PROMOTING THE SOCIAL:

CULTIVATING CHARACTER IN URBAN PUBLIC CHARTER ELEMENTARY

SCHOOL CLASSROOMS

by

Mayme Kendrick Hostetter

Dissertation Committee:

Professor Thomas Hatch, Sponsor
Professor Douglas Ready

Approved by the Committee on the Degree of Doctor of Education

Date May 16, 2018

Submitted in partial fulfillment of the
requirements for the Degree of Doctor of Education in
Teachers College, Columbia University

2018

ABSTRACT

PROMOTING THE SOCIAL:

CULTIVATING CHARACTER IN URBAN PUBLIC CHARTER ELEMENTARY

SCHOOL CLASSROOMS

Mayme Kendrick Hostetter

The study examined the reading and character growth of elementary school students (n = 2144) in urban public charter school classrooms (n = 88), focusing on the relationship between the students' growth and a character-focused lesson. Reading growth, as measured by in-classroom reading assessments, and character growth, as measured by self- and teacher- report surveys focused on either grit or self-control, were the outcomes of interest. The study employed a mixed-methods design, combining quantitative methods (i.e., descriptive statistics, correlations, and multi-variable linear regression) and qualitative methods (i.e., video observations, surveys, and interviews) to both describe and better understand the relationship between these outcomes.

The average reading and character growth of the students in the study was notable, with students—on average—making 1.24 years of reading progress (as measured by grade level equivalency) over one school year and demonstrating character growth

beyond expectation. Students in classrooms focused on the character strength of grit grew more with respect to strength of character than did their peers in classrooms focused on self-control. Also, students demonstrating higher levels of grit grew more with respect to reading than did their peers with lower levels of grit.

Counter to the study's hypothesis, the character-focused lessons were negatively (though weakly) associated with students' character growth. Qualitative examination of a subset of the lessons indicated that (a) grit was often positioned as the more "academic" strength, while self-control was often positioned as the more "social" strength and (b) stronger lessons may have heightened students' reference bias, such that students had a more ambitious vision of grit or self-control as a result.

In interviews with a subset of the teachers in the sample, those who led their students to notably above-average character growth all had consistent, robust character education in their schools, in stark contrast to the teachers who led their students to notably below-average character growth.

The study's findings suggest that particular non-cognitive strengths—in this case, grit—are associated with desirable academic outcomes, even in young school children, and that in-classroom and school-wide character education may help to support the growth and development of these strengths.

DEDICATION

for Crista, Vivian, and George

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

ix

LIST OF FIGURES

PREFACE

I started teaching in the fall of 2001. I did not know it at the time, but that was a notable school year for our profession. It was the year of the No Child Left Behind (NCLB) Act (2002), the reauthorization of the Elementary and Secondary Education Act (ESEA). Consequently, it was the year when language like "accountability," "achievement gap," and "highly qualified teacher" began to seep into the discourse of our profession. By the time I entered the public school system as a teacher in 2003, these words were prevalent in the popular press's reports on our profession. Seventeen years later, these words practically define the debates in public education in America.

During this time, I consider myself lucky to have worked at schools and with people who focused on their students' academic achievement *and* character[1] growth. Interestingly, NCLB (2002) emphasized the importance of kindergarten through 12th grade (K-12) students' academic achievement *and* character development (§5431 (a) (1)). Yet K-12 academic achievement was the most prominent focus of the NCLB legislation and subsequent action, and it remains the most notable legacy of NCLB (U.S. Department of Education, 2013, p. 1). Why the character education (CE) element of NCLB was, ironically, left behind in much of the discourse and reaction that followed the passing of the act has always been a point of some confusion and concern for me.

---

[1] "Character" is just one term for the concept in question here. People use different terms for this concept (and related concepts, like "virtue"), depending on the philosophical, research, policy, or practice traditions they represent. For now, I use the term "character" as proxy, and I will define this term more precisely when I introduce the theoretical framework for this study in Chapter I. I will also define and describe several alternatives to "character" as the terminology of choice (e.g., socio-emotional development and non-cognitive skills) in the literature review in Chapter II.

