = .62, t = 2.82,\ p < .01)$.

However, the relationships between teacher professional learning and student achievement on NAEP did not differ significantly based on the form of professional development. The difference between the fixed effects for traditional and communities of practice professional learning was not significantly different from zero ($z = .74, p = .303$). Furthermore, the standardized effect size for both forms of professional development was almost exactly the same. For both forms of professional development a one standard deviation increase was associated with a .039 standard deviation increase in student NAEP achievement; approximately the difference between scoring in the 50th percentile and scoring in the 52nd percentile. This suggests that traditional and communities of practice professional learning do not have different relationships with student achievement on NAEP civics.

Table 21

HLM Results: Professional Learning and Student Achievement on NAEP Civics

Parameter	Model 7		Model 8	
	Coeff	(SE)	Coeff	(SE)
Intercept	160.71***	(2.13)	160.38***	(2.11)
Student Variables				
Black [Ref=White]	-11.95***	(1.13)	-11.91***	(1.13)
Hispanic	-3.41**	(1.09)	-3.39**	(1.08)
Asian	0.30	(1.70)	0.32	(1.69)
Native-American	-1.35	(4.05)	-1.44	(4.04)
Other	-2.10	(3.47)	-2.11	(3.48)
FRPL	-6.22***	(0.82)	-6.21***	(0.82)
IEP	-30.35***	(1.46)	-30.37***	(1.46)
ELL	-32.46***	(2.06)	-32.45***	(2.06)
Female	-1.09	(0.63)	-1.10	(0.63)
Home Index‡	3.83***	(0.21)	3.82***	(0.21)
Parent Education‡	2.14***	(0.33)	2.15***	(0.33)
Days absent‡	-2.96***	(0.33)	-2.96***	(0.33)
Teacher Variables				
Teacher Non-White	-3.25*	(1.33)	-3.34*	(1.34)
Year Teaching ‡	0.07	(0.07)	0.06	(0.07)
Year Teaching Squared‡	<0.01[b]	(<0.01[b])	<0.01[c]	(<0.01[c])
Regular/Standard Certification	-1.81	(1.61)	-1.69	(1.61)
Nationally Board Certified	0.41	(1.43)	0.22	(1.37)
Undergraduate Social Science	1.34	(0.93)	1.23	(0.93)
Graduate Non Social Science	-0.53	(1.01)	-0.72	(1.00)
[Ref=No Graduate Degree]				
Graduate Social Science	-0.61	(1.21)	-0.63	(1.19)
Only Teaches Social Studies	2.14	(1.15)	2.34*	(1.16)
Number Students in Sample	0.08	(0.09)	0.09	(0.09)
School Variables				
City [Ref=Suburb]	1.95	(1.18)	1.95	(1.18)
Rural	-1.31	(1.16)	-1.21	(1.15)
Northeast [Ref=West]	0.41	(1.46)	0.76	(1.44)
Southeast	-0.83	(1.29)	-0.81	(1.29)
Central	1.56	(1.45)	1.56	(1.43)
Charter [Ref=Public]	13.68***	(3.67)	12.61***	(3.61)
Private	0.19	(2.15)	0.61	(2.16)
Student Absences‡	-2.07**	(0.67)	-2.09**	(0.67)
% Free Reduced Priced Lunch‡	-0.14***	(0.03)	-0.14***	(0.03)
% Minority‡	-0.03	(0.03)	-0.03	(0.03)
Number Teachers in Sample‡	1.17**	(0.39)	1.23**	(0.39)
Professional Learning				
Traditional	0.90**	(0.31)		
Communities of Practice			0.62**	(0.22)
Variance Components				
Level-1 variance	546.51		546.58	
Level-2 variance	30.65[a]		31.75[a]	
Level-3 variance	31.29***		31.40***	
FRLP	47.11		46.74	
IEP	211.48***		211.27***	
ELL	132.28		132.41	
Intraclass correlation coefficient				
Level-2	0.05		0.05	
Level-3	0.05		0.05	
Level-2 Reliability	0.18		0.19	
Level 3 Reliability	0.40		0.40	
% Additional Variance Explained				
% Variance Explained Level-1	0.0%		0.0%	
% Variance Explained Level-2	4.4%		1.0%	
% Variance Explained Level-3	0.2%		-0.2%	

Note. % Variance explained calculated in comparison to Model 5. FRPL=Student eligible for free or reduced priced lunch, IEP=Student has an IEP, ELL=Student is an English Language Learner [a] Significance of Level-2 Variance could not be computed because there was not sufficient degrees of freedom [b] Actual values are -0.0049 (0.0036) [c] Actual values are -0.0045 (0.0036)

‡ Grand Mean Centered **p<.05 **p<.01 ***p<.001*

RQ4. How are teachers' classroom practices related to student outcomes, and what role does the amount and type of professional development play in that relationship?

RQ4a. Do students of teachers who use interactive civics activities in the classroom have higher levels of civic achievement on NAEP, controlling for school, teacher, and school characteristics? Teachers' use of interactive civics activities was a significant positive predictor of student achievement on the NAEP civics assessment $\left(\hat{\beta} = .21, \mathrm{t} = 2.41,\ p < .05\right)$ (Table 22). However, the magnitude of the relationship was small. For every one-unit increase in interactive civics use students' score on NAEP increased by an average of one-fifth of a point. In standardized terms, a one standard deviation increase in interactive civics use increased student achievement on NAEP by .03 standard deviations, which is a very small effect size. This is the equivalent of increasing from the 50th percentile to the 51st percentile in student achievement.

The small magnitude of this relationship is further illustrated by thinking about the change in terms of the 44 point difference between students scoring at the Basic achievement level and students scoring at the Proficient achievement level. Conducting an additional high intensity activity such as a debate or a mock trial, would only increase student achievement by an average of six-tenths of a point. Thus, a difference of six-tenths of a point is the equivalent of reducing the gap between by only 1.44%. Similarly, the difference in expected student achievement on NAEP between students who had a bottom quartile teacher in terms interactive civics use and those with a top quartile teacher was 1.27 points. A difference of this magnitude would only reduce the difference between scoring Basic and Proficient by 2.89%.

However, one of the limitations of comparing teachers by standard deviation or quartile is that it is not based on any external criterion. For example, the standard deviation for interactive civics is 4.82 which is approximately the equivalent of one high and one medium intensity

activity. This difference, though meaningful, does not fully capture the spectrum of interactive activities use among teachers. By contrast, if we compare non-users of interactive civics to "high" users, teachers who used five high intensity interactive civics activities such as conducting a debate, mock trial, or community project, the effect size of interactive civics activities was much larger. A student whose teachers used five high intensity interactive civics activities would be expected to score .095 standard deviations higher on NAEP civics than a student whose teacher used no activities. Such a difference would reduce the gap between the Basic and Proficient achievement levels by 7.19%.

There was also some evidence that the relationship between interactive civics use and student achievement was non-linear. The quadratic term for interactive civics was negative and statistically significant$\left(\hat{\beta} = -.03, t = -2.03,\ p < .05\right)$ indicating that after a certain point using additional interactive civics activities in the classroom was associated with decreased student achievement (Table 22). The maxima of the quadratic function was at 14.38 on the interactive civics scale (Figure 4). This is the equivalent of approximately five high impact activities. This finding reflected what subjects reported during the cognitive interviews about the relationship between interactive civics and student learning: that although interactive activities are generally beneficial using them too frequently may actually reduce student achievement.

Table 22

HLM Results: Interactive Civics Use and Student Achievement on NAEP Civics

Parameter	Model 9		Model 10	
	Coeff	(SE)	Coeff	(SE)
Intercept	160.16***	(2.12)	159.94***	(2.12)
Student Variables				
Black [Ref=White]	-11.87***	(1.13)	-11.82***	(1.13)
Hispanic	-3.33**	(1.08)	-3.31**	(1.08)
Asian	0.24	(1.70)	0.23	(1.70)
Native-American	-1.61	(3.97)	-1.48	(3.97)
Other	-2.00	(3.48)	-2.05	(3.48)
FRPL	-6.25***	(0.82)	-6.25***	(0.82)
IEP	-30.36***	(1.46)	-30.37***	(1.46)
ELL	-32.50***	(2.05	-32.47***	(2.05)
Female	-1.09	(0.63)	-1.08	(0.63)
Home Index‡	3.82***	(0.21)	3.82***	(0.21)
Parent Education‡	2.16***	(0.33)	2.15***	(0.33)
Days absent‡	-2.95***	(0.33)	-2.96***	(0.33)
Teacher Variables				
Teacher Non-White	-3.06*	(1.31)	-3.13*	(1.30)
Year Teaching‡	0.10	(0.07)	0.10	(0.07)
Year Teaching Squared‡	-0.01	(<0.01[b])	<0.01[d]	(<0.01[d])
Regular/Standard Certification	-1.63	(1.60)	-1.75	(1.60)
Nationally Board Certified	0.19	(1.45)	0.28	(1.43)
Undergraduate Social Science	1.34	(0.95)	1.37	(0.94)
Graduate Non Social Science	-0.68	(1.01)	-0.72	(1.01)
[Ref=No Graduate Degree]				
Graduate Social Science	-0.55	(1.19)	-0.46	(1.20)
Only Teaches Social Studies	2.48*	(1.15)	2.28	(1.17)
Number Students in Sample	0.10	(0.09)	0.10	(0.09)
School Variables				
City [Ref=Suburb]	1.96	(1.19)	1.98	(1.19)
Rural	-1.40	(1.15)	-1.42	(1.15)
Northeast [Ref=West]	0.60	(1.45)	0.64	(1.45)
Southeast	-0.74	(1.29)	-0.67	(1.29)
Central	1.49	(1.43)	1.57	(1.43)
Charter [Ref=Public]	12.74***	(3.50)	12.29***	(3.49)
Private	<0.01[c]	(2.13)	0.11	(2.14)
Student Absences‡	-2.04**	(0.67)	-2.03**	(0.67)
% Free Reduced Priced Lunch‡	-0.14***	(0.03)	-0.14***	(0.03)
% Minority‡	-0.03	(0.03)	-0.03	(0.03)
Number Teachers in Sample‡	1.32***	(0.39)	1.31***	(0.39)
Interactive Civics	0.21*	(0.09)	0.20*	(0.09)
Interactive Civics Squared			-0.03*	(0.01)
Variance Components				
Level-1 variance	546.54		546.52	
Level-2 variance	31.79[a]		31.41[a]	
Level-3 variance	30.58**		31.03**	
FRLP	46.74		46.39	
IEP	213.07**		212.63**	
ELL	131.88		132.60	
Intraclass correlation coefficient				
Level-2	0.05		0.05	
Level-3	0.05		0.05	
Level-2 Reliability	0.19		0.19	
Level 3 Reliability	0.39		0.40	
% Additional Variance Explained				
% Variance Explained Level-1	0.0%		0.0%	
% Variance Explained Level-2	0.9%		2.1%	
% Variance Explained Level-3	2.5%		1.0%	

Note, % Variance explained calculated in comparison to Model 5. FRPL=Student eligible for free or reduced priced lunch, IEP=Student has an IEP, ELL=Student is an English Language Learner [a] Significance of Level-2 Variance could not be computed because there was not sufficient degrees of freedom [b] Actual value is 0.0036 [c] Actual values is -0.0015 [d] Actual values are -0.0050 (.0036)

‡ Grand Mean Centered †*p<.1 *p<.05 **p<.01 **p<.001*

Figure 4. Predicted NAEP Civics Score by Teachers' Use of Interactive Civics with Quadratic Term

RQ4b. Does teachers' engagement in more overall professional learning activities moderate the relationship between teachers' use of interactive civics activities in the classroom and student achievement on the NAEP civics assessment, controlling for school, teacher, and student characteristics? Teachers' professional learning activities did not significantly moderate the relationship between interactive civics and student achievement on NAEP civics (Table 23). The interaction term for professional learning and interactive civics use was not statistically significant$\left(\hat{\beta} = -0.04\,, \mathrm{t} = -1.370, \mathrm{p} = .172\right)$. The effect size, though small, was not negligible: a one standard deviation increase in teacher professional learning reduced the slope of interactive civics by .12 points. This meant that for teachers whose professional development was one standard deviation below the mean the slope for interactive civics was .24, whereas for teachers who were one standard deviation above the mean the slope was 0 (Figure 5).

These results suggest that the relationship between interactive civics and student achievement may be weaker for teachers who participated at higher levels in professional learning. However, because the interaction term was not statistically significantly the null hypothesis that professional learning does not moderate the slope for the relationship between interactive civics use and student achievement could not be rejected.

Table 23

HLM Results: Interaction between Professional Learning and Interactive Civics Use

	Model 11	
Parameter	Coeff	(SE)
Intercept	160.65***	(2.10)
Student Variables		
Black [Ref=White]	-11.93***	(1.13)
Hispanic	-3.37***	(1.08)
Asian	0.30	(1.69)
Native-American	-1.43	(3.98)
Other	-2.13	(3.47)
FRPL	-6.20***	(0.82)
IEP	-30.31***	(1.46)
ELL	-32.41***	(2.05)
Female	-1.11	(0.63)
Home Index‡	3.82***	(0.21)
Parent Education‡	2.14***	(0.33)
Days absent‡	-2.95***	(0.33)
Teacher Variables		
Teacher Non-White	-3.32*	(1.34)
Year Teaching‡	0.07	(0.07)
Year Teaching Squared‡	<0.01[b]	(<0.01[b])
Regular/Standard Certification	-1.63	(1.60)
Nationally Board Certified	0.07	(1.41)
Undergraduate Social Science	1.15	(0.94)
Graduate Non Social Science	-0.58	(1.01)
[Ref=No Graduate Degree]		
Graduate Social Science	-0.67	(1.22)
Only Teaches Social Studies	2.12	(1.16)
Number Students in Sample	0.08	(0.09)
School Variables		
City [Ref=Suburbs]	1.96	(1.18)
Rural	-1.10	(1.16)
Northeast [Ref=West]	0.50	(1.44)
Southeast	-0.83	(1.29)
Central	1.59	(1.43)
Charter [Ref=Public]	13.00***	(3.57)
Private	0.56	(2.14)
Student Absences‡	-2.07**	(0.67)
% Free Reduced Priced Lunch‡	-0.14***	(0.03)
% Minority‡	-0.03	(0.03)
Number Teachers in Sample‡	1.24**	(0.39)
Total Professional Learning	0.44**	(0.15)
Interactive Civics	0.12	(0.09)
Interactive Civics X Total Professional Learning	-0.04	(0.03)
Variance Components		
Level-1 variance	546.37	
Level-2 variance	31.11[a]	
Level-3 variance	30.58	
FRLP	46.40	
IEP	209.99***	
ELL	131.96	
Intraclass correlation coefficient		
Level-2	0.05	
Level-3	0.05	
Level-2 Reliability	0.18	
Level 3 Reliability	0.39	
% Additional Variance Explained		
% Variance Explained Level-1	0.1%	
% Variance Explained Level-2	3.0%	
% Variance Explained Level-3	2.5%	

Note. % Variance explained calculated in comparison to Model 5. FRPL=Student eligible for free or reduced priced lunch, IEP=Student has an IEP, ELL=Student is an English Language Learner [a] Significance of Level-2 Variance could not be computed because there was not sufficient degrees of freedom [b] Actual values are -.004 (.004)

‡ Grand Mean Centered *$p<.05$ **$p<.01$

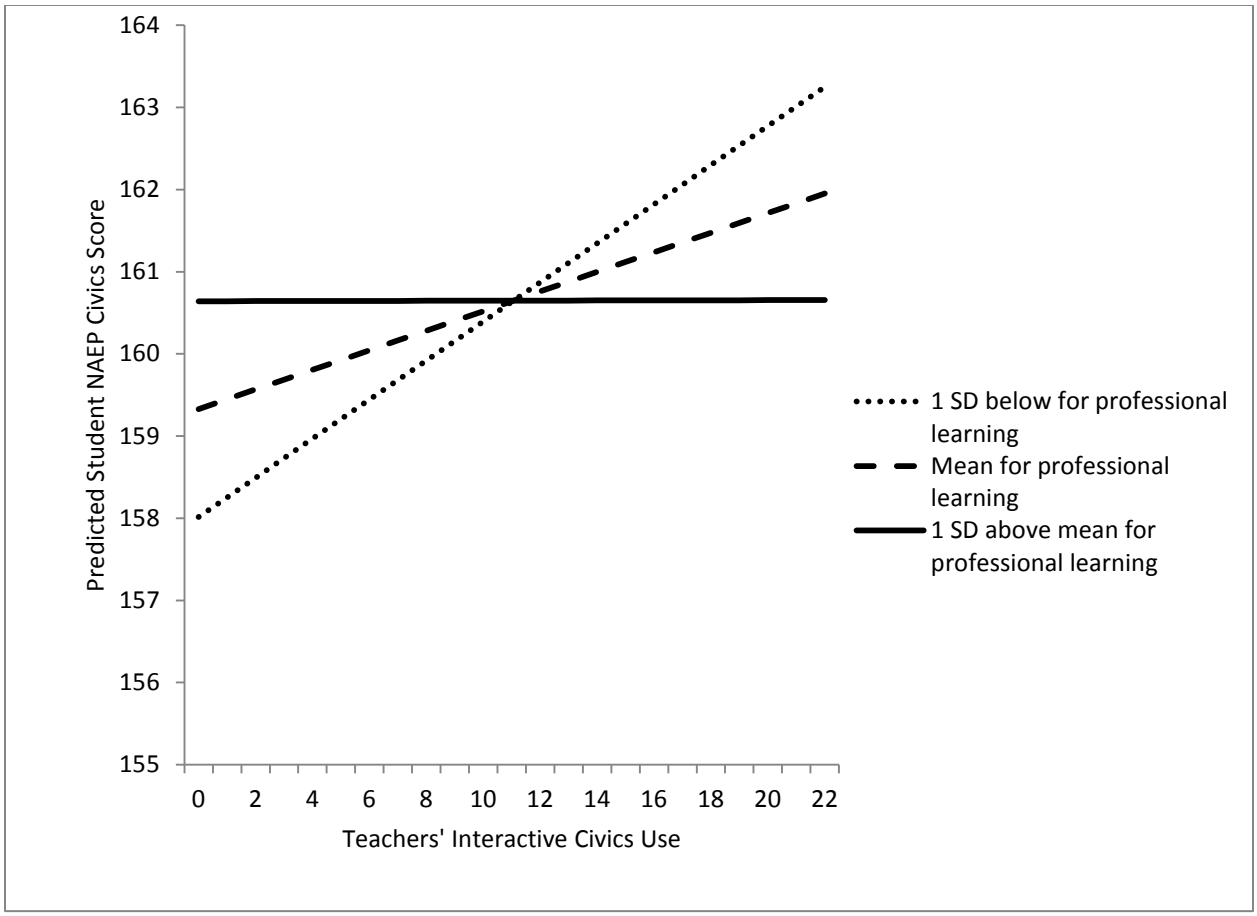

Figure 5. Predicted NAEP Civics Score by Teachers Interactive Use and Professional Learning with Interaction Term

RQ4c. Does the form of professional learning moderate the relationship between teachers' use of interactive civics activities in the classroom and student achievement on the NAEP civics assessment, controlling for school, teacher, and student characteristics?

Participation in traditional professional development activities negatively moderated the slope between teachers' interactive civics usage and students' achievement on NAEP civics, though the relationship was only marginally statistically significant $\left(\hat{\beta} = -.097, \mathrm{t} = -1.789, p = .074\right)$. A one standard deviation increase in traditional professional learning decreased the slope for interactive civics and student achievement by .14 points. This meant that for teachers who

were one standard deviation below the mean in their participation in traditional professional learning the slope for interactive civics was .29, whereas it was -0.011 for teachers one standard deviation below the mean.

Engagement in communities of practice professional development activities did not significantly moderate the relationship between interactive civics use and student achievement on NAEP civics $\left(\hat{\beta} = -.038\ \mathrm{t} = -.846, p = .400\right)$. A one standard deviation increase in communities of practice professional learning decreased the slope for interactive by .08 points. Teachers who participated in communities of practice one standard deviation below the mean had a slope for interactive civics of .22 points whereas it was -.055 for teachers' one standard deviation above the mean.

There was no significant difference between the two forms of professional learning. A z-test indicated that the mean difference between the interaction terms for traditional professional development and communities of practice was not statistically significantly different from zero $(\mathrm{z} = -.26, p = .385)$. This suggests that there were no significant difference between the types of professional learning in how much they moderated the relationship between interactive civics use and student achievement on NAEP civics.