In the spring of 2013, however, I attended a series of conferences in San Francisco—a small gathering of deans of schools of education, a larger meeting for leaders of teacher preparation programs, a large conference for educational entrepreneurs and reformers, and the annual meeting of the American Education Research Association. At each of these gatherings, people were talking about "grit."[2] Of course, they were also talking about students' test scores, accountability, and achievement gaps, but for the first time since I became aware of the larger discourse in and about the field of education, people—many people from many places—were talking about "grit." They were also talking about character strengths, non-cognitive skills, social-emotional learning, and social-psychological interventions. They had many different names for this set of skills (or strategies or mindsets or traits)[3], but it was a decidedly different approach to the conversation about education than the one that has dominated the prior 15 years.

This dissertation was inspired by my experiences during the spring of 2013. It is my attempt to examine how and to what ends contemporary teachers are building both academic skills and strength of character in their students. These two foci of schools, i.e., the academic and the social—so often pitted against one another in the political discourse of schooling—both seem so crucial to children's development. That spring seemed to mark a point at which the discourse was converging. This dissertation seeks to examine how that discourse plays out in teachers' classrooms…and to what ends for their students.

---

[2] Duckworth, Peterson, Matthews, and Kelly (2007) define "grit" as "perseverance and passion for long-term goals" (p. 1087).

[3] In addition to a difference in vocabulary with respect to "character" vs. "non-cognitive" vs. "socio-emotional" (to name just a few prominent examples), there is also a difference in description with respect to whether these concepts represent a set of skills, strategies, mindsets, traits or some combination thereof. I will discuss these differences in more detail in Chapter II.

I – INTRODUCTION

Background of the Study

Since their earliest incarnations, schools have prepared students both academically and socially. The balance, though, between the academic and social foci of schooling has shifted across the eras. Some recent research, however, has begun to show that these two primary foci of schooling—the academic and the social—are not as separate as the somewhat bifurcated schooling of earlier eras might indicate. The character-focused aspects of schooling may promote social growth *and* have strong relationships with students' academic outcomes. Thus, in the early 21st century, one sees promising prospects for more synthesis and less shifting with respect to the academic and social foci of schooling. In the sub-sections that follow, I will briefly describe the shifts from one focus to the other, and finally, the more recent attempts to synthesize the two.

Early Origins

For hundreds of years, the focus of schooling erred on the side of the social, i.e., preparing children to be good people and citizens, with the academic taking a secondary or supporting role. Some of the earliest writings about teaching and learning—and the purposes, practices, and outcomes thereof—come to us from Aristotle. In his *Nichomachean* Ethics, he opens by stating that "the good" is that "at which all things aim" (Aristotle, trans. 2009, I.1.2-3). From Aristotle into the 20th century, learning to be good remains an oft-cited ultimate purpose of schooling. Though the people writing

about leading and teaching in schools have used many different terms—from ethics, to virtues, to morals, to values, to character—the predominant perspective over the last two millennia has been that the development of students' character is a definitive aim of education.

Formal schooling in America also had building students' character as a definitive aim. In the earliest days of American schooling, the focus was predominantly on the moral, with the academic playing a supporting role (McClellan, 1999). From Colonial times through the 1940s, the emphasis on character in American classrooms changed, but did not fade. In the country's earliest schools, American students were encouraged (if not exhorted) to develop strength of character, as well as basic academic skills. Late in the 19th century, Horace Mann's Common Schools, the precursors to today's public schools, sought a citizenry united around shared skills and values (Tyack & Hansot, 1982). In short, from Aristotle through to the Progressive era, many philosophers, theologians, publishers, principals, and teachers have all seen developing children's character as a—if not *the*— primary purpose of schooling.

The Shift to the Academic

Standardized tests started to become popular and proliferate in the U.S. in the early 20th century, such that by the end of World War II, character education (CE) had largely faded from public school curricula (Leming, 1993). The Russians launching Sputnik in 1957 (before Americans had ventured into outer space) marked one way of benchmarking a paradigm shift in the purpose and outcomes of American schooling. No longer were the shared ethics and skills of Mann's Common Schools sufficient; American

students now needed to be prepared to compete intellectually in an international field, and the country needed proof that schools were heeding this call, in the form of standardized test scores, preferably rising ones, for which students, schools, districts, states, and the country would be ever more responsible in the decades to come (Koretz, 2008, pp. 54–55).

Beyond the international competition symbolized by the launch of Sputnik, educational historians have posited a number of theories for the increased focus on academics (and academic accountability, in particular) in American public schools of the latter 20th century. Such theories included an increased awareness of educational inequality within America, improvement in measurements to better understand the effect of academic outcomes on life outcomes, and shifts in the nation's labor market, to name a few (Levine, 2006). Regardless of the causes, the post-Sputnik era has seen the authorization of the National Defense Education Act (1958), the ESEA (1965), the publication of *Equality of Educational Opportunity* (Coleman, 1966), the creation of the Department of Education in 1980, the publication of *A Nation at Risk* (United States, 1983), the reauthorization of ESEA in the form of NCLB (2002) and the Every Student Succeeds Act (ESSA) (2015). Each represents an effort at redirecting and redoubling efforts toward the academic purposes and outcomes of schooling, accompanied most recently by an attendant rise of academic data-tracking and outcomes-based accountability in schools. Thus, in the mid-20th century, students' academic outcomes became the focus of educational research, policy, and K-12 schooling, and have since remained in that privileged position over character development.

A Few Voices for Character

In the midst of the post-Sputnik push for American schools to attend to and raise students' academic outcomes, a few researchers, philosophers, and pundits pushed for the social aspects of schooling to also remain a focus. Though theirs was not the dominant discourse—nor the dominant focus of most American schools—their work largely fueled the CE that occurred in schools in the latter half of the 20th century. Kohlberg's (1969) research on moral development; Noddings's (1988) work on caring; and the more conservative perspectives of Bennett (e.g., Bennett & Delattre, 1978), Ryan (1988), and Lickona (1991), for example, each represent different approaches to the work of educating students toward moral ends.

While the aforementioned historical context likely contributed to the de-privileging of CE in schools, so too did the limited and fractious nature of the advocacy for and understanding of CE. These late-20th century CE advocates represented a relatively small (and a highly splintered[1]) socially-focused counter-current in the tidal shift towards academic outcomes. The CE advocates of this era opposed or de-emphasized a focus on academic outcomes for a variety of reasons, but there were other reasons CE efforts did not accelerate as the nation's focus on accountability did. For example, scholarly and practitioner understanding of non-academic outcomes was, and remains, far more tentative than understandings of academic outcomes. Multiple perspectives on non-academic outcomes existed—from caring (e.g., Noddings, 1988) to

---

[1] Even the term "CE" is a contested and loaded one for many. Some scholars and practitioners whose work might be classified as CE resist the label. For example, Noddings (1988), who identifies as a "care theorist" and promotes "care ethics," seeks to distance herself and her practice from what she sees as the indoctrination inherent in the type of CE promoted by more conservative scholars and policy-makers, such as Bennett and Delattre (1978) and Ryan (1988).

moral education (e.g., Nucci, 2001)—and the idea of uniting around a shared vocabulary, let alone a shared set of character-focused outcomes, would have been anathema to many in the field. The result was a disconnected field of study with competing goals, theories, concepts, and key terms. This lack of coherence provided a less than compelling argument for an investment in these outcomes, especially when placed in the context of the rising focus and agreement on the importance of academic outcomes. Thus, in the second half of the 20th century, the social purposes and outcomes of schooling took on a supporting role to academic purposes and outcomes.

However, some of these voices promoting CE of various forms, particularly the more politically and socially conservative among them, were perhaps part of the reason that NCLB (2002) includes reference to and funding for CE (§5431). In contrast to the earlier shift from a focus on character to a focus on academics, NCLB supports CE in the context of academic endeavors, suggesting that academics and CE can co-exist, if not work synergistically. NCLB states:

> The Secretary is authorized to award grants to eligible entities for the design and implementation of character education programs that —
> > (A) are able to be integrated into classroom instruction and to be consistent with State academic content standards; and
> > (B) are able to be carried out in conjunction with other educational reform efforts. (§5431 (a) (1)).