Table 24

HLM Results: Interaction between Form of Professional Learning and Interactive Civics

	Model 12		Model 13	
Parameter	Coeff	(SE)	Coeff	(SE)
Intercept	160.73***	(2.12)	160.40***	(2.10)
Student Variables				
Black [Ref=White]	-11.94***	(1.13)	-11.90***	(1.13)
Hispanic	-3.37**	(1.08)	-3.35**	(1.08)
Asian	0.27	(1.70)	0.29	(1.69)
Native-American	-1.41	(3.96)	-1.51	(3.99)
Other	-2.11	(3.47)	-2.10	(3.48)
FRPL	-6.22***	(0.82)	-6.21***	(0.82)
IEP	-30.29***	(1.46)	-30.34***	(1.46)
ELL	-32.42***	(2.06)	-32.44***	(2.05)
Female	-1.10	(0.63)	-1.11	(0.63)
Home Index‡	3.82***	(0.21)	3.82***	(0.21)
Parent Education‡	2.15***	(0.33)	2.15***	(0.33)
Days absent‡	-2.94***	(0.33)	-2.95***	(0.33)
Teacher Variables				
Teacher Non-White	-3.16*	(1.33)	-3.30*	(1.34)
Year Teaching Civics‡	0.08	(0.07)	0.07	(0.07)
Year Teaching Civics Squared‡	<0.01[b]	(<0.01[b])	<0.01[c]	(<0.01[c])
Regular/Standard Certification	-1.61	1.61)	-1.62	(1.60)
Nationally Board Certified	0.14	1.43)	0.05	(1.40)
Undergraduate Social Science	1.18	0.95)	1.18	(0.95)
Graduate Non Social Science	-0.52	1.01)	-0.67	(1.01)
[Ref=No Graduate Degree]				
Graduate Social Science	-0.60	(1.22)	-0.66	(1.21)
Only Teaches Social Studies	2.06	(1.15)	2.30*	(1.16)
# Students in Sample	0.09	(0.09)	0.09	(0.09)
School Variables				
City [Ref=Suburb]	1.95	(1.18)	1.96	(1.18)
Rural	-1.25	(1.16)	-1.13	(1.15)
Northeast [Ref=West]	0.34	(1.46)	0.66	(1.43)
Southeast	-0.83	(1.28)	-0.79	(1.29)
Central	1.54	(1.44)	1.58	(1.43)
Charter [Ref=Public]	13.36***	(3.59)	12.61***	(3.57)
Private	0.15	(2.13)	0.59	(2.13)
Student Absences‡	-2.06**	(0.66)	-2.06**	(0.67)
% Free Reduced Priced Lunch‡	-0.14***	(0.03)	-0.14***	(0.03)
% Minority‡	-0.03	(0.03)	-0.03	(0.03)
# Teachers in Sample‡	1.24**	(0.39)	1.27**	(0.39)
Traditional Professional Learning	0.79*	(0.31)		
Communities of Practice Professional Learning			0.56*	(0.23)
Interactive Civics	0.15	(0.09)	0.14	(0.09)
Interactive Civics X Traditional Professional Learning	-0.10	(0.05)		
Interactive Civics X Communities of Practice Professional Learning			-0.04	(0.04)
Variance Components				
Level-1 variance	546.43		546.42	
Level-2 variance	30.50[a]		31.84[a]	
Level-3 variance	30.42***		30.74***	
FRLP	46.51		46.36	
IEP	210.46***		210.45***	
ELL	131.20		132.06	
Intraclass correlation coefficient				
Level-2	0.05		0.05	
Level-3	0.05		0.05	
Level-2 Reliability	0.18		0.19	
Level 3 Reliability	0.39		0.40	
% Additional Variance Explained				
% Variance Explained Level-1	0.1%		0.1%	
% Variance Explained Level-2	4.9%		0.7%	
% Variance Explained Level-3	3.0%		1.9%	

Note. % Variance explained calculated in comparison to Model 6. FRPL=Student eligible for free or reduced priced lunch, IEP=Student has an IEP, ELL=Student is an English Language Learner [a] Significance of Level-2 Variance could not be computed because there was not sufficient degrees of freedom [b] Actual values are -.0045 (.0036) [c] Actual values are -.0044 (.0036)

‡ Grand Mean Centered *$p<.05$ **$p<.01$ ***$p<.001$

Summary

The goal of this study was to better understand the relationships between teacher professional learning, teacher use of interactive civics activities and student achievement on the NAEP civics assessment. The study found that teacher engagement in professional learning was positively related to their use of interactive civics activities. Moreover, both traditional and communities of practice professional learning were associated with increases in teachers' use of interactive civics activities. The effect sizes of these relationships were fairly robust ranging from .25 to .31 standard deviations.

However, the relationships between teacher professional learning and instructional practices and student achievement on NAEP, though positive and statistically significant, were not very large. The effect sizes for the relationships between these predictors and student achievement on NAEP were relatively small ranging from .03 to .045 standard deviations. This suggests that teacher professional learning and use of interactive civics activities may not be strongly related to students' civic skills and knowledge; at least by conventional measures of effect size. Additionally, teacher professional learning did not significantly moderate the relationship between interactive civics use and student achievement on NAEP civics.

Chapter 5

There is growing concern that schools are not doing an adequate job preparing the next generation of democratic citizens. In 2014, for example, only 21.5% of eligible voters aged 18-29 voted in the midterm elections (CIRCLE, 2014). Concerns about the low levels of youth civic engagement coupled with ever-growing gaps along racial and socio-economic lines have led some to call for rethinking approaches to civic education (Levine, 2013; Levinson, 2012). This dissertation explored two critical aspects of this rethinking process: teacher professional development and classroom instructional practices. Although the relationships between these variables have been explored in other academic subjects (D. Cohen & Hill, 2000; Desimone, Smith, & Phillips, 2013; Wenglinsky, 2001) few previous studies have examined these factors within the context of civic education As a result, it is unclear how best to approach these issues with the goal of preparing students for the rights and responsibilities of democratic citizenship.

This chapter summarizes and builds upon the research findings described in the previous chapter. First, it will describe the main findings of the study and situate them within the broader literature on teacher professional development and classroom practices. Second, it will describe how these findings might be utilized by policymakers and practitioners to improve students' civic education outcomes. Finally, it will list some of the limitations of this study and propose some avenues for future research.

Findings

Teachers Participated at High Rates in Professional Learning

Teachers in the study reported participating in numerous content-related professional development activities. On average, teachers participated in almost six professional development activities over the past two years. Certain professional learning experiences were particularly

common among teachers. For example, nearly three-quarters (74.8%) of teachers completed a workshop and 77.0% of teachers in the sample independently read material on social studies instruction. Even activities that required higher levels of commitment and initiative were fairly common among teachers in the study. More than half of teachers in the study (53.3%) reported having attended a conference or professional meeting, 44.3% had conducted individual or collaborative research, and 50.6% had participated in a committee or task force.

Additionally, substantial partitions of teachers in the study participated in the types of communities of practice professional learning activities advocated in the literature (Desimone, 2009; Little, 2006). On average, teachers engaged in a little more than three communities of practice professional development activities. For example, almost half of all teachers in the study received or participated in mentoring, coaching, and peer observation (48.7%) and similar numbers of teachers participated in regular discussion groups (48.1%) and worked in instructional teams (42.5%). Additionally, more than one-quarter of teachers (25.8%) participated in a teacher network specifically devoted to teaching social studies.

The high level of participation professional learning in this study is surprising given that many researchers have been critical of the professional learning opportunities available to U.S. teachers (Borko, 2004; Hill H. , Fixing teacher professional development, 2009; Little J. W., Professional community and professional development in the learning-centered school, 2006; Wilson & Berne, 1999). Researchers studying civics education have expressed similar concerns about the quality of professional learning for civics teachers (Campaign for the Civic Mission of Schools, 2011; CIRCLE, 2013; Hess & Zola, 2012). Hess and Zola (2012), for example, note that although sporadic examples of high quality professional learning in civics do exist: “the

quality and quantity of this professional development must improve dramatically and quickly" (p.2).

On one hand, the high levels of participation in this study seem to offer evidence that professional learning opportunities in civics instruction may be more readily available than previously thought. Clearly, many teachers are taking advantage of opportunities to participate in professional development on civics instruction. On the other hand, participation rates do not provide information about the quality of the professional learning that teachers are experiencing. Teachers may be participating in many different types of activities but not receiving much benefit from them. This question clearly cannot be resolved based on the evidence in this study. Additional survey measures about the quality of teacher professional learning are needed to fully understand the landscape of teacher professional learning in civics.

Teacher Education and Experience was Highly Related to Professional Learning Participation

Teachers' educational background was highly associated with participation in professional learning on social studies instruction. Teachers with undergraduate or graduate degrees in a social science (e.g., history, political science) were significantly more likely to participate in 10 out of the 13 professional learning activities examined in this study. Having a graduate degree in a non-social-science, such as a master's in education, was a negative predictor of attending content-related conferences and participating in workshops.

A number of different factors may account for this relationship. Teachers who majored in a social science may have had a greater underlying interest in the content and may have been more likely to pursue professional learning opportunities along those lines (Scribner, 1999). Additionally, teachers who did not major in their content area may have been more likely to have

to be assigned by their principal to teach an out-of-field subject, for example, a teacher with a background in English teaching social studies (Ingersoll, 2001). These teachers may be less invested in improving their instructional practices than a teacher teaching in their field of study (J.A. Ross, Cousins, Gadalla, & Hannay, 1999).

This study also found that more experienced teachers were more likely to attend a conference, serve on a task force, participate in a workshop, hold a school leadership position, and conduct independent reading. Teachers with more teaching experience often have greater social capital (Leana, 2011) or professional capital (Hargreaves & Fullan, 2012) which may increase their access to professional learning opportunities. With the high attrition rate within teaching (Ingersoll & Smith, 2003), teachers who remain in teaching after the first few years may also have some individual trait or characteristic, such as grit or passion, that makes them more likely to pursue certain professional opportunities. Schools are also notoriously hierarchical institutions and teachers with more experience may have more opportunities to participate in professional learning opportunities because of their seniority.

Additionally, one school characteristic of particular interest was the percent of low-income and minority students within the school. The relationship between school demographics and teacher professional learning are somewhat disputed in the literature. Some have argued that teachers in schools serving low-income and minority populations are less likely to participate in professional learning activities because those schools have fewer resources devoted to teacher professional learning (Greenwald, Hedges, & Laine, 1996; Scribner, 1999). Others have claimed that teachers in large districts, which often serve more diverse student population, have more opportunities to participate in professional learning because there is greater infrastructure for professional development than in smaller districts (J.W. Little, 2006).

The findings of this study appear to side more with the latter position. This study did not find any relationship between student demographics and participation in teacher professional learning, with the exception of independent reading. Teachers in schools with high percentage of low-income and minority students were just as likely to participate in professional learning as teachers in schools with few low-income or minority students.

Professional Learning was Positively Associated with Interactive Civics Use

Teachers who participated in more professional learning activities were more likely to use interactive civics activities in their classrooms. The effect size of this relationship was robust: a one standard deviation increase in professional learning was associated with a predicted .32 standard deviation increase in use of interactive civics activities.

This finding aligns with studies in other content areas that have found that teacher professional learning is associated with interactive teaching practices. For example, Supovitz and Turner (2000) found that the number of hours that teachers participated in professional learning predicted use of inquiry-based teaching practices in science. Using a randomized controlled trial, Gersten et al., (2010) found that first-grade teachers who participated in a teacher study group were more likely to use interactive practices to teach vocabulary. Another experimental study found that fifth grade math teachers who received professional development were more likely to report using reform-based math practices with their students (Dash et al., 2012).

The few studies that have examined this topic in civics have also found a relationship between professional learning and interactive instruction in civics. A study of a nationally representative sample of social studies teachers observed that teachers who had participated in multi-day professional learning programs were more likely to discuss politics with their students (CIRCLE, 2013). Moreover, a randomized control trial of Facing History and Ourselves found

that teachers who participated in professional development were significantly more likely to feel that they had the knowledge to create student-centered learning environments and promote students' deliberation skills (Boulay, et al., 2010).

There are two plausible explanations for why professional learning was associated with greater use of interactive civics. First, professional learning may have increased teachers' exposure to interactive civics practices. The "egg-crate" structure of modern schools often means that teachers have few opportunities to observe and learn from the practices of other teachers (Darling-Hammond et al., 2009; J.W. Little, 2006; O'Day, 2002). Professional learning may therefore have provided teachers with opportunities to learn from the experiences of other teachers and expand their repertoire of teaching practices in the classroom (Ball & Cohen, 1999; Borko, 2004; Cochran-Smith & Lytle, 1999).

Alternatively, participation in professional learning and use of interactive civics activities may both be related to some third unobserved teacher characteristic such as passion, grit, or commitment to teaching civics. Because this study was an observational study, and not a randomized experiment, it is difficult to determine whether it was the professional learning alone that caused changes in teachers' instructional practices. The type of teacher who voluntarily commits to attending conferences on weekends, for example, may also be the type of teacher who spends extra time at night preparing source material for a mock trial. However, very little research has been conducted on how these personal qualities may be related to participation in professional learning or instructional practices (Robertson-Kraft & Duckworth, 2014). As a result, more research is needed in the role these traits may play in influencing teacher behavior.

Significant but Small Relationship between Professional Learning and Achievement

Teacher participation in professional learning was significantly related to student achievement on NAEP civics, but the effect size was marginal. A one standard deviation increase in professional learning was associated with a predicted .045 standard deviation increase in student achievement on NAEP civics. This difference is the equivalent of scoring in the 52nd percentile rather than in the 50th percentile. One way to understand this effect size is to compare it to the point difference between students who scored at the cut-score for the "Basic" achievement level and students who scored at the cut-score for the "Proficient" achievement level. The differences between scoring at the Basic and Proficient levels was 44 points or 1.32 standard deviations (National Center for Education Statistics, 2010). Thus, a one standard deviation increase in teacher professional learning would reduce the difference between the Basic and Proficient achievement levels in civics by only 3.45%.

The effect size for professional learning found in this study is similar to what other observational studies of teacher professional learning have found. For example, Lubienski et al., (2008) found that each additional professional learning activity that teachers participated in increased student achievement on the 4th grade NAEP math assessment by an estimated .3 points (.02 standard deviations).[14] Cohen and Hill (2000) observed that a one standard deviation in participation in curriculum workshops focused on reform-oriented math instruction was associated with a predicted .10 increase in student achievement in math.

However, the relationship between professional learning and student civic achievement was lower than the effect sizes found in experimental and quasi-experimental studies of teacher professional learning. For example, Yoon et al., (2007) in a meta-analysis of experimental and quasi-experimental studies, found that professional learning, on average, increased student achievement by 21 percentile points.

[14] Effect size calculated by author.

One explanation for the small effect size is that teacher professional learning is difficult to measure using the NAEP data. NAEP only provides information about what activities teachers participated in, not how often they participated or the quality of the program. In contrast, experimental and quasi-experimental studies of teacher professional learning are generally studying a specific program (Wayne et al., 2008). Teachers in an experimental or quasi-experimental study are therefore more likely to have a consistent professional development experience than teachers who gave similar responses about professional development on a survey. Consequently, if a form of professional learning does have a positive effect, the effect size would likely be larger in an experimental or quasi-experimental study where the quality of intervention is more controlled.

Additionally, the outcome of interest in this study was likely more distally related to professional learning than it would be an experimental or quasi-experimental study. In an experimental or quasi-experimental study, the student outcome of interest is generally aligned to the intervention (Wayne et al., 2008). For example, if the professional learning is focused on instructional practices for teaching fractions, students should be tested on their knowledge of fractions, not their knowledge of long division. However with the NAEP civics data the relationship between professional learning and the content of the tests was not so easily aligned. Teachers may have engaged in professional learning related to the topics on NAEP civics or they may have participated in professional learning on another social studies topic entirely. Based on the available data, there was no way to know for sure. It is therefore possible that lack of content alignment weakened the relationship between professional learning and student achievement.

Finally, NAEP civics was administered to a large heterogeneous sample of teachers of a nationally representative population of students. Most studies of teacher professional learning are

conducted using small teacher samples at only a few sites (Yoon et al., 2007). This is less costly than large scale studies and it ensures that the intervention will be carefully monitored (Wayne et al., 2008). However, there is often a tradeoff between fidelity and adaptation: An intervention that works only in a specific context is less useful for the field than a study that looks at an intervention across multiple setting (Borko, 2004). However, most recent large-scale experimental studies have failed to find a significant relationship between teacher professional learning and student achievement (Garet, et al., 2008; Garet, et al., 2011; Gersten et al., 2010; Glazerman, et al., 2008; O'Dwyer, Dash, Kramer, Humez, & Russell, 2010). This suggests that the relationship between professional learning and student outcomes may be weakened as teacher samples become larger and more diverse and interventions are applied across multiple contexts. For NAEP civics, the heterogeneity of the teacher sample may have attenuated some of the relationship between professional learning and student achievement.

Yet, rather than view these factors as limitations, they may actually be seen as a strength of the study. This study was grounded in the day-to-day professional learning experiences of teachers which are often quite diverse; ranging from the quick conversation in the hallway to week-long summer institutes (Borko, 2004; Desimone, 2009). Although experimental studies can provide much more accurate information about the effect of a specific program for a specific population, their generalizability to other settings is often circumscribed. By contrast, the teachers in this sample reflected a wide range of teaching contexts and student populations. Their professional learning was similarly diverse; some teachers engaged in numerous professional learning activities while others engaged in almost no professional learning.

Because of this diversity, it is meaningful that teacher professional learning was significantly related to student achievement on NAEP civics even though the effect size was

small. There are a number of plausible explanations for this relationship. Professional learning may increase teachers' skills and knowledge allowing them to be more effective in the classroom (Desimone, 2009). Additionally, professional learning may expand the size of teachers' instructional networks, allowing them to draw on the knowledge and experience of others in planning their instruction (DuFour, 2004; Giles & Hargreaves, 2006; King, 2002). Professional learning may also provide a sense of community for teachers, enhancing their professional identity and commitment to teaching (Hur & Brush, 2009; Noble & Littenberg-Tobias, 2014).

No Differences between Forms of Professional Learning

This study examined two different forms of teacher professional learning: traditional and communities of practice. Traditional professional learning includes activities such as workshops, conferences, or courses where teachers take time away from the classroom for professional learning. Some have criticized this form of professional learning for removing teachers from their school context and presenting information in generalized one-shot forms (Grossman, Wineburg, & Woolworth, 2001; J.W. Little., 1993; Wilson & Berne, 1999). In contrast, in communities of practice forms of professional learning teachers learn through ongoing interactions with their peers (Grossman, Wineburg, & Woolworth, , 2001; J.W. Little., 1993; Wilson & Berne, 1999). In the past decade, communities of practice forms of professional learning such as discussion groups, educator networks, peer coaching, and collaborative research have grown in popularity (Berry, Norton, & Byrd, 2007; Gersten et al., 2010; Guiney, 2001; King, 2002).

This study found that there was no significant difference between traditional and communities of practice professional learning in their relationships to instructional practice and student achievement. Effect sizes were also very similar across the different forms of

professional learning. A one standard deviation in traditional professional learning was associated with a predicted .26 standard deviation increase in interactive civics instruction. The same increase for communities of practice professional learning resulted in a predicted .29 standard deviation in interactive civics instruction. Additionally, a one standard deviation increase in both traditional and communities of practice professional learning was related to a predicted .039 standard deviation increase in student achievement on NAEP civics. This means that the difference between the two forms of professional learning in their relationship to student achievement was less than .001 of a standard deviation.

This finding stands in stark contrast to the critical sentiments toward traditional professional learning in the literature. Traditional professional learning is often described in pejorative terms: "disconnected from practice" (Darling-Hammond et al., 2009, p.9), "intellectually superficial" (Borko, 2004, p. 3), "the professional equivalent of yo-yo dieting" (Ball & Cohen, 1999, p. 4). These characterizations are likely true about many traditional forms of teacher professional development. Yet, they may also reflect the general lack of quality in professional learning opportunities for teachers (H. Hill, 2009), rather than a problem endemic to the form. Indeed, some studies have found that traditional forms of professional development can be beneficial if they are of high quality (Fishman et al., 2003; Harris, et al., 2012; Walker et al., 2011).

By contrast, descriptions of teacher communities of practice are often overly idealistic. For example, McLaughlin and Talbert (2006) describe school-based professional learning communities as:

> Teachers work collaboratively to reflect on practice, examine evidence about the relationship between practice and student outcomes, and make changes that improve teaching and learning for the particular students in their classes (p.4)."

Although these images of teacher collaboration are appealing, they are often difficult to implement in practice. Teachers may be hesitant to honestly reflect on their work (Valli, 1997), critique a colleague's teaching (Giles & Hargreaves, 2006) or change their own instructional approach (Ball & Cohen, 1999; Coburn, 2001). Teachers may also lack skills and knowledge to draw inferences about instruction based on student data (Ingram, Louis, & Schroeder, 2004; Wayman & Stringfield, 2006). Moreover, schools may not provide teachers with sufficient time to participate in these communites of practice forms of professional development (H. Hill, 2009). To make matters more complicated, the language of communities of practice is often adopted by schools without changing the content of these forms of professional learning, for example, calling a department meeting a "professional learning community" (DuFour, 2004; Grossman, Wineburg, & Woolworth, 200; H. Hill, 2009). This can create confusion when examining the relationship between a particular form of professional learning and student outcomes.

Accordingly, the findings of this study suggest that the form of professional learning that teachers participate may be less relevant for student achievement than how much they participate in professional learning. Distinctions between the forms of traditional and communities of practice forms of professional learning have shrunk with many "traditional" programs adopting aspects of communities of practice (Glazerman, et al., 2008; Garet, et al., 2011). Additionally, traditional professional development and communities of practice can have a symbiotic relationship where high quality professional development can provide resources and benefits to the teacher communites of practice (Lumpe, 2007; McLaughlin & Talbert, 2006). Indeed, many

educators who engage in communities of practice also participate in traditional professional development activities. For example, Carpenter and Krutka (2014) found that educators who used social media for professional learning also participated in many traditional forms of professional development. As a result, both forms of teacher professional learning may afford potential benefits for teachers.

Interactive Civics Significantly Related to Student Achievement but Small Effect Size

The study found that teachers' use of interactive civics activities was a significant predictor of their students' performance on the NAEP civics assessment. This finding is consistent with other studies that have similar positive relationships between participation in interactive civics activities and civic skills and knowledge (CIRCLE, 2013; Kahne & Sporte, 2008; Kawashima-Ginsberg, 2013; Torney-Purta, 2002; T. Zhang, Torney-Purta, & Barber, 2012). However, the effect size of interactive civics activities was very small; a one standard deviation increase in interactive civics was associated with only a predicted .03 standard deviation increase in student achievement on NAEP, or the difference between scoring in the 50th and 51st percentile. This meant that a one standard deviation increase in interactive civics use would be the equivalent of reducing the point difference between students scoring at the Basic and Proficient levels by 2.31%.