While the social outcomes of NCLB were largely left behind, the conception of CE in the context of academic endeavors was important for the next phase of both the related research and practice: the synthesis phase which I first encountered on a large scale in the spring of 2013 and described in the preface to this study.

The Synthesis of Academics and Character

The work of Peterson and Seligman (2004) and Duckworth et al. (2007) expanded the discourse and research on CE from the fields of education, philosophy, sociology, and policy to the field of psychology. These psychologists were not, of course, the first to attend to issues that were more typically the realm of education and policy. Williams James (1899/2008) in his *Talks to Teachers on Psychology* wrote of "education and behavior," "the laws of habit," and "the will" over 100 years before Peterson and Seligman (2004) dubbed "persistence" a character strength and Duckworth et al. (2007) coined "grit" and studied it as a related concept. What recent social psychologists have done (that James did not) was to start assessing, measuring, and trying to build these strengths in individuals through empirical studies and interventions with reproducible and generalizable findings.

Much of the recent work in the field has focused on identifying key "non-cognitive"[2] strengths and determining if and how the strengths are associated with positive outcomes like health (e.g., Giltay, Geleijnse, Zitman, Hoekstra, & Schouten, 2004), wealth (e.g., Moffitt et al., 2011), academic performance (e.g., Duckworth & Seligman, 2005), long-lasting relationships (e.g., Eskreis-Winkler, Shulman, Beal, & Duckworth, 2014), and job retention and performance (e.g., Robertson-Kraft & Duckworth, 2014). The studies conducted by these authors are descriptive and correlational. In other words, the researchers were not testing interventions designed to

---

[2] As noted above, "non-cognitive" is another term that researchers and policy-makers have used in the last decade to describe the beyond-academic-content knowledge and skills that students might bring to and/or develop in school. Examples of "non-cognitive" factors include academic behaviors (like going to class), grit and self-control, and academic mindsets (like a sense of belonging at school) (Farrington et al., 2012). I will define and analyze this term in more detail in Chapter II.

build the strengths of interest in these studies (i.e., optimism, self-control, and grit, respectively), but rather they were determining the extent to which these strengths—such that they already existed, as measured in these studies, in the participants—were associated with positive outcomes like health, wealth, and academic achievement.

In addition to the descriptive research, some researchers have also developed and tested a variety of character-related interventions. Some interventions have focused on building character strengths in individuals as an end unto itself (e.g., Emmons & McCullough, 2003), while others have attempted to harness strengths to different ends. Duckworth and colleagues conducted some of this intervention-focused research. For example, Duckworth and Oettingen's work (e.g., Duckworth, Kirby, Gollwitzer, & Oettingen, 2013) suggests that teaching students to use a goal-visualization and planning strategy known as Mental Contrasting with Implementation Intentions (MCII)—a process that is highly related to grit and self-control—increases students' grade point average (GPA), school attendance, and school conduct relative to a control group who does not receive such instruction. Similarly, Cohen, Garcia, Apfel, and Master (2006) demonstrated that affirming students' interests and values through a simple, one-time writing exercise leads to meaningful academic gains relative to control groups. This synergy between the non-cognitive and the cognitive (i.e., character and academics) has been the focus of much of the recent intervention-based research in the field. The fact that several of these interventions were conducted in schools in very brief timeframes (e.g., a one-class-period, goal-setting exercise or a 15-minute writing exercise, as in Cohen et al. [2006]) suggests intriguing possibilities for future school-based interventions.

In their review of such classroom interventions, Seligman, Ernst, Gillham, Reivich, and Linkins (2009) began with a brief description of a dozen or so studies, both correlational and experimental, that associate character strengths with desirable outcomes. They ended this litany of studies with a decisive conclusion and a question: "So we conclude that well-being should be taught in school. But *can* it?" (p. 297). Their question is a great one in its construction. It implies that nothing is taught if nothing is learned. In other words, thousands of schools in this country (and others) teach some form of CE, so quite literally, character can be and is taught in school every day. Seligman et al. (2009) do not question that. What they question is whether or not anything is gained, learned, or strengthened as a result of those CE lessons.