Comparing "high users" to "non-users" increased this effect size. A teacher who used five "high impact" activities would be expected to have student scores a predicted .095 standard deviations higher than a teacher who used no interactive civics activities. This is the equivalent of reducing the difference between students scoring Basic and Proficient levels by 7.19%. However, this is still a relatively small effect size for an educational intervention (C. Hill et al., 2007).

The effect size of interactive civics found in this study was smaller than those found in other studies. Niemi and Junn (1998) observed that 12th grade students who reported participating in mock trials scored a predicted 2.6 points higher (.21 standard deviations)[15] on the 1988 NAEP civics assessment. They also found that students who reported discussing current events every day scored 4 points (.32 standard deviations) higher than those who never discussed currents in class. Similarly, Kawashima-Ginsberg (2013), in a descriptive analysis of the 2010 NAEP civics data, found that 8th grade students who reported discussing current events in class frequently had NAEP civics scores that were .37-.44 standard deviations higher than those who did not discuss current events in class. In a related subject, Smith and Niemi (2001) found that a one standard deviation increase in students' reported active instruction in history was related to a .35 standard deviation increase in student achievement on the 1994 12th grade NAEP history assessment.

One possible reason that the effect size in this study was smaller than other investigations is that this study relied on teacher-reported rather than student-reported data. Student-reported data may be an unreliable indicator of a teacher's classroom practices. Students may be less likely to recall what activities they did in class and there may be discrepancies between students. For example, Desimone et al., (2010) found low correlations between student-reported and teacher-reported measures of math classroom practices on the NAEP math assessment. The study authors also observed that the differences between student and teacher-reported measures were related to student characteristics such as parental education, gender, and interest in math, and being in advanced classes. Student achievement in math also predicted how similar students' responses were to their teacher. As a result, student-reported measures of interactive classroom

[15] Effect sizes calculated by the author. The standard deviation of the 1988 NAEP civics assessment was derived from Johnson, Eugene G. & Zwick, Rebecca. (1990). *Focusing the new design: The NAEP 1988 technical report*. Princeton: Educational Testing Services.

practice may be confounded by student achievement. Students who had higher achievement may be more likely to report engaging in interactive classroom practices. Teacher-reported measures may therefore be a more reliable indicator than student-reported measures of interactive classroom practices.

Another reason that the study may have found lower effect sizes is that it used an index of interactive civics practices, rather than examining individual activities. Indices may have lower effect sizes because they examine a range of practices that may have different relationships with student achievement. Other studies that have used indices to examine the relationship between instructional practices and student civic achievement have found smaller effect sizes. For example, Campbell (2008) found that the openness of the discussion climate in the classroom had a weak relationship with student civic achievement on the 1999 CIVED study; a one standard deviation increase in the discussion climate was associated with a .06 standard deviation increase in student achievement. Using the same data, Homana (2009) observed that a one standard deviation increase in students' perceptions of opportunities for active involvement was associated with a .05 standard deviation increase in students' scores. It is therefore plausible that indices may be less sensitive than individual items in capturing the relationship between classroom practices and student achievement in civics.

Finally, the small effect size of interactive civics may reflect a misalignment between what students learn through interactive civics and the skills and knowledge tested by NAEP. Advocates for interactive civics argue that these activities help students learn how to do things like evaluate claims, develop persuasive arguments, work together in groups, and address community problems (Gingold, 2013; Hess, 2002; Levinson, 2012; Shiller, 2013). These skills are either not assessed by NAEP or only addressed superficially. As a result, NAEP civics may

not be fully capturing changes in students' knowledge and skills as a result of participating in interactive civics activities.

Yet, traditional forms of civic knowledge are also important. Content knowledge provides a framework for students to develop more participatory skills. In order to be effective advocates, students need to understand the separation of powers in the Constitution in order to know which political representative they should contact and what types of requests they should make. Moreover, teachers cannot neglect facts entirely. As Wilson and Wineburg (1993) noted "History *is* interpretation, but interpretation must be backed by a solid knowledge of facts." (p.758) Understanding basic facts about how government works is necessary to being an informed democratic citizen.

It is therefore important that interactive civics activities had a small, but significant, relationship with student achievement on NAEP civics. Using interactive civics in class not only did not reduce students' factual knowledge of civics but may have actually slightly enhanced it. There are several reasons why this may have occurred. Interactive civics activities may have increased student engagement allowing them to develop a greater understanding of the content (Hess, 2002; Kahne, Chi, & Middaugh, 2006). These activities may have also focused more on conceptual understanding (T. Zhang, Torney-Purta, & Barber, 2012) allowing students to better synthesize disparate areas of knowledge rather than just memorizing facts. Finally, interactive civics experiences may have increased student interest in civic participation, causing them to seek out and pay closer attention to civic knowledge (Kahne & Sporte, 2008; Levinson, 2012).

Non-Linear Relationship between Interactive Civics and NAEP Civics Achievement

Teachers' use of interactive civics also had a significant quadratic relationship with student achievement on NAEP. Interactive civics use was associated with increases in student

achievement up to a certain point, around 5 high impact activities, after which student achievement began to decline. This finding is consistent with the results of the study's cognitive interviews conducted with current social studies teachers. For most activities, teachers weighted frequent use (e.g., almost every day) as being less effective than conducting the same activity once or twice a month. Taken together, these findings suggest that interactive civics may be hurt student achievement in civics if they are used too frequently by teachers.

There are a number of reasons why interactive civics, if used too frequently, might reduce student academic achievement. Interactive activities may "crowd out" other forms of instruction such as lectures or independent reading which may be helpful in facilitating student understanding of the content. Frequent use of interactive civics activities may leave students and teachers with little time to adequately prepare for activities overburdening students and causing burnout. Finally, overuse may reflect a lack of flexibility on the part of the teacher. Some researchers have found that effective teachers constantly adapt curriculum in response to their students' needs (Drake & Sherin, 2006; Parke & Coble, 1997; Remillard, 1999). Teachers who use one type of activity all of the time may be too rigid in their teaching practice and may not be responsive to the instructional needs of their students.

Professional Learning was Not a Significant Moderator of the Relationship between Interactive Civics and NAEP Civics Achievement

This study did not find sufficient evidence that professional learning moderated the relationship between teachers' use of interactive civics and student achievement on NAEP civics. The interaction terms for total, traditional, and communities of practice professional learning were all not statistically significant at the $p<.05$ level. This indicates that there was not enough evidence in the data to support the proposition that professional learning changed the magnitude

of the relationship between their use of interactive civics activities and student achievement on NAEP civics.

Although the relationships were not statistically significant, the effect sizes of the interactions were not negligible. Depending on the form, a one standard deviation change in professional learning moderated the slope for interactive civics anywhere from between .08 and .14 points. Given that the slope for interactive was fairly small these are quite substantial changes. Moreover, the direction of the effect was opposite of what was initially predicted. The interaction terms for teacher professional learning were negative, indicating that teachers who engaged in more professional learning had a weaker relationship between interactive civics use and student achievement.

This finding diverges from the conclusions of other studies that have suggested that professional learning may help teachers be more effective in their use of instructional strategies (Ball & Cohen, 1999; Desimone, 2009; J.W. Little, 2006; Wilson & Berne, 1999). There are several reasons why this may not have been the case in this study. This study assumed that the relationship interactive civics between and student civic achievement varied linearly based on the continuous values of the moderator. However, these types of inferences often have low statistical power when there is measurement error in the predictor and moderator as there was in this study (Baron & Kenny, 1986; Jaccard, Wan, & Turrisi, 1990). As a result, the study may have been statistically underpowered to detect moderation effects. Underpowered studies are more likely to produce non-significant results. Moreover, in the event that an underpowered study produces a significant result in it is likely to be vastly overinflated, or even in the opposite direction, of the true effect (Gelman & Weakliem, 2009)

Additionally, the presence of an interaction effect was also hampered by restriction of range. Specifically, there were few teachers in the study who had high levels of interactive civics use and low levels of participation in professional learning. When such restrictions of range exist it often reduces the statistical power of the analysis, decreasing the probability that a given interaction will be statistically significant (Frazier, Tix, & Barron, 2004).

Finally, the quadratic relationship between interactive civics use and student achievement may have resulted in a spurious negative interactions. Lubinski and Humphreys (1990) observed that quadratic relationships in the data can sometimes produce spurious interaction terms. This may have occurred in this study and may explain the magnitude and direction of the interaction terms.

It is also possible that there were substantive reasons why the direction of the moderator was different than originally theorized. Teachers who engage in greater professional learning may be more effective teachers, so how much they use interactive civics activities may not matter as much. Yet, given the lack of statistical significance, low reliability, restriction of range, and potential spurious nature of the interaction, this interpretation is tenuous at best. It is more likely that the negative interaction is an artifact of this particular data rather than reflecting real trends in population.

Research Implications

With the growing focus on teacher quality in education (Green, 2014; Hanushek & Rivkin, 2010; Kane & Steiger, 2008), policymakers are increasingly paying attention to how teachers might influence the educational trajectories of their students. This study found that there was significant variation between teachers in their students' civic skills and knowledge. Furthermore, it identified two factors that were related to a teachers' effectiveness in civics: their

participation in professional learning opportunities and their use of interactive civics activities in the classroom. Students whose teachers participated in more professional learning opportunities and used more interactive civics activities had higher achievement on the NAEP civics assessment.

These findings suggest that policymakers should increase access to professional development activities specifically focused on teaching civics. These opportunities might take different forms. For example, policymakers might bring in outside organizations like Facing History and Ourselves or Project Citizen that have an established track record of improving teachers' skills and knowledge (Boulay, et al., 2010; Tolo, 1998). Policymakers might also consider promoting internal forms of professional learning such as Socratic seminars (Hess & Zola, 2012) or teacher reading groups (Grossman, Wineburg, & Woolworth, 2001). These types of experiences may improve teachers' instructional skills by increasing their content knowledge, learning from experiences of other teachers, and modeling new instructional strategies that they can use in their classroom (CIRCLE, 2013; Hess & Zola, 2012) .

However the study also found that although participation in professional learning was a significant predictor of student achievement, the effect size was relatively small. This suggests that participation alone is not sufficient and that more attention should be paid to the quality of the professional learning experience. Quality is not the same as form. The study did not find any significant difference between traditional activities such as workshops and communities of practice activities such as discussion groups. Indeed the literature suggests that the quality of a professional learning experience is affected by characteristics such as duration, content focus, and being based in practice (Ball & Cohen 1999; Desimone, 2009). However, these factors are not deterministic; simply increasing the length or learning style is no guarantee that a

professional learning experience will be effective. Rather, professional learning experiences that rigorously examine what teachers need to know and be able to do in the classroom are more likely to succeed than those that are removed from practice (Ball & Cohen, 1999; H. Hill, 2009).

Additionally, the study found students who experience interactive civics activities in their social studies classes have higher achievement on NAEP civics. This suggests that interactive civics activities such as debates, mock trials, and working on community projects may improve students' civic knowledge and skills in addition to their benefits for other civic outcomes. Policymakers should find ways to encourage teachers to integrate these types of activities into their curriculum. There are a number of different ways to accomplish this goal. For example, state legislators in Tennessee passed a bill requiring students to complete a project-based assessment based on the Project Citizen curriculum (CIRCLE Staff, 2014). The assessment ensured that all students in the state had an opportunity to learn how to research and craft solutions to problems in their community. As another example, in Boston, the 8th grade social studies curriculum is designed around developing students' active citizenship skills and encourages students to explore ideas and issues based on their interest (Boston Public Schools, 2014). Schools may also want to consider partnering with non-profit and community organizations to bring volunteers into the classroom to teach an interactive civics activity. For example, Citizen Schools brings professional lawyers into low-income middle schools to teach students a 10-week lawyer "apprenticeship" focused on conducting a mock trial (Citizen Schools, 2014).

Yet, as with professional learning, the effect size of interactive civics activities on student achievement was very small. Additionally, the study found that overuse of interactive civics activities may actually decrease student achievement. This suggests that simply increasing the

number of interactive civics activities may not be sufficient. More attention needs to be paid by civic educators as to how the interactive activity might facilitate greater student civic skills and knowledge.

Civics educators may want to follow the example of math educators who have spent a great deal of time reflecting on how instructional strategies might promote greater student learning in math. Math educators have found that instructional strategies that focus on students' conceptual understanding are more effective in promoting mathematical knowledge (D. Cohen & Hill, 2000; Hill, Rowan, & Ball, 2005; Ball, Thames, & Phelps, 2008).

Interactive civics activities may be more effective if those activities are linked to deepening students' conceptual understanding of civic institutions, systems, and processes. For example, a teacher may conduct an in-class debate where students are asked to debate whether the Constitution gives the principal the right to search a students' backpack for drugs. To prepare for the debate students are asked to research previous cases, develop arguments, and think about they would respond to possible counter-arguments. Such a process would not just teach students about the Constitution; it would also help them develop an understanding of how laws are debated and interpreted within the judicial system.

Limitations

All studies have some limitations and this study is no exception. This section will review the limitations of the data, measures, and analysis procedures, and describe how these limitations may affect some of the findings from this study.

Teachers Not Randomly Sampled

Teachers, unlike students and schools, were not randomly sampled by NAEP. Instead teachers were included in the study if their students are sampled by NAEP (Rogers, Stoeckel, &

Sikali, 2012). As a result, NAEP data cannot be used to make inferences about the national population of 8th grade social studies teachers. Instead, the teacher sample should be interpreted as the teachers of a representative sample of 8th grade students. This sample likely differed slightly from a nationally representative sample of teachers. For example, teachers who teach small classes were likely underrepresented and teachers who teacher larger classes were likely overrepresented. Consequently, teacher estimates may not have accurately reflected the overall population of 8th grade public and private school teachers in the United States.

Alignment Between Instruction and Assessment Content

The NAEP civics assessment frameworks was developed by the National Assessment Governing Board and was not specifically aligned to any state social studies standards (National Assessment Governing Board, 2010). Additionally, most states do not assess middle school students in social studies, so districts and schools have more flexibility in terms of the types of content that students might learn in 8th grade social studies (CIRCLE, 2014). Although 85% of students in the NAEP sample reported having learned civics and government in 8th grade (National Center for Education Statistics, 2010), some of the civics content and skills assessed on the NAEP civics assessment may not reflect what students had covered in their classes. As a result, some students may not perform as well on certain questions on the NAEP assessment because they have not had the opportunity to learn the content.

A sensitivity analysis (Appendix H) was conducted in order to assess how the alignment between instruction and practice might influence the results of this study. The analyses was rerun without teachers who reported spending less than 10% of their class time teaching civics to determine if excluding those teachers changed the relationships between the predictors and outcomes of interest. Overall, excluding these teachers reduced the magnitude of the observed

effect sizes in the analysis. Additionally, the analyses was also conducted without schools where the principal reported that students do not typically take a class specifically focused on civics or government in 8th grade to determine if excluding those schools altered the relationships between the predictors and outcomes of interest. Excluding these schools also reduced the magnitude of the observed effect sizes, although the reductions were not large.

These findings suggest that the lack of alignment between instruction and assessment content may have had some effect on the magnitude of the observed findings in the study. Specifically, teachers who used interactive civics activities in the classroom may have also been more likely to have been teaching curriculum that was more aligned with the NAEP civics assessment. However, without knowing more about the curriculum that students received within each classroom, it is difficult to ascertain the extent that curricular misalignment affected the analysis in the study.

Limitations of Observational Data

This study examined observational data that was not designed nor intended to make causal inferences about any of the predictors in the analysis. Since teachers were not randomly assigned to conditions, there is no way to rule out the possibility that differences between teachers are not caused by unobservable variables. For example, students who have a teacher who participates in a lot of professional development or who regularly uses interactive activities in their classroom may already be predisposed to doing better on a standardized assessment like NAEP. This means that differences on NAEP between students may have reflected pre-existing differences between the students rather than the effects of teacher professional learning or interactive civics instruction.

Additionally, schools with more resources to send teachers to conferences and workshops may also have a higher socio-economic student population, and therefore are more likely to have higher academic achievement regardless of whether the training is effective. Controlling for possible confounding variables such as race and socio-economic status may partially account for some of these characteristics, but they cannot explain them entirely (Shadish, Cook, & Campbell, 2002).

Construct Validity of Predictors and Outcomes of Interest

NAEP civics assessment. NAEP only measured students' civic skills and knowledge, not their civic attitudes or behaviors (National Assessment Governing Board, 2010). As a result, the study was not able to capture differences in students' civic self-efficacy or their participation within their community. Furthermore, the types of civic knowledge and skills that were measured by NAEP were limited to what can be reliably captured through a standardized written assessment. This means the study was not able to measure whether students have developed important civic skills such as oral communication or the ability to work productively within a group. Additionally, the questions on NAEP civics likely reflect the civic values and expectations of the test developers who are predominantly White, middle-class, and college-educated (Levinson, 2010). As a result, NAEP may assess a certain set of civic values that not all civic educators might share. Indeed, students may be developing knowledge and skills within interactive civics activities that are not being adequately captured by NAEP civics. Consequently, this study may have underestimated the relationship between interactive civics activities and a broader set of civic skills and knowledge.

Teacher professional learning. Professional learning activities was measured by the number of different types of subject-matter professional development activities that teachers

engaged in over the past two years. Although the list is extensive it may not fully account for all the different ways that teachers might engage in professional learning. Additionally, the teacher background survey did not measure the quality of those activities or how much teachers learned from them. A teacher who participated every day in a teacher online network on civics and one who attended a two-hour seminar on the Constitution would be assigned the same value for those activities on the teacher professional development measure. Consequently, this analysis may have underestimated the true relationship between teacher professional learning and interactive civics and student achievement on the NAEP civics assessment.

Interactive civics. For interactive civics instruction, teachers were only asked about how frequently they used each type of activity. The teacher background survey did not measure how well the lesson was executed or how "active" students' participation was within the activity. There are also some forms of interactive civics instruction, such as making a video or developing an advocacy campaign, that are not measured by NAEP. Additionally, instructional practices were self-reported by the teacher. Although some studies have shown that teacher self-reports can be accurate (D. Mayer, 1999; Supovitz & Turner, 2000) there is always some level of concern about the accuracy of self-reported data. For example, one study found that teacher underestimated the frequency of using different instructional practices on surveys when compared with observations and instructional logs (Mullens, 1998). In the other direction, others have found that teachers are more likely to report using instructional practices than their students report participating in them (Desimone, Smith, & Frisvold, 2010). Since the measures of teacher use of interactive may be inaccurate, the study may underestimate the true relationship between interactive civics and student achievement on the NAEP civics assessment.

Low Reliability of Teacher Professional Learning and Interactive Civics Measures

Both the measures of teacher professional learning and interactive civics use had relatively low reliability. This meant that there was likely a substantial amount of variance in the underlying construct that was not being captured by the instrument. As a result, the relationship between these constructs and student achievement may have been attenuated in this analysis. This may have been a result of limited response categories on the survey: Teachers only had two options for professional learning and four response options for the interactive civics measures. Furthermore, the response options for the interactive civics activities were skewed toward the higher frequencies (e.g., once or twice a week, almost every day) which were not applicable for many of the higher intensity interactive civics activities. Alternatively, the low reliability may reflect the multidimensionality of the constructs. It may be that teacher professional learning and interactive civics actually reflect multiple dimensions that could not be adequately assessed using a single measure.

Areas for Future Research

The findings of this study raise a number of questions and topics that could be explored in future research studies. The study observed that there were correlational relationships between teachers' professional learning and interactive civics use and students' civic skills and knowledge. A logical next step would be to design an experimental study that examines whether professional learning in civics and use of interactive civics activities improves student civic achievement. Although there have been some experimental studies on this topics (e.g., Boulay, et al., 2010; Kawashima-Ginsberg, 2012; Syvertesen et al., 2009) none have examined their effects on students' civic knowledge using a full-length standardized assessment. One option would be to take advantage of state standardized tests in civics/government that are currently being offered

in nine states (CIRCLE, 2014). Schools could be randomly assigned to receive professional learning focused on teaching using interactive civics activities. The study would then assess whether students in schools that participated in the professional learning were exposed to more interactive civics activities and whether students had higher achievement on the state standardized test.

Another avenue for future research is designing a performance-based assessment focused on civics skills such as analyzing a community problem or crafting an argument about a particular issue. Such an assessment would be better aligned with the types of skills and knowledge that are advanced through interactive civics activities. There are some examples of performance-based assessments in civics that already being developed. For example, Tennessee recently passed a law that requires students to complete a project-based assessment in civics (CIRCLE Staff, 2014). Additionally, there have been some efforts among proponents of "action civics" to develop performance-based tasks to assess students' competencies to take action in their communities (Gingold, 2013). However, the psychometric properties of these assessments have not yet been reviewed. Future studies may therefore seek to develop a performance-based assessment for civics and assess evidence of its reliability and validity.

Additionally, future studies should examine the predictive validity of using NAEP civics as a measure of students' future civic engagement. Although some studies have found a relationship between civic achievement and intended engagement (A. Cohen & Chafee, 2012; Torney-Purta & Amadeo, 2003), no study has examined whether students' scores on NAEP civics predict future civic engagement. Such a study would provide additional evidence about the validity of using student score on the NAEP civics assessment as a way of drawing inferences about whether students are adequately prepared for the responsibilities of democratic citizenship.