Statement of the Problem

The problem this study seeks to address is related to Seligman et al.'s (2009) question: can character be effectively taught, such that something is measurably, meaningfully learned? We do not know to what extent —and how—character might be built in school. Additionally, we do not know if or how this social learning may be related to academic learning. In the rush toward greater academic achievement, schools have not abandoned the character-focused portion of their work with students. Rather, they have focused more on the academics and less on the character. CE is still there; it is just not typically the primary focus anymore (Seider, 2012). What these efforts look like and what children are learning from them is less clear. Moreover, whether or not (and how) the character-related aspects of schooling are related to character growth and academic outcomes remains an open question worth further examination.

Rationale for the Study

Teachers have been enacting CE (of various names and types) in their classrooms for centuries. Today, there are structured CE curricula that one can easily access and describe. There are published checklists of "principles of effective character education" (e.g., Berkowitz & Bier, 2005) and "what works in character education" (Character Education Partnership, 2010). However, only in the last decade or so has a critical mass of researchers begun examining the effects of such instruction on social growth and the relationship between character and academic outcomes, and that research is suggestive, but not conclusive.

Several studies have suggested a positive relationship between CE and desirable social outcomes, such as decreased disciplinary infractions or increased feelings of connection to school (e.g., Berkowitz & Bier, 2004; Seider, 2012). Other recent studies have posited a relationship between brief, "non-cognitive" interventions and positive social-emotional change, such as feeling more optimistic and grateful (e.g., Froh, Sefick, & Emmons, 2008). The studies showing a relationship between CE and academic outcomes include research that links specific strengths of character such as self-control with positive academic outcomes like GPA (e.g., Duckworth & Seligman, 2005). Some studies have found that school-wide CE correlates positively with school-wide academic achievement (e.g., Benninga, Berkowitz, Kuehn, & Smith, 2003), while other studies have focused on specific non-cognitive strengths that are associated with increased academic achievement (e.g., Cohen et al., 2006). Despite all these findings of positive and significant relationships, a recent U.S. Department of Education (US DOE) (2010)

meta-analysis of seven school-level CE programs' effects on elementary school students found no significant effect of program enrollment on students' social or academic outcomes. In short, while there are many studies suggesting that one can promote social growth and, relatedly, academic outcomes, these findings are neither entirely consistent nor conclusive.

Part of the problem may be that even from the studies finding a positive relationship between character-focused instruction and student outcomes, we do not yet have a good idea of the particular mechanisms—in the school, in the curriculum, and/or in the classroom—that might drive these positive relationships. Recent work has given us *some* insight into potential CE-related levers that propel academic outcomes. For example, Cohen et al.'s (2006) work suggests that affirming children's strengths through writing mitigates stereotype threat and is associated with corresponding increases in their academic achievement. Seider's (2012) work indicates that a context-specific, whole-school approach to CE, in which character-focused instruction is part of the explicit curriculum, is associated with both character development and improved academic outcomes.

Cohen et al.'s (2006) work represents the type of "one-shot" interventions that have proliferated in the social-psychological literature of the last decade, while Seider's (2012) work represents the sort of schoolwide approach to CE that has been a long-standing aspect of education in many schools. If both a "one-shot" and a longer-term, all-school approach to CE seem capable of reaping both social and academic benefits for students, we need to better understand the concrete mechanisms that underlie "what works" in these types of efforts, so that schools and teachers can help students build both

academic skills and strength of character more effectively and potentially synergistically in their classrooms.