Finally, the findings of this suggest that more research should be conducted on the quality of professional learning experiences and interactive civics practices. This could be addressed through qualitative research studies. For example, observations could be conducted of interactive civics activities to better understand how these types of activities are used in practice. This data would provide a more thorough understanding about what makes for a high-quality interactive civics experience. Additionally, interviews could be conducted with civics teachers about their own experiences with professional learning and how their professional learning experiences translate into classroom instructional practices. Such a study would likely illuminate some of the characteristics of high-quality effective professional learning in civics. This could then be used to help design future professional learning opportunities for civics teachers.

Conclusion

A recent study described the current state of college student civic engagement as a "crucible moment" (The National Task Force on Civic Learning and Democratic Engagement, 2012). This is also an apt metaphor for the current state of affairs in K-12 civic education. On the one hand, current trends in youth civic engagements are troubling. Millennials are less likely than previous generations to participate in their communities and engage in social activism (Caren, Ghoshall, & Ribas, 2011; CIRCLE, Harvard Institute of Politics, and Mobilize.org, 2012). Youth voter participation has been relatively flat for the last thirty years and is still far too low: almost 80% of eligible 18-29 years did not vote in the 2014 midterm elections (CIRCLE, Harvard Institute of Politics, and Mobilize.org, 2012; CIRCLE, 2014). Civic engagement is also highly stratified by race, socio-economic status, and educational attainment and these gaps have expanded over time (CIRCLE, Harvard Institute of Politics, and Mobilize.org, 2012; Flanagan, 2004; Hart & Atkins, 2002; Ingels et al., 2012; Levinson, 2012). These trends in civic

engagement raise serious doubts about the viability of American democracy in the future. This may be happening already. One group of researchers recently found that the positions of economic elites and business interests substantially affected the policies of the U.S. government, while the views of the average citizen and mass-based interest groups had little effect (Gilens & Page, 2014).

On the other hand, there are also many policymakers, researchers, and educators who are seeking to change this downward civic trajectory. Through service-learning, mock trials, class discussions, debates, and action research projects students are learning the types of civic skills and knowledge that are necessary to be an engaged citizen. Teachers are also being offered professional learning opportunities to learn how to teach using these methods (Boulay, et al., 2010; Hess & Zola, 2012). Many advocates hope that such efforts may alter our country's current course of civic engagement creating a new generation of active democratic citizens.

This study added to the knowledge base for this effort. It observed that teachers who engaged in more professional learning were more likely to use interactive civics activities in the classroom. Additionally, it found that teachers' professional learning and use of interactive civics were both significant predictors of student achievement on the NAEP civics assessment. However, it also found that using interactive civics activities too frequently may decrease civic achievement and that there was no evidence that professional learning strengthened the relationship between interactive civics and students' civic achievement. These findings will hopefully help inform civic researchers, policymakers, educators going forward as they strive to educate tomorrow's democratic citizens.

Bibliography

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95-135.

Abedi, J. (2004). The No Child Left Behind Act and English Language Learners: Assessment and accountability issues. *Educational Researcher, 33*(1), 4-14.

Anderson, J. R., Reder, L., & Simon, H. (1996). Situated learning and education. *Educational Researcher, 25*(4), 5-11.

Andolina, M., Jenkins, K., Zukin, C., & Keeter, S. (2003). Habits from home, lessons from school: Influences on youth civic engagement. *PS: Political Science and Politics, 36*(2), 275-280.

Arnold, N., & Paulus, T. (2010). Using a social networking site for experiential learning: Appropriating, lurking, modeling, and community building. *Internet and Higher Education*(13), 188-196.

Bachner, J. (2010). *From classroom to voting booth: The effect of high school civic education on turnout.* Cambridge, MA: Harvard University: Manuscript.

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., . . . Shepard, L. (2010). *Problems with the use of student test scores to evaluate teachers. EPI Briefing Paper #278.* Washington D.C.: Economic Policy Institute.

Ball, D., & Cohen, D. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Fancisco, CA: Jossey Bass.

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

Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Barr, D. (2010). *Continuing a tradition of research and learning on the foundations of a democratic education: The national professional development and evaluation project.* Brookline, MA: Facing History and Ourselves National Foundation.

Bartels, L. M. (1996). Uninformed votes: Information effects in presidential elections. *American Journal of Political Science, 40*(1), 194-230.

Beane, C., Kawashima-Ginsberg, K., Kiesa, A., & St.Rose, A. (2014). *Closing the leadership gap: How educators can help girls lead.* Washington, D.C.: National Education Association.

Beck, I., & McKeown, M. (2002). Questioning the author: Making sense of social studies. *Educational Leadership, 60*(3), 44-47.

Belanger, A., & Stein, S. (2013, Spring). Closing the academic divide through debate. *Communities and Banking*, pp. 16-19.

Bell, B., Morgan, G., Kromrey, J., & Ferron, J. (2010). The impact of small cluster size on multilevel models: A Monte Carlo examination of two-level models with binary and continous predictors. *JSM Proceedings, Section on Survey Research Methods, American Statistical Association*, 4057-4067.

Bell, B., Morgan, G., Schoenberger, J., Kromkey, J., & Ferron, J. (2014). How low can you go? An investigation of the influence of sample size and model complexiy on point and interval estimates in two-level linear models. *Methodology, 10*(1), 1-11.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289–300.

Bennett, S. E. (2003). Is the public's ignorance of politics trivial? *Critical Review: A Journal of Politics and Society, 15*(3-4), 307-337.

Bernstein, J. (2007). Simulations and the dynamics of racial and gender gaps in civic competence. *New Directions for Teaching and Learning, 111*, 89-96.

Berry, B., Norton, J., & Byrd, A. (2007). Lessons from networking. *Educational Leadership, 65*(1), 48-52.

Bickel, R. (2007). *Multilevel Analysis for Applied Research: It's Just Regression.* New York: The Guildford Press.

Biddle, P. (2014, February 18). *The probelms with little white girls (and boys): Why I stopped being a voluntourist*. Retrieved April 27, 2014, from Pippa Biddle: Life is an Adventure: http://pippabiddle.com/2014/02/18/the-problem-with-little-white-girls-and-boys/

Billig, S., Root, S., & Jesse, D. (2005). *The Impact of Participation in Service-Learning on High School Students' Civic Engagement.* College Park, MD: CIRCLE Working Paper 33.

Billing, S. (2000). Research on K-12 school-based service-learning: The evidence builds. *Phi Delta Kappan, 81*(9), 658-664.

Blitz, C. (2013). *Can online learning communities achieve the goals of traditional professional learning communities? What the literature says.* Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Educational Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. Retrieved from http://ies.ed.gov/ncee/edlabs

Blouin, D. B., & Perry, E. M. (2009). Whom does service-learning really serve? Community based organizations' perspectives on service-learning. *Teaching Sociology, 37*(2), 120-135.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15.

Boss, J. (1999). *Outdoor education and the development of civic responsibility. Eric Digest.* Charleston, WV: Appalachia Educational Laboratory.

Boston Public Schools. (2014). *Civics in Action: Course Description.* Boston, MA: Boston Public School.

Boulay, B., McCormick, R., Gamse, B., Barr, D., Selman, R., Lowenstein, E., & Fine, M. (2010). *Effects of Facing History and Ourselves on Teachers and Students: Findings from the National Professional Development and Evaluation Project.* Cambridge, MA: Abt Associates.

Brady, H., Verba, S., & Schlozman, K. L. (1995). Beyond SES: A resource model of political participation. *American Political Science Review, 89*(2), 271-294.

Bransford, J., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Research Council, National Academy of Sciences.

Braun, H. (2005). *Using Student Progress to Evaluate Teachers: A primer on Value-Added Models.* Princeton, NJ: Educational Testing Services.

Braun, H., Jenkins, F., & Grigg, W. (2006). *A closer look at charter schools using hierarchical linear modeling.* Washington, D.C.: National Ceter for Education Statistics.

Brown, J. S., & Duguid, P. (1991). Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation. *Organization Science, 2*(1), 40-57.

Bryk, A., Camburn, E., & Louis, K. S. (1999). Professional Community in Chicago Elementary Schools: Faciliating Factors and Organizational Consequences. *Educational Administration Quaterly, 35*, 751-781.

Bryk, S., Camburn, E., & Louis, S. K. (1999). Professional communities in Chicago Elementary Schools: Facilitating factors and organizational consequences. *Education Administration Quaterly, 35*(5), 751-781.

Burroughs, S., Groce, E., & Webekj, M. L. (2005). Social studies education in the age of testing and accountability. *Educational Measurement: Issues and Practice, 24*(3), 13-20.

Cai, X. (2008). *Missing data treatment of a level-2 variable in a 3-level hierarchical linear model.* Kalamazoo, MI: Doctoral Dissertation, Western Michigan University.

Callahan, R., Muller, C., & Schiller, K. (2010). Preparing the next generation for electoral engagement: Social studies and the school context. *American Journal of Education, 116*(4), 526-556.

Cammarota, J., & Fine, M. (2008). Youth participatory action research: A pedagogy for transformational resistance. In J. Cammarota, & F. Michelle, *Revolutionizing Education: Youth Participatory Action Research in Motion* (pp. 1-12). New York, NY: Taylor & Francis.

Campaign for the Civic Mission of Schools. (2011). *Guardians of Democracy: The Civic Mission of Schools.* Philadelphia, PA: The Leonore Annenberg Institute for Civics of the Annenberg Public Policy Center at the University of Pennsylvania.

Campbell, D. (2008). Voice in the classroom: How an open classroom climate fosters political engagement among adolescents. *Political Behavior, 30*, 437-454.

Carbonaro, W., & Covay, E. (2010). School sector and student achievement in the era of standards based reforms. *Sociology of Education, 83*(2), 160-182.

Caren, N., Ghoshall, R. A., & Ribas, V. (2011). A social movement generation: Cohort and period trends in protest attendence and petition signing. *American Sociological Review, 76*(1), 125-151.

Carle, A. (2009). Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC Medical Research Methodology, 9*, 49.

Carnegie Corportation of New York and CIRCLE. (2003). *The Civic Mission of Schools.* New York: Carnegie Corportation of New York and Center for Information and Research on Civic Learning and Engagement. Retrieved 4 October, 2013, from http://civicmission.s3.amazonaws.com/118/f7/1/172/2003_Civic_Mission_of_Schools_Report.pdf

Carpenter, J., & Krutka, D. G. (2014). The assorted affordances and uses of Twitter: A survey of educators. *Paper presented at the annual meeting of American Educational Research Association.* Philadelphia, PA.

Celio, C., Durlak, J., & Dymnicki, A. (2011). A meta-analysis of the impact of service learning on students. *Journal of Experiential Education, 34*(2), 164-181.

Cheong, Y. F., Fotiu, R., & Raudenbusch, S. (2001). Efficency and robustness of alternative estimators for two- and three-level models: The case of NAEP. *Journal of Educational and Behaviorial Statistics, 26*(4), 411-429.

Chetty, R., Friedman, J., & Rockoff, J. E. (2012). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood.* Cambridge, MA: NBER Working Paper No.17699.

Cho, V., Ro, J., & Littenberg-Tobias, J. (2013). What Twitter will and will not do: Theorizing about teachers' professional communities online. *Learning Landscapes, 6*(2), 45-62.

CIRCLE. (2013). *All Together Now: Collaboration and Innovation for Youth Engagement: The Report of the Commission on Youth Voting and Civic Knowledge.* Medford, MA: Center for Information & Research on Civic Learning and Engagement. Retrieved from www.civicyouth.org/about-circle/commission-on-youth-voting-civic-knowledge/

CIRCLE. (2013). *Civic Learning through Action: The Case of Generation Citizen.* Medford, MA: The Center for Informationa& Research on Civic Learning and Engagement.

CIRCLE. (2014). *State Civic Education at a Glance.* Retrieved April 30, 2014, from CIRCLE: The Center for Information and Research on Learning and Engagement: http://www.civicyouth.org/state-civic-education-at-a-glance/

CIRCLE. (2014). *Update-21.5% Youth Turnout: Two Day Estimate Comparable to Recent Years.* Medford, MA. Retrieved from http://www.civicyouth.org/21-3-youth-turnout-preliminary-estimate-comparable-to-recent-midterm-years/

CIRCLE Staff. (2013). *What do Young Adults Know about Politics: Evidence from a National Survey Conducted After the 2012 Election.* Medford, MA: Center for Information & Research on Civic Learning and Engagement.

CIRCLE Staff. (2014). *Civic Education Policy Change: Case Studies of Florida, Tennessee, and Hawaii (CIRCLE Working Paper #79).* Medfod, MA: The Center for Information and Research on Civic Learning and Engagement.

CIRCLE, Harvard Institute of Politics, and Mobilize.org. (2012). *Millennials Civic Health Index.* Washington, D.C.: National Conference on Citizenship.

Citizen Schools. (2014). *Mock Trials*. Retrieved November 19, 2014, from Citizen Schools: http://www.citizenschools.org/apprenticeship/mock-trial/

Clarke, P., & Wheaton, B. (2007). Addressing data sparesness in contexual population research: Using cluster analysis to create synthetic neighberhoods. *Sociological Methods & Research, 35*(3), 311-351.

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology, 100*(5), 1251-1293.

Coburn, C. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis, 23*(2), 145-170. doi:10.3102/01623737023002145

Cochran-Smith, M., & Lytle, S. L. (1999). Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education, 24*, 249-305.

Cohen, A., & Chafee, B. (2012). The relationship between adolescents' civic knowledge, civic attitude, and civic behavior and their self-reported future likelihood of voting. *Education, Citizenship, and Social Justice*, 1-15.

Cohen, D., & Hill, H. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record, 102*(2), 294-343.

Cohen, J. (1992). A power primer. *Quantitative Methods in Psychology, 112*(1), 155-159.

Cohen, J. (1994). The Earth is Round p<.05. *American Psychologist*, 997-1003.

Coleman, J. (1968). The concept of equality of educational opportunity. *Harvard Educational Review, 38*(1), 7-22.

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research, 12*(3), 229-238.

Collins, L. M., Schafter, J., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*(4), 330-351.

Connor, C. M., Son, S.-H., Hindman, A., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Pscyhology, 43*, 343-375.

Corcoran, S. (2010). *Can Teachers be evaluated by their Students' Test Scores? Should They Be? The Use of Value-Added Measures of Teacher Effectivness in Policy and Practice.* Providence, RI: Annenberg Institute for School Reform at Brown University.

Daly, J. K., Devlin-Scherer, R., Burroughs, G., & McCartan, W. (2010). The status of civic education: A preservice program response. *The Education Forum, 74*(2), 117-128.

Darling-Hammond, L. (1999). *Teacher Quality and Student Achievement: A Review of State Policy Evidence.* Palo Alto, CA: Center for the Study of Teaching and Policy: Stanford University.

Darling-Hammond, L. (2012). Two futures of educational reform: What strategies will improve teaching and learning. *Revue suisse des sciences de l'éducation, 34*(1), 21-38.

Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *The Phi Delta Kappan, 76*(8), 597-604.

Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis,, 23*(1), 57-77.

Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preperation: How well do different pathway prepare teachers to teach. *Journal of Teacher Education, 53*(4), 286-302.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad.* Palo Alto, CA: National Staff Development Council and The School Redesign Network at Stanford University.

Dash, S., Magidin de Kramer, R., O'Dwyer, L., Masters, J., & Russell, M. (2012). Impact of online professional development on teacher quality and student achievement in fifth rgade mathematics. *Journal of Research on Technology in Education, 45*(1), 1-26.

Davis, M. R. (2010). Social networking goes to school. *Educaiton Digest, 76*(3), 14-19.

de Jong, K., Moerbeek, M., & Rien, v. d. (2010). A priori power analysis in longitudinal three-level multilevel models: An example with therapist effects. *Psychotherapy Research, 20*(3), 273-284.

Dee, T. (2004). Are there civic returns to education? *Journal of Public Economics, 88*, 1697-1720.

Dee, T. (2004). Teachers, race, and student achievement in a randomized experiment. *The Review of Economics and Statistics, 86*(1), 195-210.

Delli Carpini, M., & Keeter, S. (1996). *What Americans know about politics and why it matters.* New Haven, CT: Yale University Press.

Desimone, L. (2009). Improving impact studies of teacher' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199.

Desimone, L., & Le Floch, K. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis, 26*(1), 1-22.

Desimone, L., & Long, D. (2010). Teacher effects and the achievement gap: Do teachers and teaching quality influence the achievement gap between black and white and high and low SES students in the early grades. *Teachers College Record, 112*(12), 3024-3073.

Desimone, L., Porter, A., Garet, M., Yoon, K. S., & Birman, B. (2002). Effects of professional development on teachers' instruction: Results from a three-year study. *Educational Evaluation and Policy Analysis, 24*(2).

Desimone, L., Smith, T., & Frisvold, D. E. (2010). Survey measures of classroom instruction: Comparing student and teacher reports. *Educational Policy, 24*(2), 267-329.

Desimone, L., Smith, T., & Phillips, K. (2013). Linking student achievement growth to professional development participation and changes in instruction: A longitudinal study of elementary students and teachers in Title I schools. *Teachers College Record, 115*(5), 1-46.

Dewey, J. (1903). Democracy in education. *The Elementary School Teacher, 4*(4), 193-204.

Discovering Justice. (2014). *Our Programs*. Retrieved April 28, 2014, from Discovering Justice: The James D. St. Clair Court Education Project: http://www.discoveringjustice.org/?p=pgms_afterschool_mocktrial

Drake, C., & Sherin, M. G. (2006). Practicing change: Curriculum adaptation and teacher narrative in the context of mathematics education reform. *Curriculum Inquiry, 36*(2), 153-187.

DuFour, R. (2004). What is a "professional learning community"? *Educational Leadership, 61*(8), 6-11.

Duncan, A. (2009). Elevating the teaching profession. *American Educator*, 3-5.

Duncan, A. (2012, January 10). *Secretary Arne Duncan's Remarks at "For Democracy's Future" Forum at the White House.* Retrieved May 1, 2014, from U.S Department of Education: http://www.ed.gov/news/speeches/secretary-arne-duncans-remarks-democracys-future-forum-white-house

Eaton, D., Brener, N., & Kann, L. K. (2008). Associations of health risk behaviors with school absenteeism. Does having permission for the absence make a difference? *Journal of School Health, 78*(4), 223-229.

Eby, J. W. (1998). *Why service learning is bad.* Retrieved September 25, 2011, from http://glennblalock.org/~gblalock/glennblalock/wiki/uploads/ACSM1110f09/whySLbad.pdf

Engelhardt, L. T., & Steinbrink, J. E. (2001). Teaching students about their civic obligation--Jury Duty. *The Social Studies, 92*(2), 61-64.

Enslin, P., Pendlebury, S., & Tjiattas, M. (2001). Deliberative democracy, diversity, and the challenges of citizenship education. *Journal of Philosophy of Education, 35*(1), 115-130.

Eyler, J., & Giles, D. (1999). *Where's the learning in service-learning.* San Fransisco: Jossey-Bass.

Eyler, J., Giles, D., & Braxton, J. (1997). The impact of service-learning on college students. *Michigan Journal of Community Service Learning, 4*, 5-15.

Facing History and Ourselves. (2014). *Our History*. Retrieved April 28, 2014, from Facing History and Ourselves: https://www.facinghistory.org/get-to-know-us/history

Farkas, S., & Duffet, A. M. (2010). *High schools, Civics, and Citizenship.* Washington D.C.: American Enterpise Institute.

Fine, M. (2009). Postcards from metro america: Reflections on youth participatory action research for urban justice. *Urban Review, 41*(1), 1-6.

Finlay, A., Wray-Lake, L., & Flangan, C. (2010). Civic engagement during the transition to adulthood: Development opportunities and social policies at a critical juncture. In L. Sherrod, J. Torney-Purta, & C. Flanagan, *Handbook of Research on Civic Engagement* (pp. 277-306). Hoboken, NJ: John Wiley & Sons.

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

Fitchett, P., & Heafner, T. L. (2010). A national perspective on the effects of high-stakes testing and standardization on elementary social studies marginalization. *Theory and Research in Social Education, 38*(1), 114-130.

Flanagan, C. (2004). Volunteerism, leadership, political socialization, and civic engagement. In R. Lerner, & L. Steinberg, *Handbook of Adolescent Psychology* (2nd ed., pp. 721-745). Hoboken, NJ: Wiley.

Flanagan, C., & Levine, P. (2010). Civic engagement and the transition to adulthood. *The Future of Children, 20*(1), 159-179.

Forte, A., Humphreys, M., & Park, T. (2012). Grassroots professional development: How teachers use Twitter. *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 106-113.

Foster-Fishman, P. G., Law, K. M., Lichy, L., & Aoun, C. (2010). Youth ReACT for social change: A method for youth participatory action research. *American Journal of Community Psychology, 46*, 67-83.

Frazier, P., Tix, A., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology, 51*(1), 115-134.

Furco, A. (1996). Service-Learning: A Balanced Approach to Experiential Education. Corportation for National Service Learn and Serve America: Higher Education Cooperative Education Association.

Garcia, E. E., Jensen, B. T., & Scribner, K. (2009). English language learners represent a growing proportion of US. students. To meet these students' needs, we must understand who they are. *Educational Leadership*, 9-13.

Gardella, J. A., Mililo, T., Sinha, G., Oh, G., & Manns, D. C. (2009). Linking advanced public service learning and community participation with enviromental analytical chemistry. In D. Redlwask, & T. Rice (Eds.), *Civic Service: Service Learning with State and Local Government Partners* (pp. 98-118). San Francisco, CA: Jossey-Bass.