This study seeks to combine the promise of CE as seen in schools like the ones Benninga et al. (2003) and Seider (2012) have studied with the promise of "one-shot" character-focused interventions like those of Cohen et al. (2006) and Duckworth et al. (2013). In short, this study focuses on a classroom-based intervention designed to distill the most crucial factors of CE into a one-lesson experience for K-12 students. Participating teachers designed these interventions as informed by reading selections from some of the studies cited above and with the guidance of their professors as part of standard coursework at the graduate school of education (GSE) that served as the study site. I will describe the intervention, the participants, and the site in much further detail in Chapter III.

Statement of Purpose and Research Questions

This study will examine efforts to develop elementary school students' strength of character in a modest number of urban public charter school classrooms. The study will investigate one main research question with several sub-components: In what ways might teachers promote students' character development and academic outcomes through classroom-based, character-focused instruction?

- How do changes in grit or self-control among the students in the study sample compare to those in students in other studies using similar measures?

- How does reading growth among the students in the study sample compare to that of students in other studies using similar measures?

- What, if any, relationship exists between students' academic and character skills and growth?

- What aspects of CE curriculum, planning, pedagogy, and teachers' choices might be associated with students' academic and character growth?

The first and second questions are relatively straightforward measurements. With respect to character growth, the study focuses only on two character strengths, i.e., grit and self-control, in the analysis. This choice is motivated by two factors: (1) among the possible character foci, grit and self-control are associated with the most validated measures and are the most widely cited in the research literature and (2) over half of the teachers in the study chose to focus on either grit or self-control with their students. This study will report beginning-of-year to end-of-year character changes for the sample on those standard measures, as compared with those measured in comparable populations.

With respect to academic growth, measuring reading growth is a standard classroom practice, and this study will report students' beginning-of-year to end-of-year reading gains, as compared with those gains as measured similarly in comparable populations. Consequently, I will first and foremost be interested in describing the social and academic growth of the 2100+ students in this study relative to comparable peers. Secondarily, these data will allow me to investigate whether there are any differences in growth by strength, grade, teacher, and other plausible influencing factors.

The third question is again a purely quantitative one that will allow for the investigation of a relationship between the social and academic measures described above. In other words, is there any correlation between students' character growth and reading growth? Do students' point-in-time character levels have any relationship with

their reading gains (and vice versa)? While some studies have demonstrated a relationship between self-control and grit and improved academic outcomes (like GPA) (e.g., Duckworth & Seligman, 2005), this study will examine a heretofore uninvestigated academic outcome (i.e., reading growth and achievement).

Through the fourth question, I will investigate the various factors that compose a single CE lesson in an attempt to identify and unpack those that seem most potent for social and academic growth. As curricular and pedagogical examples: is it the character strength on which a teacher focuses? is it a clear definition of the character strength, a compelling rationale for the role that this strength plays in one's life, a chance to analyze it in a student-led model, or a chance to flex it in an authentic practice opportunity? The ability to identify some of the most potent curricular and pedagogical levers in character-focused instruction would give teachers much-needed concrete guidance. Additionally, I will examine teachers' choices around these issues. For example, in the CE lessons in this study, teachers were presented with a basic lesson framework, but they had significant choice regarding content and pedagogy within that framework. From the strength of focus, to the definition presented to students, to the opportunities to better understand and/or practice the strength, the teachers made all these choices themselves. This study will add teachers' answers to the questions of how and why they made their instructional choices and what they see as the results.

Theoretical Perspective

This study is rooted in the theoretical traditions of social psychology, and more particularly, positive psychology, the scientific study of positive subjective experiences,

individual strengths, and the institutions that enable both (Peterson & Seligman, 2004).

Within the context of this broader theoretical framework, I have developed a two-part

conceptual framework. The first part is specific to the implementation of CE that relies

heavily on the work of Peterson and Seligman, as well as Seider (2012). The second part

of the conceptual framework is grounded in school-based, non-cognitive intervention

research, such as that of Cohen et al. (2006) and Duckworth et al. (2013). The

framework combines these two approaches, i.e., school-based CE and related social-

psychological interventions, into one intervention designed to reap the promising

outcomes of both methods. I will first outline the broader points and implications of a

theoretical framework grounded in positive psychology before describing the particular

conceptual frameworks associated with CE and brief, social-psychological interventions.