Garet, M., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., . . . Sztenjnberg, L. (2008). *The impact of two professional development interventions on early reading instruction and achievement (NCEE 2008-4030).* Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S Department of Education.

Garet, M., Porter, A. C., Desimone, L., Birman, B., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915-945.

Garet, M., Wagner, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., . . . Warner, E. (2011). *Middle School Mathematics Professional Development Study.* U.S. Department of Education, Washington, D.C.: National Center for Education Evaluation and Regional Assistance.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York, NY: Cambridge University Press.

Gelman, A., & Weakliem, D. (2009). Of beauty, sex, and power: Statistical challenges in estimating small effects. *American Scientist, 97*(4), 310-316.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectivness, 5*, 189-211.

Gelman, A., Shor, B., Bafumu, J., & Park, D. (2007). Rich state, poor state, red state, blue state: What's the matter with Connecticut. *Quaterly Journal of Political Science, 2*, 345-367.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and ereference 11.0 update (4th ed.).* Boston, MA: Allyn & Bacon.

Gersten, R., Dimino, J., Jayanthi, M., Kim, J. S., & Santoro, L. E. (2010). Teacher study group: Impact on the professional development model on reading instruction and student outcomes in first grade classrooms. *American Educational Research Journal, 47*(3), 694-739.

Gibson, N. M., & Olejnik, S. (2003). Treatment of missing data at the second level of hierarchical linear models. *Educational and Psychological Measurement, 63*, 204-238.

Gilens, M., & Page, B. (2014). Testing theories of American politics: Elites, interest groups, and average citizens. *Perspectives on Politics, 12*(3), 564-581.

Giles, C., & Hargreaves, A. (2006). The sustainability of innovative schools as learning organizations and professional learning communities during standardized reform. *Educational Administration Quarterly, 42*(1), 124-156.

Gingold, J. (2013). *Building an evidence-based practice of action civics: The current state of assessments and recommendations for the future.* Medford, MA: The Center for Information and Research on Civic Learning & Engagement.

Ginsburg, A., & Chudowsky, N. (2012). *Time for learning: An exploratory analysis of NAEP Data.* Washington, D.C.: National Assessment Governing Board.

Ginwright, S., & Cammarota, J. (2002). New terrain in youth development: The promise of a social justice approach. *Social Justice, 29*(4), 82-95.

Glazerman, S., Dolfin, S., Bleeker, M., Johnson, A., Isenbeg, E., Lugo-Gil, J., . . . Ali, M. (2008). *Impact of a Comprehensive Teacher Induction Program.* U.S. Department of Education, Washington, D.C.: National Center for the Evaluation of Education and Regional Assistance.

Godsay, S., Henderson, W., Levine, P., & Littenberg-Tobias, J. (2012). *State Civic Education Requirements.* Medford, MA: Center for Information on Civic Learning and Engagement.

Godsay, S., Kawashima-Ginsberg, K., Kiesa, A., & Levine, P. (2012). *That's not Democracy: How Out-of-School Youth Engage in Democracy and What Stands in Their Way.*

Medford, MA: The Center for Information and Research on Civic Learning and Engagement.

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National Board Certification as a signal of effective teaching. *American Educational Research Journal, 89*(1), 134-150.

Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification and student achievement. *Educational Evaluation and Policy Analysis, 22*(2), 129-145.

Goldhaber, D., & Hansen, M. (2012). Is it just a bad class? Assessing the long-term stability of estimated teacher performance. *Economica, 80*, 589-612.

Gordon, R. (2008). *Comments on Use of Value-Added Methodology to Improve Student Achievement (November 2008).* National Academy of Sciences Panel.

Green, E. (2014). *Building a better teacher: How teaching works (and how to teach it to everyone).* New York: W.W Norton & Company.

Greene, W. H. (2011). *Econometric Analysis* (7th edition ed.). Upper Saddle River, New Jersey: Prentice Hall.

Greenhow, C., Dexter, S., & Riedel, E. (2006). Methods for evaluating online, resource-based learning environments for teachers. *Journal of Computing in Teacher Education, 23*(1), 5-12.

Greenwald, R., Hedges, L., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361-396.

Grossman, P., Cohen, J., Ronfeldt, M., & Brown, L. (2014). The test matters: The relationship between classroom observation scores and teacher value added on multiple types of assessment. *Educational Researcher, 43*(6), 293-303.

Grossman, P., Loeb, S., Cohen, J., Hammerness, K., Wycoff, J., Boyd, D., & Lankford, H. (2010). *Measure for Measure: The Relationship between Measures of Instructional Practice in Middle School English Language Arts and Teachers' Value-Added Scores (Working Paper 20).* Washington, D.C.: National Center for Analysis of Longitudinal Data in Education Research.

Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a theory of teacher community. *Teachers College Record, 103*, 942-1012.

Guarino, C., Hamilton, L. S., Lockwood, J., Rathbun, A., & Hausken, E. (2006). *Teacher qualifications, instructional practices, and reading and mathematics gains of*

Kindergartners. Washington, D.C.: National Center for Education Statistics: U.S. Department of Education.

Guiney, E. (2001). Coaching isn't just for athletes: The role of teacher leaders. *The Phi Delta Kappan, 82*(10), 740-743.

Guskey, T. (2003). What makes professional development effective? *The Phi Delta Kappan, 84*(10), 748-750.

Gutmann, A., & Thompson, D. (2004). *Why Deliberative Democracy.* Princeton, N.J.: Princeton University Press.

Hambleton, R., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory.* Sage Publications: Newbury Park, CA.

Hamilton, L., & Martinez, J. F. (2007). What can the TIMSS Surveys tells us about the mathematics reforms in the United States during the 1990s. In T. Loveless (Ed.), *Lessons learned: What international assessments tell us about math achievement* (pp. 127-174). Washington, D.C.: The Brookings Institution.

Hamilton, L., McCaffrey, D. F., Stecher, B., Klein, S. P., Robyn, A., & Bugliari, D. (2003). Studing large-scale reforms of instructional practice: An example from mathematics and science. *Educational Evaluation and Policy Analysis, 25*(1), 1-29.

Hanushek, E. (2011). The economic value of high quality teaching. *Economics of Education Review, 30*, 466-479.

Hanushek, E., & Rivkin, S. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review, 100*(2), 257-271.

Hargreaves, A., & Fullan, M. (2012). *Professional Capital: Transforming Teaching in Every School.* New York: Teachers College Press.

Harris, K. R., Lane, K. L., Graham, S., Driscoll, S., Sandmel, K., Brindle, M., & Schatschneider, C. (2012). Practice-based professional development for self-regulated strategies development in writing: A randomized control trial. *Journal of Teacher Education, 63*, 103-119.

Hart, D., & Atkins, R. (2002). Civic competance in urban youth. *Applied Developmental Science, 6*(4), 227-236.

Hart, D., Donnelly, T. M., Younss, J., & Atkins, R. (2007). High school community service as a predictor of adult voting and volunteering. *Amercian Educational Research Journal, 44*(1), 197-219.

Harwell, M., & LeBeau, B. (2010). Student eligibility for a free lunch as an SES measure in education research. *Educational Researcher*, 120-131.

Heafner, T., & Fitchett, P. G. (2012). National trends in elementary instruction: Exploring the role of social studies curriculum. *The Social Studies, 103*, 67-72.

Hess, D. (2002). Discussing controversial public issues in secondary school social studies classrooms: Learning from skilled teachers. *Theory and Research in Social Education, 30*(1), 10-41.

Hess, D. (2009). *Controversy in the Classroom: The Democratic Power of Discussion.* New York: Routledge.

Hess, D., & Zola, J. (2012). Professional development as a tool for improving civic education. In D. Campbell, M. Levinson, & F. Hess, *Making civics count: Citizenship education for a new generation* (pp. 183-206). Cambridge, MA: Harvard University Press.

Hiebert, J., & Stigler, J. W. (2000). A proposal for improving classroom teaching: Lessons from the TIMSS Video Study. *The Elementry School Journal, 101*(1), 3-20.

Hill, C., Bloom, H., Black, A. R., & Lipsey, M. W. (2007). *Empirical Benchmarks for Interpreting Effect Sizes in Research.* New York, NY: MRDC Working Papers on Research Methodology.

Hill, H. (2009). Fixing teacher professional development. *The Phi Delta Kappan, 90*(7), 470-476.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Hill, H., Kapitula, L., & Umland, K. (2010). A validity argument approach to evaluating teacher value-aded scores. *American Educational Research Journal, 48*(3), 794-831.

Hillygus, D. S. (2005). The MISSING LINK: Exploring the Relationship between higher education and political engagement. *Political Behavior, 27*(1), 25-47.

Hinde, E. (2008). Civic education in the NCLB Era: The contested mission of elementary and middle schools. *Journal of Curriculum and Instruction, 2*(1), 74-86.

Ho, A. (2012). Variety and drift in the functions and purposes of assessment in K-12 Education. *The Gordon Commission of the Future of Assessment in Education: Works in Progress*, 10-13.

Hollis, S. (2004). Blaming me, blaming you: Assessing service-learning and participants' tendency to balme the victim. *Sociologicall Spectrum: Mid-South Sociological Association, 24*(5), 575-600.

Homana, G. (2009). *Communities of practice for the development of adolescent civic engagement: An empirical study of their correlates in Australia and the United States.* University of Maryland, College Park, MD: (Unpublished dissertation).

Horton, N., & Kleinman, K. (2007). Much ado about nothing: A comparison of missing data methods and software. *The American Statistician, 61*(1), 79-90.

Howard, G., Schmek, R., & Bray, J. H. (1979). Internal invalidity in studies employing self-report instruments: A suggested remedy. *Journal of Educational Measurement,, 16*(2), 129-135.

Hox, J. (1998). Multilevel modeling: When and why. In R. Balderjahn, R. Mathar, & M. (. Schader, *Classficiation, Data Analysis, and Data Highways* (pp. 147-154). New York, NY: Springer Verlag.

Hox, J. J. (2013). Multilevel regression and multilevel structural equation modeling. In *Oxford handbook of quantitative methods* (Vol. 2, pp. 281-294).

Huffman, D., Thomas, K., & Lawrenz, F. (2003). Relationship between professional development, teachers' instructional practices, and the achievement of students in sciience and mathematics. *School Science and Mathematics, 103*(8), 378-387.

Hur, J. W., & Brush, T. A. (2009). Teacher participation in online communities: Why do teachers want to participate in self-generated online communities of K-12 teachers. *Journal of Research on Technology in Education, 41*(3), 279-303.

Ingels, S., Glennie, E., Lauff, E., & Wirt, J. (2012). *Trends among young adults over three decades.* Washington D.C.: U.S. Department of Education, National Center for Education Statistics. Retrieved April 27, 2014, from http://nces.ed.gov/pubsearch

Ingersoll, R. M. (2001). The realities of out-of-field teaching. *Educational Leadership, 58*(8), 42-45.

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational Leadership, 60*(8), 30-33.

Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability politics and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record, 106*(6), 1258-1287.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med, 2*(8), 0696-0701.

Jaccard, J., Wan, C., & Turrisi, R. (1990). The detection and interpretation of interaction effects between continuous variable in multiple regression. *Multivariate Behavioral Research, 25*(4), 467-478.

Jackson, P. S., Hinde, E. R., & Hass, N. S. (2008). Teaching the constitution to twenty-first century students. *Social Education*, 377-378.

Jacobs, V., Franke, M. L., Carpenter, T., & Levi, L. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education, 38*(3), 258-288.

Jefferson, T. (1982). *Notes on the State of Virgnia.* New York: W.W. Norton & Company.

Jones, L. (1996). A history of the National Assessment of Education Progress and some questions about its future. *Educational Researcher, 25*(7), 15-22.

Kahne, J., & Middaugh, E. (2008). *Democracy for Some: The Civic Opportunity Gap in High School.* Medford, MA: The Center for Information and Research on Civic Learning and Engagement.

Kahne, J., & Sporte, S. (2008). Developing citizens: The impact of civic learning opportunities on students' commitment to civic participation. *American Educational Research Journal, 45*(3), 738-766.

Kahne, J., & Westheimer, J. (1996). In service of what? The politics of service learning. *Phi Delta Kappan, 77*, 592-599.

Kahne, J., & Westhimer, J. (2006). The limits of political efficacy: Educating citizens for a democratic society. *PS: Political Science and Politics*(2), 289-296.

Kahne, J., Chi, B., & Middaugh, E. (2006). Building social capital for civic and political engagement: The potential of high-school civics courses. *Canadian Journal of Education, 29*(2), 387-402.

Kane, T., & Steiger, D. O. (2008). *Estimating Teacher Impacts on Student Achievement.* Cambridge, MA: NBER Working Paper Series.

Kane, T., Taylor, E. S., Tyler, J. H., & Wooten, A. (2010). Identifying effective classroom practices using student achievement data. *The Journal of Human Resources, 46*(3), 587-613.

Kawashima-Ginsberg, K. (2012). *Summary of Findings from the Evaluation of iCivics' Drafting Board Intervention.* Medford, MA: The Center for Information & Research on Civic Learning and Engagement.

Kawashima-Ginsberg, K. (2013). *Do Discussion, Debate, and Simulations Boost NAEP Civics.* Medford, MA: CIRCLE Fact Sheet with support from S.D. Bechtel, Jr. Foundation.

King, M. (2002). Professional development to promote schoolwide inquiry. *Teaching and Teacher Education, 18*(3), 243-257.

Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86.

Koliba, C., & Gajda, R. (2009). "Communities of Practice" as an analytical construct: Implications for theory and practice. *International Journal of Public Administration, 32*, 97-135.

Kraft, M. A., & Papay, J. P. (2014). Can professional environment in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis, 36*(4), 476-500.

Krutka, D., & Milton, M. K. (2013). The enlightenment meets Twitter: Using social media in the social studies classroom. *Ohio Social Studies Review, 50*(2), 22-29.

Kukla-Acavedo, S. (2009). Do teacher charachteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review, 28*, 49-57.

Kupermintz, H. (2003). Teacher effects and teacher effectivness: A validity investigation of the Tennessee value added assessment system. *Educational Evaluation and Policy Analysis, 25*(3), 287-298.

Langston, K., & Jennings, M. K. (1968). Political socialization and the high school civics curriculum in the United States. *American Review of Political Science, 62*(3), 852-867.

Lave, J., & Wegner, E. (1991). *Situated learning: Legitimate peripheral participation.* New York, NY: Cambridge University Press.

Le, V.-N., Lockwood, J., Stecher, B., Hamilton, L. S., & Martinez, J. F. (2009). A longitduinal investigation of the relationships between teachers' self reports of reform-oriented instruction and mathematics and science achievement. *Educational Evaluation and Policy, 31*(3), 200-220.

Leana, C. (2011). The missing link in school reform. *Stanford Social Innovation Review*, 30-35.

Lee, V., & Smith, J. (1996). Collective responsibility for learning and its effects on gains in achievement for early secondary school students. *American Journal of Education, 104*(2), 103-147.

Lee, V., & Smith, J. (1997). High school size: Which works best for whom. *Educational Evaluation and Policy Analysis, 19*, 205-227.

Levine, P. (2013). *We are the ones we have been waiting for: The promise of civic renewal in America.* New York, NY: Oxford University Press.

Levine, P. (2013). *What the NAEP Civics Assessment Measures and How Students Perform.* Medford, MA: Center for Information and Research on Civic Learning and Engagement.

Levine, P., & Higgins-D'Alessando, A. (2010). Youth civic engagement: Normative issues. In L. Sherrod, J. Torney-Purta, & C. Flanagan (Eds.), *Handbook of Research on Civic Engagement in Youth* (pp. 115-138). Hoboken, NJ: John Wiley & Sons.

Levinson, M. (2010). The civic empowerment gap: Defining the problem and locating solutions. In L. Sherrod, C. Flanagan, & J. Torney-Purta, *Handbook of Research on Civic Engagement in Youth* (pp. 331-362). Hoboken, NJ: John Wiley $ Sons.

Levinson, M. (2012). *No Citizen Left Behind.* Cambridge, MA: Harvard University Press.

Lewis, C. C., Perry, R., Friedkin, S., & Roth, J. (2012). Improved teaching does improve teachers: Evidence from lesson study. *Journal of Teacher Education, 63*(5), 368-375.

Li, F., Duncan, T. E., Harmer, P., Acock, A., & Stoolmiller, M. (1998). Analyzing measurement models of latent variables through multilevel confirmatory factor analysis and hierarchical linear modeling approaches. *Structural Equation Modeling: A Multidisciplinary Journal, 5*(3), 294-306.

Lieberman, A. (2000). Networks as learning communities: Shaping the future of teacher development. *Journal of Teacher Education, 51*(3), 221-227.

Lieberman, A., & McLaughlin, M. W. (1992). Networks for educational change: Powerful and problematic. *The Phi Delta Kappan, 73*(9), 673-677.

Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis, 15*(2), 129-151.

Little, J. W. (2006). *Professional community and professional development in the learning-centered school.* Washingotn, DC: National Education Association.

Little, R. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198-1201.

Little, R., & Rubin, D. (2002). *Statistical Analysis with Missing Data (2nd ed).* New York: Wiley.

Lock, J. (2006). A new image: Online communities to facilitate teachers' professional development. *Journal of Technology and Teacher Education, 14*(4), 663-678.

Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences. *American Journal of Education, 106*(4), 532-575.

Loveless, T. (2014). *What do we know about professional development?* Retrieved February 21, 2014, from EducationNext: http://www.educationnext.org/what-do-we-know

Lubienski, S. T., Lubienski, C., & Crane, C. C. (2008). Achievement differences and school type: The role of school climate, teacher certiciation, and instruction. *American Journal of Education, 115*, 97-138.

Lubinski, D., & Humphreys, L. G. (1990). Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ('synergistic") relation between spatial and mathematical ability. *Psychological Bulletin, 107*(3), 385-393.

Lumpe, A. T. (2007). Research-based professional development: Teachers engaged in professional learning communities. *Journal of Science Teacher Education, 18*, 125-128.

MacNeil, A. J., Prater, D., & Busch, S. (2009). The effects of school culture and climate on student achievement. *International Journal of Leadership in Education, 12*(1), 73-84.

Manning, N., & Edwards, K. (2013). Does civic education for young people increase political participation? A systematic review. *Educational Review*, 1-24.

Marullo, S., & Edwards, B. (2000). From charity to justice: The potential of university-community colloboration for social change. *Amerian Behavioral Scientist, 43*(5), 895-912.

Marzano, R. J. (2000). *What works in classroom instruction.* Aurora, CO: Mid-continent Research for Education and Learning.

Mayer, D. (1999). Measuring instructional practice: Can policymakers trust survey data. *Educational Evaluation and Policy Analysis, 21*, 29-45.

Mayer, D. P. (1998). Do new teaching standards undermine performance on old tests? *Educational Evaluation and Policy Analysis, 20*(2), 53-73.

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist, 59*(1), 14-19.

McDevitt, M., & Chaffee, S. (1998). Second chance political socialization: "Trickle-up" effects of children on parents. In T. Johnson, C. Hays, & S. Hays (Eds.), *Engaging the public: How government and the media can reinvigorate American democracy* (pp. 57-77). Lanham, MD: Rowman and Littlefield.

McDevitt, M., & Kiousis, S. (2006). Deliberative learning: An evaluative approach to interactive civic education. *Communication Education, 55*(3), 247-264.

McDevitt, M., & Kiousis, S. (2006). *Experiments in Political Socialization: Kids Voting USA as a Model of Civic Education Reform.* University of Maryland, College Park, MD: CIRCLE Working Paper 49.

McLaughlin, M. W., & Talbert, J. E. (2006). *Building School-Based Teacher Learning Communities.* New York, NY: Teachers College Press.

Moely, B., & Ilustre, V. (2013). Stability and change in the development of college students' civic attitudes, knowledge, and skills. *Michigan Journal of Community Service Learning, 19*(2), 21-35.

Monk, D. (1994). Subject area preparation of secondary mathematrics and science teachers and student achievement. *Economics of Education Review, 13*(2), 125-145.

Mullens, J. (1998). Validating teacher responses on instructional practice self-report surveys: Findings from a recent field test. *Paper presented at the annual conference of the Association for Public Policy Analysis and Management: October 31, 1998.* New York, NY.

Mullis, I., Bohrnstedt, G., Preuschoff, A., Reyes, I. d., Stancavage, F., & Martin, M. O. (2012). *Examining the NAEP Achievement in Relation to School Testing Conditions.* San Mateo, CA: American Institutes for Research: NAEP Validity Studies.

Mutz, D. (2002). The consequences of cross-cutting networks for political participation. *American Journal of Political Science, 46*(4), 838-855.

Myers-Lipton, S. (1998). Effect of a comprenhensive service-learning program on college student's civic responsibility. *Teaching Sociology, 26*(4), 243-258.

NAEP. (2009). *Welcome to NAEP 2010! Facts for School NAEP Coordinators.* Washington, D.C.: National Assessment of Educational Progress. Retrieved December 5, 2013, from https://www.mynaep.com/dms/published/myschool/setup/doc_2009105_Gr_4_and_8_Welcome_to_NAEP_2010_final.pdf

NAEP. (2011, January 24). *How the Samples of Schools and Students Are Selected for the Main Assessments (State and National)*. Retrieved from National Assessment of Educational Progress: http://nces.ed.gov/nationsreportcard/about/nathow.asp

NAEP. (2011). *NAEP Questions Tool*. Retrieved January 9, 2014, from National Center for Education Statistics: http://nces.ed.gov/nationsreportcard/itmrlsx/search.aspx

NAEP. (2011). *NAEP Technical Documentation.* Retrieved December 5, 2013, from http://nces.ed.gov/nationsreportcard/tdw/instruments/

National Assessment Governing Board. (2010). *Civics Frameworks for the 2010 National Assessment of Educational Progress.* Washington, D.C.: US Department of Education.