At its simplest, positive psychology seeks to identify "what goes right" with

people and maximize or build upon those conditions (Peterson, 2006). Unlike some

earlier, more trait- or assignation-based views of character, positive psychologists believe

that an individual's character and intelligence are not fixed, but rather malleable, they

aim to use scientific methods to create effective interventions that promote positive life

outcomes (Peterson, Park, & Seligman, 2005). Consequently, one of the goals of positive

psychology is to identify and create institutions that enable good character (Peterson,

2006). To move toward this goal, the field has devoted much time and energy to creating

assessments of positive traits like, for example, the Values in Action (VIA) survey

(Peterson et al., 2005) and the Grit scale (Grit-O) (Duckworth et al., 2007). Positive

psychologists and others then use these assessments of positive traits, along with more

traditional measures like IQ, GPA, earnings, etc., to evaluate the associations between

positive traits and positive outcomes. More recently, researchers have begun to investigate interventions designed to strengthen positive traits and positive outcomes.

Positive traits—one example of "what goes right" with people—have many names and come in many grain sizes. One example of a positive trait, the concept of "character," can be large, unwieldy, and vague. Berkowitz and Bier (2004) defined character as "the complex set of psychological characteristics that enable an individual to act as a moral agent" (p. 73). Berkowitz (2011) later expanded upon this definition to describe a trichotomous conception of character that includes cognitive, affective, and behavioral components.

The plurality inherent in Berkowitz's conception of character as a "set of psychological characteristics" that span people's thoughts, feelings, and actions is echoed in Peterson and Seligman's (2004) conception of character. Through comparative, cross-cultural research, Peterson and Seligman defined character as comprising six virtues, which in turn comprise 24 character strengths. These 24 strengths are the unit of operation for their research, and this cross-context, multi-strength pluralistic vision of character differentiates Peterson and Seligman's conception of character from earlier, more context-driven, all-or-nothing conceptions of character, such as those of early American CE. In Peterson and Seligman's framework, character is the biggest, broadest category; virtues are less broad, but still difficult to observe in action (e.g., justice); and character strengths are the most concrete, observable component of the model (e.g., humor). Peterson and Seligman posit that people express character strengths in different ways depending on context, but the strengths are ubiquitous and can be assessed and promoted via intentional intervention (pp. 13-14). Their overarching definition of

character as being composed of virtues, which in turn are composed of character strengths, is both more expansive than early conceptions of character, but also broader and more concrete than Berkowitz's (and many others in the field, e.g., Lickona, 2004). Consequently, this study will adopt Peterson and Seligman's (2004) simultaneously broad and concrete view of character as comprising a multitude of observable, measurable, and ubiquitous strengths, though as previously mentioned, the study's analysis will focus on two of these strengths in particular: grit and self-control.

According to Peterson (2006), there are many types of institutions—from family to work places—that are potential sites for expressing, enabling, and building character strengths. He identified school as primary among these potential sites. Thus, if we operate from the theoretical perspective that character strengths are observable, measurable, and ubiquitous representations of the broad concept of character, and that school is ripe for their expression, assessment, and promotion, the next set of conceptual concerns deals with envisioning how one might express, assess, and promote character strengths in a school setting, and to what ends.

The over-arching theoretical framework of positive psychology can be applied to two separate fields of school-based endeavor: comprehensive, ongoing CE, and brief social-psychological interventions. The comprehensive, ongoing CE is represented in Seider's (2012) study of three schools with CE as a curricular and pedagogical focus. The social-psychological interventions are represented by studies like Cohen et al. (2006) and Duckworth et al. (2013), where a brief, in-classroom intervention, such as a writing assignment or a goal-visualization exercise, reaped significant academic advantages that lasted at least the duration of an academic year for participating students. The conceptual

framework for this study focuses on the latter, i.e., brief, in-classroom interventions, but also tries to take into account that school context and culture matter, particularly as it pertains to CE and prosocial behavior in diverse school environments (Seider, 2012; Spivak, White, Juvonen & Graham, 2015). In the sections that follow, I will first describe the key concepts underlying both approaches (i.e., brief interventions and school-wide CE), and then describe this study's focus.