National Center for Education Statistics. (2010). *Civics 2010: National Assessment of Educational Progress at Grades 4, 8, 12.* Washington, D.C.: US Department of Education.

National Center for Education Statistics. (2010). *State Nonfiscal Public Elementary/Secondary Education Surevey.* Washington, D.C.: U.S. Department of Education.

National Center for Learning Disabilities. (2014, March 23). *What Is an IEP?* Retrieved from National Center for Learning Disabilities: http://www.ncld.org/students-disabilities/iep-504-plan/what-is-iep

NCSS. (2002). *National Standards for Social Studies Teachers.* Retrieved November 1, 2014, from http://www.socialstudies.org/standards/teacherstandards

NCTM. (2014). *Principles to Action: Ensuring Mathmetical Success for All.* Retrieved November 1, 2014, from http://www.nctm.org/standards/default.aspx?id=58

Niemi, N. S., & Niemi, R. G. (2007). Partisanship, participation, and political trust as taught (or not) in high school history and government classes. *Theory & Research in Social Education, 35*(1), 32-61.

Niemi, R. G., & Junn, J. (1998). *Civic education: What makes students learn.* New Haven, CT: Yale University Press.

Noble, A., & Littenberg-Tobias. (2014). A lifelong classroom: Social studies educators' engagement with professional learning networks on Twitter. *UMASS Dartmouth TRED Conference.* Dartmouth, MA.

Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237-257.

Oakes, J., Ormseth, T., Bell, R., & Camp, P. (1990). *Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science.* Santa Monica, CA: The RAND Corporation.

O'Day, J. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review, 72*(3), 293-329.

O'Dwyer, L., Dash, S., Kramer, M. d., Humez, A., & Russell, M. (2010). *e-Learning for Educator Effects of On-Line Professional Development for Teachers and their Students: Findings from Four Randomized Trials.* Boston, MA: Technology and Assesment Study Collaborative, Boston College .

Opfer, V. D., & Pedder, D. (2011). Conceptualising teacher professional learning. *Review of Educational Research, 81*, 376-407.

Ozer, E. J., Ritterman, M., & Wanis, M. (2010). Participatory action research (PAR) in middle school: Opportunities, constraints, and key processes. *American Journal of Community Psychology, 46*, 152-166.

Parke, H., & Coble, C. (1997). Teachers designing curriculum as professional development: A model for transformational science teaching. *Journal of Research in Science Teaching, 34*(8), 773-789.

Pederson, S., & Liu, M. (2003). Teachers' beliefs about issues in the implementation of a student-centered learning environment. *Education Technology Research and Development, 51*(2), 57-76.

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

Peterson, R. (1994). A meta-analysis of Cronbach's Coefficient Alpha. *Journal of Consumer Research*, 381-391.

Pfeffermann, D., Skinner, C., Holmes, D., Goldstein, H., & Rabash, J. (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology),, 60*(1), 23-40.

Pil, F. K., & Leana, C. (2009). Applying organizational research to school reform: The effects of teacher human and social capital on student performance. *Academy of Management Journal, 52*(6), 1101-1124.

Plutzer, E. (2002). Becomming a habitual voter: Intertia, resources, and growth in young adulthood. *American Political Science Review, 96*(1), 41-56.

Pondisco, R., Sewall, G. T., & Stotsky, S. (2013). *Shortchanging the Future: The Crisis of History and Civics in American School No.100.* Boston, MA: Pioneer Insitute Public Policy Research.

Pope, A., Stolte, L., & Cohen, A. (2011). Closing the civic engagement gap: The potential of action civics. *Social Education*, 265-268.

Popkin, S., & Dimock, M. (1999). Policak knowledge and civic competence. In I. E. (eds), *Civic Competence and Democratic Institutions* (pp. 117-146). University Park, PA: Pennsylvania State University Press.

Project Citizen. (2014). *Project Citizen: A Portfolio-Based Civic Education Program.* Retrieved April 25, 2014, from Center for Civic Education: http://new.civiced.org/resources/publications/center-brochures

Puma, M., Olsen, R., Bell, S., & Price, C. (2009). *What to do when data are missing in group randomized controlled trials.* Washington, D.C.: Institute of Education Sciences, National Center for Educational Evaluation and Regional Assistance.

Putnam, R. D., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4-15.

Putnam, R. T. (2001). *Bowling alone: The collapse and revival of American community.* Simon & Schuster.

Quigley, C. (2007). *Response to the May 16, 2007 Release of the 2006 National Assessment of Education Progress Civics Report Card.* Calabass, CA: Center for Civic Education.

Raudenbush, S. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods, 2*(2), 173-185.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods.* Thousand Oaks, CA: Sage Publications, Inc.

Raudenbush, S., Bryk, A., & Congdon, R. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Raykov, T., & Macroulides, G. A. (2006). On multilevel reliability estimation from the perspective of structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(1), 130-141.

Remillard, J. (1999). Curriculum materials in mathematics education reform: A framework for examining teachers' curriculum development. *Curriculum Inquiry, 29*(3), 315-342.

Richardson, W. (2006). Combining cognitive interviews and social science surveys. In K. B. (Ed.), *Research methods in social studies education: Contemporary issues and perspectives* (pp. 159-182). Greenwich, CT: Information Age Publishing Company.

Robertson-Kraft, C., & Duckworth, A. (2014). True grit: trait level perseverance and passion for long-term goals predicts effectivness and retention among novice teachers. *Teachers College Record, 116*(3), 1-27.

Rockoff, J. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247-252.

Rogers, A., Stoeckel, J., & Sikali, E. (2012). *NAEP 2010 Civics, U.S. History, and Geography Restricted-Use Data Files.* Washington, D.C.: U.S Department of Education.

Ross, E. W. (1997). The struggle for the social studies curriculum. In E. W. Ross, *The social studies curriculum: Purposes, problems, and possibilities* (pp. 3-19).

Ross, J. A., Cousins, J. B., Gadalla, T., & Hannay, L. (1999). Administrative assignment of teachers in restructuring secondary schools: The effect of out-of-field course responsibility on teacher efficacy. *Educational Administration Quaterly, 35*, 782-804.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, 175-214.

Rowan, B., Correnti, R., & Miller, R. J. (2002). *What large-scale survey research tells us about teacher effects on student achievement: Insights from the "Prospects" study of elementary schools.* Philadelphia, PA: Consortium for Policy Research in Education.

Rubin, B. C. (2012). *Making Citizens: Transforming Civic Learning For Diverse Social Studies Classrooms.* New York: Routledge.

Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record, 107*(9), 1999-2045.

Russell, M., Carey, R., Kleiman, G., & Venable, J. D. (2009). Face-to-face and online professional development for mathematics teachers: A comparitive study. *Journal of Asynchronous Learning Networks, 13*(2), 71-87.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 143-151.

Sanders, D., & McCutcheon, G. (1986). The development of practical theories of teaching. *Journal of Curriculum and Supervision, 2*(1), 50-67.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247-256.

Sato, M., Wei, R. C., & Darling-Hammond, L. (2009). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal, 45*(3), 669-700.

Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical methods in medical research, 8*(1), 3-15.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147-177.

Schlomer, G., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology, 57*(1), 1-10.

Scribner, J. P. (1999). Professional development: Untangling the influence of work context on teacher learning. *Educational Administration Quaterly, 35*(2), 238-266.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs.* New York, NY: Houghton Miffin Company.

Shiller, J. (2013). Preparing for democracy: How community-based organizations build civic engagement among urban youth. *Urban Education, 48*, 69-91.

Shin, Y. (2013). Efficient handling of predictors and outcomes having missing values. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large scale assessment: Background, technical issues, methods of data analysis* (p. 451). Boca Raton, FL: CRC Press.

Silva, D. Y., & Mason, T. C. (2003). Developing pedagogical content knowledge for civics in elemenary teacher education. *Theory and Research in Social Education, 31*(3), 366-397.

Sireci, S., Scarpati, S., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research, 75*(4), 457-490.

Smith, J., & Niemi, R. (2001). Learning history in school: The impact of course work and instructional practices on achievement. *Theory and Research in Social Education, 29*(1), 18-42.

Sondheimer, R. M., & Green, D. P. (2010). Using experiments to estimate the effects of education on voter turnout. *American Journal of Political Science, 54*(1), 174-189.

Spillane, J., & Louis, K. S. (2002). School improvement processes and practices: Professional learning for building instructional capacity. *Yearbook of the National Society for the Study of Education*, 83-104.

Staiger, D. O., & Rockoff, J. E. (2010). Searching for effective teachers with imperfect information. *The Journal of Economic Perspectives, 24*(3), 97-117.

Stamps, D. (1997). Communities of practice. *Training, 34*(2), 34.

Stapleton, L. M. (2009). The incorporation of sample weights into multilevel structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(4), 475-502.

Stauss, H. (2013, 25 July). *National Civics, History Tests to Disappear.* Retrieved December 6, 2013, from Heartland Institue: http://news.heartland.org/newspaper-article/2013/07/25/national-civics-history-tests-disappear

Storandt, B., Dossin, L. C., & Lacher, A. P. (2012). Toward an understanding of what works in professional development for online instructors: The case of PBS Teacherline. *Journal of Asynchronous Learning Networks, 16*(2), 121-162.

Strategic Data Project. (2012). *SDP Human Capital Diagnostics: Los Angeles Unified School District.* Cambridge, MA: Center for Education Policy Research: Harvard University.

Strizek, G., Pittsonberger, J., Riordan, K., Lyter, D., & Orlofsky, G. (2006). *Characteristics of Schools, Districts, Teachers, Principals, and School Libraries in the United States: 2003-04 Schools and Staffing Survey (NCES 2006-313).* Washington, D.C.: US Government Printing Office: U.S. Department of Education, National Center for Education Statistic.

Stronge, J., Ward, T., Tucker, P., & Hindman, J. (2007). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personal Evaluation in Education, 20*, 165-184.

Supovitz, J., & Turner, H. (2000). The effects of professional development on science teaching practices. *Journal of Research in Science Teaching, 37*(9), 963-980.

Syvertesen, A., Stout, M. D., Flanagan, C., Mitra, D. L., Oliver, M. B., & Sundar, S. S. (2009). Using elections as teachable moments: A randomized evaluation of the Student Voices civic education program. *American Journal of Education, 116*(1), 33-67.

The National Task Force on Civic Learning and Democratic Engagement. (2012). *A Crucible Moment: College Learning and Democracy's Future.* Washington, D.C.: Association of American Colleges & Universities.

Theall, K., Schribner, R., Broyles, S., Yu, Q., Chotalia, J., Simonsen, N., . . . Carlin, B. (2011). Impact of small group size on neighborhood influences in multilevel models. *Journal of Epidemiology and Community Health, 65*(8), 688-695.

Thomas, S., & Heck, R. (2001). Analysis of large-scale secondary data in higher education research: Potential perils associated with complex sampling desings. *Research in Higher Education, 43*(5), 517-540.

Tolo, K. (1998). *"We the People...Project Citizen" Promoting Citizenship in Classrooms and Communities: A Report by the Policy Research Project on "An Assessment of Project Citizen" Policy Research Project Report Number 129.* Austin, TX: Lyndon B. Johnson School of Public Affairs, University of Texas.

Tolo, K. (1999). *The Civic Education of American Youth: From State Policies to School District Practices (Policy Research Project, Number 133).* Austin, TX: Lyndon B. Johnson School of Public Affairs, University of Texas.

Torney-Purta, J. (2002). The school's role in developing civic engagement: A study of adolescents in twenty-eight countries. *Applied Development Science, 6*(4), 203-212.

Torney-Purta, J., & Amadeo, J.-A. (2003). A cross-national analysis of political and civic involvement among adolescents. *PS: Political Science and Politics, 36*(2), 269-274.

Torney-Purta, J., Richardson, W. K., & Barber, C. H. (2005). Teachers' educational experience and confidence in relation to students' civic knowledge across countries. *Internal Journal of Citizenship and Teacher Education, 1*(1), 32-57.

U.S. Census Bureau. (2009). *Current Population Survey, November 2008 Civic Engagement Supplement File.* Washington, D.C.: Government Printing Office.

U.S. Census Bureau. (2014). *Current Population Survey, Aggregated November Civic Engagement Supplement File (2008-2013).* Washington, DC: Government Printing Office. Retrieved from http://www.census.gov/apsd/techdoc/cps/cpsnov10c.pdf

U.S. Department of Education. (2009). *Institute of Education Sciences, National Center for Statistics, National Assessment of Educational Progress (NAEP), 1990,2000,2005, 2009.* Washington, D.C.: U.S. Department of Education.

United States Department of Agriculture Food and Nutritition Services. (2014, March 24). *National School Lunch Program Fact Sheet.* Retrieved from http://www.fns.usda.gov/sites/default/files/NSLPFactSheet.pdf

Valli, L. (1997). Listening to other voices: A description of teacher refleciton in the United States. *Peabody Journal of Education, 72*(1), 67-88.

Van Buuren, S. (2011). Multiple Imputation of Multilevel Data. In *Handbook of advanced multilevel analysis* (pp. 173-196).

Verba, S., Burns, N., & Schlozman, K. L. (1997). Knowing and caring about politics: Gender and political engagement. *The Journal of Politics, 59*(4), 1051-1072.

Verba, S., Schlozman, K. L., & Brady, H. (1995). *Voice and Equality: Civic Voluntarism in American Politics.* Cambridge, MA: Harvard University Press.

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education, 24*, 80-91.

Walker, A., Recker, M., Robertshaw, M. B., Olsen, J., Leary, H. Y., & Sellers, L. (2011). Integrating technology and problem-based learning: A mixed methods study of two teacher professional development designs. *The Interdisciplinary Journal of Problem-Based Learning, 5*(2), 70-94.

Wang, J. (1998). Opportunity to learn: The impacts and policy implications. *Educational Evaluation and Policy Anaysis, 20*(3), 137-156.

Wayman, J. C., & Stringfield. (2006). Technology-supported involvement of entire faculties in examination of student data for instructional improvement. *American Journal of Education, 112*(4), 549-571.

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher, 37*(8), 469-479.

Wayne, A., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research,, 73*(1), 89-122.

Weingarten, R. (2010). The professional educator a new path forward: Four apporaches to quality teaching and better schools. *American Educator*, 36-39.

Wenger, E. (1998). *Communities of Practice.* Cambridge, MA: Cambridge University Press.

Wenglinsky, H. (2001). *Teacher classroom practices and student performance: How schools can make a difference.* Princeton, NJ: Educational Testing Service.

Wenzlaff, T., & Wieseman, K. (2004). Teachers need teachers to grow. *Teacher Education Quaterly, 31*, 113-124.

Westheimer, J., & Kahne, J. (2004). What kind of citizen?: The politics of educating for democracy. *American Educational Research Journal, 41*(2), 237-269.

Wilson, S., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research on Education, 24*, 173-199.

Wilson, S., & Wineburg, S. S. (1993). Wrinkles in time and place: Using performance assessments to understand the knowledge of history teachers. *American Educational Research Journal, 30*(4), 728-769.

Wooldridge, J. M. (2009). *Introductory Econometrics.* Mason, OH: South-Western Cengage Learning.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57-67.

Wu, M. (2005). The role of plausible values in large scale surveys. *Studies in Educational Evaluation, 31*, 114-128.

Yates, M., & Youniss, J. (1998). Community service and political identity: Development in Adolescence. *Journal of Social Issues, 54*(3), 495-512.

Yeh, S. S., & Ritter, J. (2009). The cost effectivness of replacing the bottom quartile of novice teachers through value-added teacher assessment. *Journal of Education Finance, 34*(4), 426-451.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement.* U.S. Department of Education, Washington, D.C.: National Center for Educational Evaluation and Regional Assistance.

Zhang, D. (2005). *A monte carlo investigation of robustness to nonnormal incomplete data of multilevel modeling.* College Station, TX: Doctoral Dissertation, Texas A&M University.

Zhang, T., Torney-Purta, J., & Barber, C. (2012). Students' conceptual knowledge and process skills in civic education: Identifying cognitive profiles and classroom correlates. *Theory and Research in Social Education, 40*(1), 1-34.

Appendix A: Cognitive Interview Protocol on the NAEP Teaching Practices in Social Studies Measures

Background

The purpose of my dissertation is to study the relationships between teacher professional learning, interactive civics activities, and students' civic skills and knowledge.

I'm using the NAEP civics assessment data in my dissertation. NAEP civics is a national assessment of students' civic skills and knowledge that is administered to a nationally representative sample of 8th graders approximately every 4 years. NAEP also collects survey data from social studies teachers about their educational background, participation in professional development.

I reached out to you because I want to better understand how social studies teachers might have answered the background survey. It's fine if you don't teach civics. I just want to get a better sense of how social studies teachers teacher so I can appropriately use the data in my analysis.

1. Do you have any questions about the background of this survey?
2. Is there anything you want me to go into more detail about?

Definition of "Interactive Civics"

Please Read Over the Following Definition

Interactive civics describes classroom activities that are:

a. *Student-centered, meaning that students, not teachers, are responsible for engaging in their learning on their own and with their peers*
b. *Requires student to communicate with others (either written or spoken)*
c. *Engage students in authentic forms of civic participation that expose students to the types of civic activities that they might engage in as adults (e.g., voting, discussing issues, volunteering for organizations, working with others to solve community problems)*

Interactive civics <u>do not</u> have to occur within a formal civics class to qualify as an interactive civics activities.

1. Do you have any questions about this definition? Is everything clear and easy to understand?

2. Based on this definition, have you used these types of activities in your classroom?

3. [If yes] How often would you use these types of activities in your classroom?

NAEP Teacher Classroom Practice Items

Now I'm going to ask you the read over the instructional activities that are listed on the NAEP Teacher Survey. This is not a test. These are just examples of activities that teachers might do. It's okay if you don't do any of these in your own classroom. I just want to get a sense of what teachers might see as an interactive civics activity.

Please check off the activities that could be examples of "interactive civics".

- ❑ Ask students to complete a worksheet
- ❑ Give a lecture to the class about social studies
- ❑ Have students participate in debates or panel discussions
- ❑ Have students participate in mock trials, role-playing, or dramatization
- ❑ Have students write letters to state an opinion or solve a community problem
- ❑ Have visitors from your community meet with the class to discuss important events and ideas
- ❑ Have students participate in community volunteer projects or services
- ❑ Have students access information through the Internet for use in the classroom
- ❑ Discuss current events
- ❑ Use student government
- ❑ Give students social studies homework
- ❑ Tests with multiple choice true/false or matching type questions
- ❑ Tests with fill-in the blank questions
- ❑ Paragraph-length written responses about what students have read
- ❑ Extended essays/papers on assigned topics
- ❑ Individual projects
- ❑ Individual presentations
- ❑ Group projects
- ❑ Group presentations

1. Now that you've read over the list of activities, what activities do you think might meet the definition of interactive civics?

2. Why did you choose those activities? What about them made them interactive?

3. Were there any activities that you didn't choose that you felt were borderline? What about them made them borderline activities?

4. Were there any interactive activities that were missing that you felt should have been included?

Frequency of Practices

1. Look back at the activities you check off, how often do you would do each of these activities in your classroom in a particular year?

2. Why might you do some activities more than others? What factors might make you more likely to do an activity more often than others?

4. How much do you think other social studies teachers in your school use these activities?

3. Think about a teacher who is in the "middle" of the spectrum of how often they would use interactive activities in social studies?

- **a.** Which, if any, of the activities would they do every day?
- **b.** Which, if any, of the activities would they do once a week?
- **c.** Which, if any, of the activities would do they do once or twice a month?
- **d.** Which, if any, of the activities would they do once or twice a year?

Intensity of Activities

In this study, intensity is defined in terms of impact on student learning. Activities with different levels of intensity may need to be used more or less frequently in order to have the same impact on students.

1. The following table has a list of all the activities on the NAEP teacher questionnaires. For each activity you identified as interactive, mark how "intense" each of the activities would be for your students if you were to use it "once or twice a month"

	0 No impact on student learning	1	2	3 High impact on student learning
Ask students to complete a worksheet				
Give a lecture to the class about social studies				
Have students participate in debates or panel discussions				
Have students participate in mock trials, role-playing, or dramatization				
Have students write letters to state an opinion or solve a community problem				
Have visitors from your community meet with the class to discuss important events and ideas				
Have students participate in community volunteer projects or services				
Have students access information through the Internet for use in the classroom				
Discuss current events				
Use student government				
Give students social studies homework				
Tests with multiple choice true/false or matching type questions				
Tests with fill-in the blank questions				
Paragraph-length written responses about what students have read				
Extended essays/papers on assigned topics				
Individual projects				
Individual presentations				
Group projects				
Group presentations				

2. Why did you choose the categories that you did? Where there any activities that you felt were borderline? Why?

3. Now indicate how the "intensity" for "once a twice a week"

	0 No impact on student learning	1	2	3 High impact on student learning
Ask students to complete a worksheet				
Give a lecture to the class about social studies				
Have students participate in debates or panel discussions				
Have students participate in mock trials, role-playing, or dramatization				
Have students write letters to state an opinion or solve a community problem				
Have visitors from your community meet with the class to discuss important events and ideas				
Have students participate in community volunteer projects or services				
Have students access information through the Internet for use in the classroom				
Discuss current events				
Use student government				
Give students social studies homework				
Tests with multiple choice true/false or matching type questions				
Tests with fill-in the blank questions				
Paragraph-length written responses about what students have read				
Extended essays/papers on assigned topics				
Individual projects				
Individual presentations				
Group projects				
Group presentations				

4. Why did you choose the categories that you did? Where there any activities that you felt were borderline? Why?

5. Now indicate how the "intensity" for "Almost every day"

	0 No impact on student learning	1	2	3 High impact on student learning
Ask students to complete a worksheet				
Give a lecture to the class about social studies				
Have students participate in debates or panel discussions				
Have students participate in mock trials, role-playing, or dramatization				
Have students write letters to state an opinion or solve a community problem				
Have visitors from your community meet with the class to discuss important events and ideas				
Have students participate in community volunteer projects or services				
Have students access information through the Internet for use in the classroom				
Discuss current events				
Use student government				
Give students social studies homework				
Tests with multiple choice true/false or matching type questions				
Tests with fill-in the blank questions				
Paragraph-length written responses about what students have read				
Extended essays/papers on assigned topics				
Individual projects				
Individual presentations				
Group projects				
Group presentations				

5. Why did you choose the categories that you did? Where there any activities that you felt were borderline? Why?

Closing Activity

1. Is there part of this interview you had questions about? Anything that was confusing?

2. Any general questions about my dissertation research?

Appendix B: Interview Response Sheet

I. Introduction

II. Definition of "Interactive Civics"

Please Read Over the Following Definition

Interactive civics describes classroom activities that are:

a. Student-centered, meaning that students, not teachers, are responsible for engaging in their learning on their own and with their peers
b. Requires student to communicate with others (either written or spoken)
c. Engage students in authentic forms of civic participation that expose students to the types of civic activities that they might engage in as adults (e.g., voting, discussing issues, volunteering for organizations, working with others to solve community problems)

Interactive civics do not have to occur within a formal civics class to qualify as an interactive civics activities.

III. NAEP Teacher Classroom Practice Items

Now I'm going to ask you the read over the instructional activities that are listed on the NAEP Teacher Survey. This is not a test. These are just examples of activities that teachers might do. It's okay if you don't do any of these in your own classroom. I just want to get a sense of what teachers might see as an interactive civics activity.

Please check off the activities that could be examples of "interactive civics".

- ❑ Ask students to complete a worksheet
- ❑ Give a lecture to the class about social studies
- ❑ Have students participate in debates or panel discussions
- ❑ Have students participate in mock trials, role-playing, or dramatization
- ❑ Have students write letters to state an opinion or solve a community problem
- ❑ Have visitors from your community meet with the class to discuss important events and ideas
- ❑ Have students participate in community volunteer projects or services
- ❑ Have students access information through the Internet for use in the classroom
- ❑ Discuss current events
- ❑ Use student government
- ❑ Give students social studies homework
- ❑ Tests with multiple choice true/false or matching type questions
- ❑ Tests with fill-in the blank questions
- ❑ Paragraph-length written responses about what students have read
- ❑ Extended essays/papers on assigned topics
- ❑ Individual projects
- ❑ Individual presentations
- ❑ Group projects
- ❑ Group presentations

IV. Frequency of Using Interactive Civics Activities

V. Intensity of Interactive Civics Activities

In this study, intensity is defined in terms of impact on student learning. Activities with different levels of intensity may need to be used more or less frequently in order to have the same impact on students.

1. The following table has a list of all the activities on the NAEP teacher questionnaires. For each activity you identified as interactive mark how "intense" each of the activities would be for your students if you were to use it "once or twice a month"

	0 No impact on student learning	1	2	3 High impact on student learning
Ask students to complete a worksheet				
Give a lecture to the class about social studies				
Have students participate in debates or panel discussions				
Have students participate in mock trials, role-playing, or dramatization				
Have students write letters to state an opinion or solve a community problem				
Have visitors from your community meet with the class to discuss important events and ideas				
Have students participate in community volunteer projects or services				
Have students access information through the Internet for use in the classroom				
Discuss current events				
Use student government				
Give students social studies homework				
Tests with multiple choice true/false or matching type questions				
Tests with fill-in the blank questions				
Paragraph-length written responses about what students have read				
Extended essays/papers on assigned topics				
Individual projects				
Individual presentations				
Group projects				
Group presentations				

2. Now indicate how the “intensity” for “once a twice a week”

	0 No impact on student learning	1	2	3 High impact on student learning
Ask students to complete a worksheet				
Give a lecture to the class about social studies				
Have students participate in debates or panel discussions				
Have students participate in mock trials, role-playing, or dramatization				
Have students write letters to state an opinion or solve a community problem				
Have visitors from your community meet with the class to discuss important events and ideas				
Have students participate in community volunteer projects or services				
Have students access information through the Internet for use in the classroom				
Discuss current events				
Use student government				
Give students social studies homework				
Tests with multiple choice true/false or matching type questions				
Tests with fill-in the blank questions				
Paragraph-length written responses about what students have read				
Extended essays/papers on assigned topics				
Individual projects				
Individual presentations				
Group projects				
Group presentations				

3. Now indicate how the "intensity" for "Almost every day"

	0 No impact on student learning	1	2	3 High impact on student learning
Ask students to complete a worksheet				
Give a lecture to the class about social studies				
Have students participate in debates or panel discussions				
Have students participate in mock trials, role-playing, or dramatization				
Have students write letters to state an opinion or solve a community problem				
Have visitors from your community meet with the class to discuss important events and ideas				
Have students participate in community volunteer projects or services				
Have students access information through the Internet for use in the classroom				
Discuss current events				
Use student government				
Give students social studies homework				
Tests with multiple choice true/false or matching type questions				
Tests with fill-in the blank questions				
Paragraph-length written responses about what students have read				
Extended essays/papers on assigned topics				
Individual projects				
Individual presentations				
Group projects				
Group presentations				

VI. Closing Activity

Thank you so much!!!

Appendix C: Cognitive Interview Protocol on the NAEP Teaching Practices in Social Studies Measures: Validation Sample

Background

The purpose of my dissertation is to study the relationships between teacher professional learning, interactive civics activities, and students' civic skills and knowledge. Specifically, I am interested in the following questions

a. Is teacher professional learning related to the use of interactive civic activities?
b. Do students' who participate in interactive civics activities in social studies class have higher levels of civic skills and knowledge?

I'm using the NAEP civics assessment to answer these questions. NAEP civics is a national assessment of students' civic skills and knowledge that is administered to a nationally representative sample of 8th graders approximately every 4 years. NAEP also collects survey data from the teachers of the students about their educational background, participation in professional development,

I reached out to you because I want to better understand how social studies teachers might have answered the background survey. This will help me understand how I can appropriately use the data in my analysis.

3. Do you have any questions about the background of this survey?
4. Is there anything you want me to go into more detail about?

Definition of "Interactive Civics"

Please Read Over the Following Definition

Interactive civics describes classroom activities that are:

d. Student-centered, meaning that students, not teachers, are responsible for engaging in their learning on their own and with their peers
e. Requires student to communicate with others (either written or spoken)
f. Engage students in authentic forms of civic participation that expose students to the types of civic activities that they might engage in as adults (e.g., voting, discussing issues, volunteering for organizations, working with others to solve community problems)

Interactive civics <u>do not</u> have to occur within a formal civics class to qualify as an interactive civics activities.

1. Do you have any questions about this definition? Is everything clear and easy to understand?

2. Have you used these types of activities in your classroom? Why or why not?

3. [If yes] How often would you use these types of activities in your classroom?

4. What factors might influence how often you might use these activities?

"Interactive Civics" Teacher Classroom Practice Items

Now I'm going to ask you the read over the instructional activities from NAEP that may be characterized as interactive. This is not a test. These are just examples of activities that teachers might do. It's okay if you don't do any of these in your own classroom. I just want to get a sense of what teachers might see as an interactive civics activity.

- ❑ Have students participate in debates or panel discussions
- ❑ Have students participate in mock trials, role-playing, or dramatization
- ❑ Have students write letters to state an opinion or solve a community problem
- ❑ Have visitors from your community meet with the class to discuss important events and ideas
- ❑ Have students participate in community volunteer projects or services
- ❑ Discuss current events
- ❑ Use student government
- ❑ Group projects
- ❑ Group presentations

1. Now that you've read over the list of activities, what if any of the activities do you think might meet the definition of interactive civics?

2. Why did you choose those activities? What about them made them interactive?

3. Were there any activities not on the list that you feel should be included? Why?

Reviewing Impact Weights

For this study, I have created **weights** based on interviews with other social studies teacher. The **weight** indicates the relative impact on student learning. **The weights are on a scale of 0-3 where 0=No impact and 3=High impact on student learning.** Some activities are weighted more heavily than others. Please review this table. **Circle the weights you agree with and put an X on the weights that you disagree with.**

	Never or hardly ever	Once or twice a month	Once or twice a week	Almost every day
Have students participate in debates or panel discussions	0	3	3	3
Have students participate in mock trials, role-playing, or dramatization	0	3	3	3
Have students write letters to state an opinion or solve a community problem	0	2	2	2
Have visitors from your community meet with the class to discuss important events and ideas	0	3	3	3
Have students participate in community volunteer projects or services	0	3	3	3
Discuss current events	0	1	2	3
Use student government	0	1	1	1
Group projects	0	2	2	2
Group presentations	0	2	2	2

1. Let's go activity by activity. Why did you give the activity the rating that you did?

2. What were some of the factors you considered? Why would some activities have more impact than others?

3. Were there any activities you were on the fence about?

4. How would increasing from once or twice a month to almost every day affect how you would teach the activity? How would it change the experience of students?

Closing Activity

1. Is there part of this interview you had questions about? Anything that was confusing?

2. Any general questions about my dissertation research?

Appendix D: Interview Response Sheet (Validation Sample)

Definition of "Interactive Civics"

Please Read Over the Following Definition

Interactive civics describes classroom activities that are:

g. *Student-centered, meaning that students, not teachers, are responsible for engaging in their learning on their own and with their peers*
h. *Requires student to communicate with others (either written or spoken)*
i. *Engage students in authentic forms of civic participation that expose students to the types of civic activities that they might engage in as adults (e.g., voting, discussing issues, volunteering for organizations, working with others to solve community problems)*

Interactive civics <u>do not</u> have to occur within a formal civics class to qualify as an interactive civics activities.

"Interactive Civics" Teacher Classroom Practice Items

Now I'm going to ask you the read over the instructional activities from NAEP that may be characterized as interactive. This is not a test. These are just examples of activities that teachers might do. It's okay if you don't do any of these in your own classroom. I just want to get a sense of what teachers might see as an interactive civics activity.

- ❑ Have students participate in debates or panel discussions
- ❑ Have students participate in mock trials, role-playing, or dramatization
- ❑ Have students write letters to state an opinion or solve a community problem
- ❑ Have visitors from your community meet with the class to discuss important events and ideas
- ❑ Have students participate in community volunteer projects or services
- ❑ Discuss current events
- ❑ Use student government
- ❑ Group projects
- ❑ Group presentations

Reviewing Impact Weights

For this study, I have created **weights** based on interviews with other social studies teacher. The **weight** indicates the relative impact on student learning. **The weights are on a scale of 0-3 where 0=No impact and 3=High impact on student learning.** Some activities are weighted more heavily than others. Please review this table. **Circle the weights you agree with and put an X on the weights that you disagree with.**

	Never or hardly ever	Once or twice a month	Once or twice a week	Almost every day
Have students participate in debates or panel discussions	0	3	3	3
Have students participate in mock trials, role-playing, or dramatization	0	3	3	3
Have students write letters to state an opinion or solve a community problem	0	2	2	2
Have visitors from your community meet with the class to discuss important events and ideas	0	3	3	3
Have students participate in community volunteer projects or services	0	3	3	3
Discuss current events	0	1	2	3
Use student government	0	1	1	1
Group projects	0	2	2	2
Group presentations	0	2	2	2

Appendix E: Subject Recruitment Letter

Dear [Name],

I am a doctoral student at Boston College studying educational research, measurement, and evaluation. For my doctoral dissertation, I will be examining the relationship between middle school student participation in interactive activities in social studies and achievement on the National Assessment of Education Progress (NAEP) civics assessment.

I am contacting you because I would like to interview middle school social studies teachers about their experiences using interactive activities in the classroom and I thought you would be a good person to interview.

Would you be interested in participating in a **20-minute** interview about your experience as a middle school social studies teacher with interactive activities? If so, please let me know what days or times would be most convenient for you! Please also let me know if you have any additional questions about this study.

This study has been approved by the Boston College Institutional Review Board. If you have any questions about your rights as a person in this research study, you may contact: Director, Office for Research Protections, Boston College at (617) 552-4778, or irb@bc.edu

All the best,

Josh Littenberg-Tobias
Doctoral Student
Educational Research, Measurement, and Evaluation
Boston College

Appendix F: Participant Consent Form

Boston College Lynch School of Education
Informed Consent for Participation as a Research Subject

Risks/Discomforts of Being in the Study:

- The study has the following risk: in the course of the interview, I might pose a question that you find offensive or which, for whatever reason, you would prefer not to discuss. If this occurs, you need not answer the question.
- Other than the matter of provocative questions, there are no reasonable foreseeable (or expected) risks. There may be unknown risks.

Benefits of Being in the Study:

- You will probably not get any direct benefit from participating in this study. You may appreciate having an opportunity to express your opinion and provide me with ideas for what my research should focus upon. I cannot guarantee that you will receive any direct benefit from this study.

Payments:

- There will be no financial payment for participating in this study.

Costs:

- There is no cost to you to participate in this research study.

Confidentiality:

- Interview recordings, transcripts, and response sheets will be stored electronically in a password-protected account on the Boston College server. Only the primary investigator will have access to these files. The principal investigator's advisor may view the data but will not have access to the file. Each subject will be assigned a randomly generated identifier that will be used in all transcripts and coding materials. All personally identifying information will be removed from the transcripts. Additionally, pseudonyms will be used in all published materials in order to protect the confidentiality of the subject.

Voluntary Participation/Withdrawal:

- You need not answer every question that is posed, for whatever reasons.
- You are free to withdraw your participation at any time, for any reason.

Contacts and Questions:

- The researcher directing this study is Joshua Littenberg-Tobias, a doctoral candidate at Boston College. You may contact him at: tobiasj@bc.edu or (516) 330-6234
- His research advisor is Dr. Laura O'Dwyer, an Associate Professor in Education Research, Measurement, and Evaluation. You may contact her at: laura.odwyer@bc.edu and (617) 552-8089.
- If you have any questions about your rights as a person in this research study, you may contact: Director, Office for Research Protections, Boston College at (617) 552-4778, or irb@bc.edu

Copy of Consent Form:

- You will be given a copy of this form to keep for your records and future reference.

Statement of Consent:

- I have read the contents of this consent form and have been encouraged to ask questions. I have received answers to my questions. I give my consent to participate in this study. I have received a copy of this form.

Signatures/Dates

_____ Check here if you will allow me to record your interview.

_____ Check here if you do not want to have your interview recorded.

_____ _____

Date *Consent Signature of Participant*

_____

Print Name of Participant

Appendix G: Missing Data Patterns

Table G1

Correlation of Missing Pattern of Student-Level Variables With Student-Level Predictors

	Missing Variables						
	Parent Education	Newspaper	Magazine	Computer	Encyclopedia	Number of books	Days absent
White Non-Hispanic only (School records)	-0.142**	0.001	0.003	0.003	-0.012	-0.004	-0.006
Black Non-Hispanic only (School records)	-0.014	-0.007	-0.005	0.011	0.007	-0.012	-0.003
Hispanic, any race (School records)	0.163**	0.012	0.005	0.000	0.018	0.014	0.020
Asian/Pacific Islander Non-Hispanic only (School records)	0.035**	-0.010	-0.001	-0.016	-0.013	0.008	-0.014
Native American (School records)	-0.005	-0.004	-0.005	-0.009	-0.005	-0.005	-0.005
Other (School records)	0.001	-0.005	-0.006	-0.011	-0.006	-0.006	-0.007
Eligible for Free or Reduced Priced Lunch (School records)	0.178**	-0.004	0.000	0.019	-0.001	0.007	-0.014
Student has Individualized Education Plan (IEP)	0.107**	0.011	0.006	0.030**	0.002	0.005	-0.024*
Student is an English Language Learner (ELL)	0.166**	0.010	0.024*	0.030**	0.011	.023*	0.015
Female	-0.054**	-0.007	-0.018	-0.013	-0.018	-0.011	-0.018
Parent Education	.c	-0.003	-0.004	-0.003	0.002	-0.009	0.016
Newspaper in home	-0.079**	.c	-0.014	-0.004	-0.002	-0.015	-0.016
Magazine in home	-0.126**	-0.010	.c	0.003	-0.01	-0.006	-0.011
Computer in home	-0.080**	0.005	0.008	.c	-.023*	-.039**	0.002
Encyclopedia in home	-0.120**	0.012	0.011	0.001	.c	-0.019	-0.006
Numbers of books in home	-0.182**	0.010	0.002	-0.025*	0.003	.c	-0.006
Days absent from school	0.021*	-0.003	-0.003	0.007	-0.009	0.01	.c
Plausible NAEP civics value #1	-0.226**	-0.054**	-0.057**	-0.089**	-0.061**	-0.064**	-0.059**
Plausible NAEP civics value #2	-0.233**	-0.055**	-0.053**	-0.089**	-0.065**	-0.067**	-0.063**
Plausible NAEP civics value #3	-0.235**	-0.064**	-0.064**	-0.095**	-0.071**	-0.072**	-0.068**
Plausible NAEP civics value #4	-0.233**	-0.074**	-0.069**	-0.094**	-0.080**	-0.081**	-0.070**
Plausible NAEP civics value #5	-0.229**	-0.059**	-0.061**	-0.092**	-0.064**	-0.073**	-0.061**

*p<.05 **p<.01 [c] Correlation cannot be computed because at least one of the variables is a constant

Table G2

Correlation of Missing Pattern of Professional Development Variables with Teacher Predictors (Part 1)

	Missing Variables					
	College course taken after first certification	Workshop or training	Conference or professional association meeting	Observational visit to another school	Mentoring/Coaching/Peer observation	Committee or Task Force
Years' experience teaching 6-12 social studies subjects	0.057	0.042	0.016	0.061*	0.032	0.002
Regular/Standard Certification	0.047	0.043	0.043	0.037	0.038	0.033
Certified by the National Board for Professional Teaching Standards	-0.013	0.013	0.025	-0.006	-0.013	0.010
Number of students per teacher	-0.021	-0.002	-0.025	0.005	-0.011	-0.013
Only teaches social studies	0.028	0	0.006	0.009	0.027	0.005
Teacher is Hispanic, any race	-0.013	0.025	0.044	-0.004	0.018	0.044
Teacher is White, Non-Hispanic	-0.008	-0.053	-0.043	-0.017	-0.047	-0.058
Teacher is Black, Non-Hispanic	0.043	0.054	0.033	0.050	0.062*	0.050
Teacher is Native American, Non-Hispanic	-0.016	-0.011	-0.015	-0.017	-0.014	-0.013
Teacher is Other/Multiracial, Non-Hispanic	0.023	0.046	0.026	0.019	0.030	0.034
Undergraduate major in social science	0.035	0.024	0.031	0.014	0.016	-0.002
Graduate major in social science	0.005	0.058	0.056	0.055	0.029	0.007
Graduate major in non social science	-0.018	-0.018	-0.025	-0.035	-0.041	-0.030
Students participate in debates or panel discussion	-0.001	-0.022	0.016	0.010	-0.002	0.027
Students participate in civic simulations	-0.041	-0.029	-0.02	-0.008	-0.026	-0.027
Students write a letter to state an opinion or solve a community problem	0.003	0.007	0.006	0.037	0.018	-0.034
Students participated in community project	-0.031	-0.053	-0.036	-0.018	-0.042	-0.050
Students discuss current events	0.022	0.007	0.004	0.038	0.004	-0.025
Students do student government	0.030	0.003	0.036	0.008	0.007	-0.015
Student work on group projects	0.021	-0.008	0.010	0.009	-0.030	0.021

Cont.	College course taken after first certification	Workshop or training	Conference or professional association meeting	Observational visit to another school	Mentoring/Coaching/Peer observation	Committee or Task Force
Students make group presentations	-0.028	-0.015	-0.029	-0.002	-0.024	-0.013
Students interact with visitors from the community	0.067*	0.057	0.015	0.041	0.041	-0.003
College course taken after first certification	.c	0.019	0.092**	0.122**	0.144**	0.126**
Workshop or training	0.090**	.c	0.081**	0.081**	0.083**	0.059
Conference or professional association meeting	0.109**	-0.003	.c	0.097**	0.072*	0.054
Observational visit to another school	0.083**	0.028	0.049	.c	0.077*	0.040
Mentoring/Coaching/Peer observation	0.113**	0.080**	0.106**	0.119**	.c	0.073*
Committee or Task Force	0.112**	0.030	0.086**	0.126**	0.087**	.c
Regularly scheduled discussion or study group	0.124**	0.037	0.116**	0.130**	0.086**	0.061*
Teacher collaborative or network (including online networks)	0.041	-0.031	0.098**	0.053	0.090**	0.034
Individual or collaborative research	0.065*	0.009	0.101**	.089**	0.075*	0.049
Independent reading on a regular basis	0.043	0.017	0.049	0.077*	0.047	0.034
Co-teaching/team-teaching	0.077*	0.047	0.094**	0.147**	0.088**	0.066*
Consultation with subject specialist	0.067*	0.049	.079**	0.122**	0.114**	0.059
School leadership position in social studies	0.027	0.027	-0.011	0.028	-0.027	0.009

*p<.05 **p<.01 [c] Correlation cannot be computed because at least one of the variables is a constant

Table G3

Correlation of Missing Pattern of Professional Development Variables with Teacher Predictors (Part 2)

	Missing Variables						
	Regularly scheduled discussion or study group	Teacher collaborative or network (including online networks)	Individual or collaborative research	Independent reading on a regular basis	Co-teaching or team-teaching	Consultation with subject specialist	School leadership position in social studies
Years' experience teaching 6-12 social studies subjects	0.016	0.058	0.043	0.024	0.036	0.036	0.00
Regular/Standard Certification	0.038	0.037	0.024	0.016	0.007	0.044	0.02
Certified by the National Board for Professional Teaching Standards	-0.015	-0.022	-0.02	0.025	-0.02	-0.008	-0.022
Number of students per teacher	0.023	-0.003	-0.007	-0.006	-0.03	-0.027	0.001
Only teaches social studies	0.018	0.012	0.038	0.008	0.026	0.009	-0.048
Teacher is Hispanic, any race	-0.002	-0.006	-0.006	0.02	0.028	0.025	-0.021
Teacher is White, Non-Hispanic	-0.006	-0.025	-0.022	-0.051	-0.022	-0.04	0.003
Teacher is Black, Non-Hispanic	0.022	0.064*	0.053	0.046	0.017	0.049	0.028
Teacher is Native American, Non-Hispanic	-0.014	-0.017	-0.015	-0.014	-0.015	-0.015	-0.005
Teacher is Other/Multiracial, Non-Hispanic	0.029	0.018	0.027	0.031	0.027	0.025	-0.007
Undergraduate major in social science	0.016	0.011	0.017	-0.009	-0.005	0.007	-0.014
Graduate major in social science	0.04	0.028	0.006	0.05	-0.008	0.028	.067*
Graduate major in non social science	-0.022	-0.019	-0.007	-0.022	0.004	-0.025	-0.039
Students participate in debates or panel discussion	0.026	-0.012	0	0.023	0.005	-0.014	0.007
Students participate in civic simulations	-0.006	-0.018	-0.007	-0.003	-0.015	-0.034	0.025
Students write a letter to state an opinion or solve a community problem	-0.003	0.013	-0.001	-0.01	-0.02	0.005	0.056
Students participated in community project	-0.056	-0.03	-0.048	-0.053	-0.035	-0.028	0.056
Students discuss current events	0.025	0.02	0.004	0.019	0.03	0.01	-0.023

Cont.	Regularly scheduled discussion or study group	Teacher collaborative or network (including online networks)	Individual or collaborative research	Independent reading on a regular basis	Co-teaching or team-teaching	Consultation with subject specialist	School leadership position in social studies
Students do student government	0.004	0.012	0.017	0.011	-0.005	0.018	-0.029
Student work on group projects	-0.002	0.003	0.026	-0.014	0.009	0.011	-0.018
Students make group presentations	-0.021	-0.015	-0.014	-0.027	-0.023	-0.01	-0.008
Students interact with visitors from the community	0.071*	.067*	0.062*	0.094**	0.065*	0.057	-0.018
College course taken after first certification	0.082**	0.093**	0.079**	0.070*	0.092**	0.105**	0.022
Workshop or training	0.032	0.085**	0.053	0.026	0.056	0.053	0.025
Conference or professional association meeting	0.054	0.070*	0.059	0.034	0.065*	0.053	0.041
Observational visit to another school	0.003	0.049	0.056	0.079**	0.122**	0.077*	0.028
Mentoring/Coaching /Peer observation	0.073*	0.104**	0.091**	0.085**	0.110**	0.101**	0.001
Committee or Task Force	0.081**	0.129**	0.087**	0.075*	0.106**	0.106**	0.043
Regularly scheduled discussion or study group	.c	0.134**	.105**	0.080**	0.097**	0.107**	0.001
Teacher collaborative or network (including online networks)	0.069*	.c	0.024	0.041	0.069*	-0.026	0.024
Individual or collaborative research	0.053	0.124**	.c	0.044	0.103**	0.075*	0.005
Independent reading on a regular basis	0.037	0.079**	0.058	.c	0.067*	0.065*	0.017
Co-teaching/team-teaching	0.093**	0.138**	0.107**	.118**	.c	0.113**	0.050
Consultation with subject specialist	0.087**	0.115**	0.073*	0.047	.092**	.c	0.012
School leadership position in social studies	-0.008	0.015	0.003	0.016	0.014	-0.004	.c

*p<.05 **p<.01

Table G4

Correlation of Missing Pattern of Interactive Civics Use Variables with Teacher Predictors

	Missing Variables								
	Students participate in debates or panel discussion	Students participate in civic simulations	Students write a letter to state an opinion or solve a community problem	Students participated in community project	Students discuss current events	Students do student government	Student work on group projects	Students make group presentations	Students interact with visitors from the community
Years' experience teaching 6-12 social studies subjects	-0.073*	-0.039	-0.061*	-0.053	-0.047	-0.038	-0.028	-0.02	-0.051
Regular/Standard Certification	0.002	0.002	0.002	-0.004	-0.012	-0.045	-0.054	-0.029	-0.001
Certified by the National Board for Professional Teaching Standards	-0.035	-0.035	-0.035	-0.031	-0.027	-0.014	-0.035	-0.033	-0.033
Number of students per teacher	-0.029	-0.026	-0.045	-0.055	-0.05	-0.024	-0.031	-0.032	-0.049
Only teaches social studies	0.010	0.01	-0.021	-0.008	0.018	0.003	0.016	0.01	0.002
Teacher is Hispanic, any race	-0.003	-0.033	-0.033	-0.03	-0.027	-0.01	-0.005	-0.003	-0.031
Teacher is White, Non-Hispanic	-0.091**	-0.070*	-0.070*	-0.064*	-.0082**	-0.088**	-0.063*	-0.070*	-0.056
Teacher is Black, Non-Hispanic	0.157**	0.157**	0.157**	0.143**	0.167**	0.162**	0.118**	0.126**	0.134**
Teacher is Native American, Non-Hispanic	-0.008	-0.008	-0.008	-0.008	-0.007	-0.009	-0.009	-0.008	-0.008
Teacher is Other/Multiracial, Non-Hispanic	-0.011	-0.011	-0.011	-0.01	-0.009	-0.013	-0.012	-0.011	-0.011
Undergraduate major in social science	-0.070*	-0.032	-0.089**	-0.073*	-0.054	-0.026	-0.042	-0.032	-0.061*
Graduate major in social science	-0.044	0.006	-0.044	-0.039	-0.034	-0.05	-0.046	-0.044	-0.015
Graduate major in non social science	0.027	0.007	0.047	0.024	0.021	0.04	0	0.007	0.015
Students participate in debates or panel discussion	.c	0.037	0.007	-0.010	0.027	-0.044	0.024	0.015	0.027
Students participate in civic simulations	0.028	.c	0.052	0.014	.c	-0.034	-0.032	-0.043	0.02
Students write a letter to state an opinion or solve a community problem	-0.020	0.006	.c	.c	.c	-0.028	-0.019	-0.008	-0.026
Students participated in community project	-0.024	0.018	0.03	.c	.c	-0.032	0.033	0.043	-0.012
Students discuss current events	-0.018	-0.001	-0.039	-0.001	.c	-0.030	-0.035	-0.039	0.021
Students do student government	-0.013	-0.013	0.044	0.036	.c	.c	-0.019	-0.013	0.044
Student work on group projects	0.013	0.036	0.012	0.008	.c	-0.021	.c	.c	0.010
Students make group presentations	0.027	0.024	0.024	0.017	.c	-0.005	0.012	.c	0.049
Students interact with visitors from the community	-0.021	-0.018	-0.015	.c	.c	-0.028	0.017	0.024	.c
College course taken after first certification	-0.034	-0.057	-0.029	-0.046	-0.042	-0.006	-0.016	-0.011	-0.050
Workshop or training	-0.087**	-0.063*	-0.077*	-0.073*	-0.099**	-0.091**	-0.068*	-0.055	-0.063*
Conference or professional association meeting	-0.056	-0.035	-0.069*	-0.079**	-0.072*	-0.041	-0.044	-0.056	-0.085**
Observational visit to another school	0.024	0.005	0.005	-0.011	-0.004	-0.015	-0.006	-0.001	-0.016

Cont.	Students participate in debates or panel discussion	Students participate in civic simulations	Students write a letter to state an opinion or solve a community problem	Students participated in community project	Students discuss current events	Students do student government	Student work on group projects	Students make group presentations	Students interact with visitors from the community
Mentoring/Coaching /Peer observation	-0.048	-0.028	-0.048	-0.055	-0.066*	-0.015	-0.055	-0.048	-0.062*
Committee or Task Force	-0.020	-0.051	-0.031	-0.035	-0.068*	-0.054	-0.078*	-0.072*	-0.022
Regularly scheduled discussion or study group	-0.035	-0.027	-0.048	-0.055	-0.066*	-0.049	-0.093**	-0.088**	-0.062*
Teacher collaborative or network (including online networks)	0.016	-0.008	-0.002	-0.016	-0.009	-0.002	-0.035	-0.031	-0.021
Individual or collaborative research	-0.041	-0.061*	-0.020	-0.025	-0.033	-0.006	-0.047	-.061*	-0.033
Independent reading on a regular basis	-0.038	-0.022	-0.046	-0.037	-0.027	-0.025	-.083**	-.094**	-0.054
Co-teaching/team-teaching	-0.017	0.004	-0.03	-0.039	-0.031	-0.055	-0.044	-0.037	-0.046
Consultation with subject specialist	-0.006	-0.006	-0.006	-0.013	-0.023	0.012	-0.033	-0.027	-0.020
School leadership position in social studies	-0.029	-0.049	-0.029	-0.017	-0.027	-0.008	-0.034	-0.029	-0.023

*p<.05 **p<.01 [c] Correlation cannot be computed because at least one of the variables is a constant

Table G5

Correlation of Missing Pattern of Teacher Demographics with Teacher Predictors

	Missing Variables						
	Teacher race	Years teaching civics	Regular/Standard Certification	National Board Certification	Undergraduate major	Graduate major	Only teaches social studies
Years' experience teaching 6-12 social studies subjects	-0.02	.c	0.053	-0.037	-0.027	-0.04	-0.091**
Regular/Standard Certification	0.016	0.014	.c	-0.019	-0.021	-0.031	0.005
Certified by the National Board for Professional Teaching Standards	0.009	-0.011	0.068*	.c	-0.016	0.025	-0.012
Number of students per teacher	-0.029	0.074*	0.001	0.011	0.010	-0.011	-0.046
Only teaches social studies	-0.014	-0.023	-0.042	-0.022	-0.03	0.031	.c
Teacher is Hispanic, any race	-0.039	-0.016	0.009	-0.024	-0.036	0.025	0.039
Teacher is White, Non-Hispanic	-0.233**	-0.056	-0.035	-0.029	-0.017	-0.061*	-0.070*
Teacher is Black, Non-Hispanic	-0.037	0.110**	0.064*	0.025	0.024	0.028	0.067*
Teacher is Native American, Non-Hispanic	-0.010	-0.004	-0.010	-0.011	-0.009	-0.005	-0.012
Teacher is Other/Multiracial, Non-Hispanic	-0.013	-0.006	-0.014	-0.016	-0.012	0.118**	-0.016

Cont.	Teacher race	Years teaching civics	Regular/Standard Certification	National Board Certification	Undergraduate major	Graduate major	Only teaches social studies
Undergraduate major in social science	-0.035	-0.010	0.024	-0.032	.c	-0.045	-.092**
Graduate major in social science	0.001	-0.019	0.020	0.071*	0.104**	.c	-0.048
Graduate major in non social science	0.028	-0.032	0.043	0.000	0.005	.c	0.054
Students participate in debates or panel discussion	0.037	-0.034	0.045	-0.010	0.003	0.066*	-0.037
Students participate in civic simulations	0.036	-0.013	0.059*	-0.028	-0.014	-0.030	-0.022
Students write a letter to state an opinion or solve a community problem	-0.001	-0.006	0.031	0.024	0.016	0.006	-0.054
Students participated in community project	-0.001	0.077*	0.080**	-0.026	-0.025	-0.021	-0.033
Students discuss current events	0.026	-0.001	0.035	0.032	0.020	0.066*	-.077*
Students do student government	0.002	0.006	0.043	-0.023	-0.038	0.019	-0.050
Student work on group projects	-0.011	0.006	0.059*	-0.007	0.021	0.010	-0.035
Students make group presentations	-0.022	0.012	0.038	-0.023	0.003	-0.008	0.002
Students interact with visitors from the community	0.022	0.049	0.088**	0.030	0.037	0.086**	-0.044
College course taken after first certification	-0.028	-0.027	-0.024	-0.060	-0.026	-0.019	-0.060
Workshop or training	-0.029	-0.024	-0.017	-0.066*	0.006	-0.052	-0.141**
Conference or professional association meeting	-0.029	-0.045	0.040	-0.034	-0.005	-0.032	-.152**
Observational visit to another school	0.003	-0.023	0.007	0.000	0.027	0.056	0.016
Mentoring/Coaching/Peer observation	-0.004	-0.042	0.011	-0.062*	-0.006	0.031	-.075*
Committee or Task Force	0.006	-0.043	-0.032	-0.027	0.017	0.030	-0.106**
Regularly scheduled discussion or study group	-0.003	-0.041	0.019	-0.021	0.056	0.031	-0.100**
Teacher collaborative or network (including online networks)	-0.007	-0.026	0.020	-0.026	0.032	-0.018	-0.041
Individual or collaborative research	0.037	-0.038	0.029	-0.010	0.012	.c	-0.089**
Independent reading on a regular basis	-0.041	-0.078*	0.028	-0.049	-0.031	.c	-0.049
Co-teaching/team-teaching	-0.003	-0.037	-0.012	0.036	0.016	.c	-0.031
Consultation with subject specialist	-0.034	-0.033	-0.013	-0.014	0.022	.c	-0.068*
School leadership position in social studies	-0.028	0.014	0.033	0.006	0.009	-0.001	-0.062*

*$p<.05$ **$p<.01$ [c] Correlation cannot be computed because at least one of the variables is a constant

Table G6

Correlation of Missing Pattern of School Variables with School Level Predictors

	Missing Variables	
	Student Absences	% Free Reduced Priced Lunch
City	0.064	-0.048
Suburb	0.000	0.025
Rural	-0.064	0.023
Northeast	0.020	0.135**
Southeast	-0.045	-0.04
Central	-0.005	-0.033
West	0.029	-0.048
Charter	0.037	-0.008
Private	-0.044	0.156**
Public	0.031	-0.147**
Student Absences	.c	.c
% Free Reduced Priced Lunch	0.092	.c
% Minority	0.065	-0.054
Number Teachers in Sample	0.009	-0.066

*p<.05 **p<.01 [c] Correlation cannot be computed because at least one of the variables is a constant

Appendix H: Sensitivity Analysis

The sensitivity of the models to the alignment between instruction and assessment content was tested by removing students who were unlikely to be exposed to the content. The findings suggest that the models were somewhat sensitive to instructional content. Restricting the analysis to only teachers who spent more than 10% of their time teaching civics reduced the relationships between teachers' professional learning and classroom practices and students' achievement on NAEP civics. Removing schools that did not teach 8th grade civics also weakened most relationships, although less severely.

Removing teachers who did taught civics 10% or less of the time. Examining only teachers who spent more than 10% of their instructional time on civics ($j = 760$) diminished the relationships between the predictors and outcomes of interest. Although total participation in professional learning was still significantly associated with interactive civics use ($\hat{\beta} = .41\ t = 7.85\ p < .001$) excluding teachers who did not spend a significant amount of time on civic reduced the magnitude of the relationship by about 12%. Similar reductions were found in the relationships between traditional ($\hat{\beta} = .71\ t = 6.02\ p < .001$) and communities of practice professional learning ($\hat{\beta} = .57\ t = 7.52\ \ p < .001$). However, the differences in the fixed effects between the full sample and coefficients from the restricted sample were not statistically significant from the zero (Table H1)

Removing teachers who spent little time on civics resulted in an even larger reduction for the relationship between teacher professional learning and classroom practices and student achievement on NAEP civics. The relationship between total participation in professional learning and student achievement was reduced by 36% ($\hat{\beta} = .30\ t = 1.90\ p = .059$).

Traditional $(\hat{\beta} = .55\ t = 1.54\ \ p = .124)$ and communities of practice professional learning $(\hat{\beta} = .40\ t = 1.67\ \ p = .448)$ also had weaker relationships with student achievement on NAEP when the sample was restricted. In addition, the relationship between interactive civics and student achievement declined by 62% $(\hat{\beta} = .08\ t = .76\ p < .448)$. When the sample was restricted none of the estimated fixed effects was statically significant. However, this may be because of the reduction in statistical power due to the decreased sample size. Additionally, the differences in the coefficients between the full sample and restricted sample were not themselves statistically significant (Table H1).

Table H1

Sensitivity Analysis Results: Teacher Who Taught Civics >10% of the Time

Outcome	Predictor	Full Sample $\widehat{\boldsymbol{\beta}}$	Restricted Sample $\widehat{\boldsymbol{\beta}}^{*}$	Difference $(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}^{*})$	Standard errors of Difference $(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}^{*})$	95% Confidence Interval	
						LO	HI
Interactive Civics Use	Total Professional Learning	0.47	0.41	0.05	0.07	-0.08	0.19
	Traditional Professional Learning	0.84	0.71	0.13	0.15	-0.17	0.43
	Communities of Practice Professional Learning	0.65	0.57	0.08	0.10	-0.12	0.27
Student NAEP Civics Achievement	Total Professional Learning	0.47	0.30	0.17	0.22	-0.26	0.59
	Traditional Professional Learning	0.90	0.55	0.35	0.48	-0.58	1.28
	Communities of Practice Professional Learning	0.62	0.40	0.22	0.33	-0.43	0.86
	Interactive Civics Use	0.21	0.08	0.13	0.14	-0.14	0.40

Note. The significance of the difference between the full and restricted sample was calculated by constructing a confidence interval around the difference between the estimated fixed effects. The confidence interval was $(\hat{\beta} - \hat{\beta}^{*}) + 1.96(SE_{\hat{\beta}-\hat{\beta}^{*}})$ where $SE_{\hat{\beta}-\hat{\beta}^{*}} = \sqrt{SE_{\hat{\beta}}{}^{2} + SE_{\hat{\beta}^{*}}{}^{2}}$

* Confidence interval does not contain zero

Removing school where civics was not taught in 8th grade. Restricting the sample to only schools where the principals reported that students learned civics in 8th grade $(k = 160)$ also reduced the estimated fixed effects; although the effects not as dramatic as they were in the previous analysis. Total professional learning $(\hat{\beta} = .43\ t = 6.48\ p < .001)$ and communities of practice professional learning $(\hat{\beta} = .53\ t = 5.11\ p < .001)$ both decreased in their relationships to interactive civics use, while traditional professional development activities increased slightly $(\hat{\beta} = .87\ t = 5.70\ p < .001)$**.** None of the differences between the full sample and the restricted sample were statistically significant.

Similarly, the relationships between teacher characteristics and student achievement decreased in the restricted sample. Compared to removing teachers who spent less than 10% of their time on civics, the reductions in this analysis were not as severe. Total professional learning activities was reduced by 15% $(\hat{\beta} = .40\ t = 1.567\ p = .118)$**,** traditional professional learning by 35% $(\hat{\beta} = .59\ t = 1.18\ \ p = .240)$**,** and communities of practice by 4% $(\hat{\beta} = .59\ t = 1.40\ \ p = .140)$**.** The estimated fixed effect for interactive civic use decreased by 21% $(\hat{\beta} = .17\ t = 1.06\ \ p = .295)$**.** None of the relationships were statistically significant, though that may be because of the large reduction in sample. The differences between the estimated fixed effects in the full sample and in the restricted sample were not statistically significant (Table H2).

Table H2

Sensitivity Analysis Results: Schools Where Students Learn Civics in 8th Grade

Outcome	Predictor	Full Sample $\widehat{\boldsymbol{\beta}}$	Restricted Sample $\widehat{\boldsymbol{\beta}}^{*}$	Difference $(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}^{*})$	Standard errors of Difference $(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}^{*})$	95% Confidence Interval	
						LO	HI
Interactive Civics Use	Total Professional Learning	0.47	0.43	0.03	0.08	-0.12	0.19
	Traditional Professional Learning	0.84	0.87	-0.03	0.18	-0.38	0.32
	Communities of Practice Professional Learning	0.65	0.57	0.08	0.10	-0.12	0.27
Student NAEP Civics Achievement	Total Professional Learning	0.47	0.40	0.07	0.30	-0.51	0.65
	Traditional Professional Learning	0.90	0.59	0.32	0.59	-0.83	1.47
	Communities of Practice Professional Learning	0.62	0.59	0.03	0.45	-0.86	0.92
	Interactive Civics Use	0.21	0.17	0.05	0.18	-0.31	0.40

Note. The significance of the difference between the full and restricted sample was calculated by constructing a confidence interval around the difference between the estimated fixed effects. The confidence interval was $(\hat{\beta} - \hat{\beta}^{*}) \pm 1.96(SE_{\hat{\beta}-\hat{\beta}^{*}})$ where $SE_{\hat{\beta}-\hat{\beta}^{*}} = \sqrt{SE_{\hat{\beta}}{}^{2} + SE_{\hat{\beta}^{*}}{}^{2}}$

* Confidence interval does not contain zero