Conceptual Framework for Character Education

Peterson and Seligman (2004) defined CE as "a deliberate intervention with the goal of teaching young people a core set of values" (p. 387). The present study will adopt this definition with two slight operational adjustments. The verb "teaching" is too imprecise given the particular pedagogical focus of this study. Therefore, in an attempt to narrow the range of possible variables associated with "teaching" in the broadest sense, this study will focus on a particular though flexible format for teaching one CE lesson: defining a character strength, modeling that strength, and giving students opportunities to better understand and/or practice that strength. Additionally, Peterson and Seligman's description of a "core set of values" is too broad to operationalize in the single-lesson format investigated in this study. As noted above, the present study focuses on two of Peterson and Seligman's (2004) 24 character strengths, i.e., grit and self-control, as identified and defined within the context of the classroom. In short, this study operationalizes the broad concept of CE as a deliberate, classroom-based lesson with the goal of enabling students to define and better understand, use, and/or build one particular character strength.

Working within the broader field of educational psychology, Seider (2012)

describes a long-term, comprehensive, and context-driven approach to school-based CE

that yields measurable effects in terms of both character and academic outcomes. The

schools he studied have context-specific character strengths of focus that are clearly

defined, promoted, and practiced in classroom learning, extracurricular activities, and

school culture more broadly. For example, Boston Prep, one of the schools Seider (2012)

studied in *Character Compass*, has a mission that emphasizes courage, compassion,

integrity, perseverance, and respect. This mission pervades the school:

> Almost every aspect of the Boston Prep school day is centered on those
> five virtues, including a weekly ethics class in which students in grades six
> through twelve learn about how philosophers such as Plato, Aristotle, and
> Rousseau understand these virtues. In both class discussion and written
> reflections, students then draw upon these perspectives to consider their own roles
> and responsibilities as students at Boston Prep and members of their respective
> families and communities. (Seider, 2012, p. 14)

In short, the schools in question identified particular character strengths that were

important to their faculty, parents, and students, and then created myriad opportunities for

students to learn about, use, and build these strengths. So too do the teachers in this

study, albeit within the context and confines of a single studied lesson.

Seider (2012) called this approach "homegrown character education" and

contrasts it with the "copying and pasting" approach many schools take to CE (p. 220).

From Seider's vantage, the power of the CE in the schools he studied came from the

customized nature of their approaches. In Seider's words:

> Stakeholders committed to effective character education in their school
> community must start by determining their overarching objectives for students'
> character development and *then* seek out (or create themselves) the highly
> customized curriculum and practices that will allow students to achieve those
> objectives. (p. 222)

In Seider's study, this "customization" happened at the school level; in this study, it happens at the classroom level.

For an educator in a school with a clear, cohesive, and consistent character-focused mission, CE may be relatively straightforward. As in the schools in Seider's (2012) study, there may be character-focused classes and/curricula that students take as a matter of course. There may be assemblies, clubs, athletic teams, community service trips, etc. that allow students opportunities to build their strengths outside of classroom endeavors. For a teacher in a school that does not have a clear, cohesive, and consistent character-focused mission, however, these resources may not be available. Because the teachers in this study worked in schools with hugely variant school-wide approaches to character (from robust, school-wide efforts to nothing), and because the study only examines one CE lesson, one cannot assume that either of the scenarios described above is the case for this study's teacher participants.

While one lesson is a small "dose" of learning, it is a conceptually analogous microcosm of the larger CE endeavors that Seider (2012) and others describe. (I will describe in the next section how small "doses" of key concepts can reap large gains for students.) Consequently, the conceptual framework depicted in Figure 1 is a classroom-level, one-lesson adaptation of the CE mechanisms that Seider described across a full school year in several schools.

As suggested by Figure 1, the beginning of the CE process includes a catalyzing process in which the teacher in question identifies character strengths that have resonance and value for the learners in question. Next, the teacher chooses one strength from among those identified on which to focus. In the case of this study, the teacher
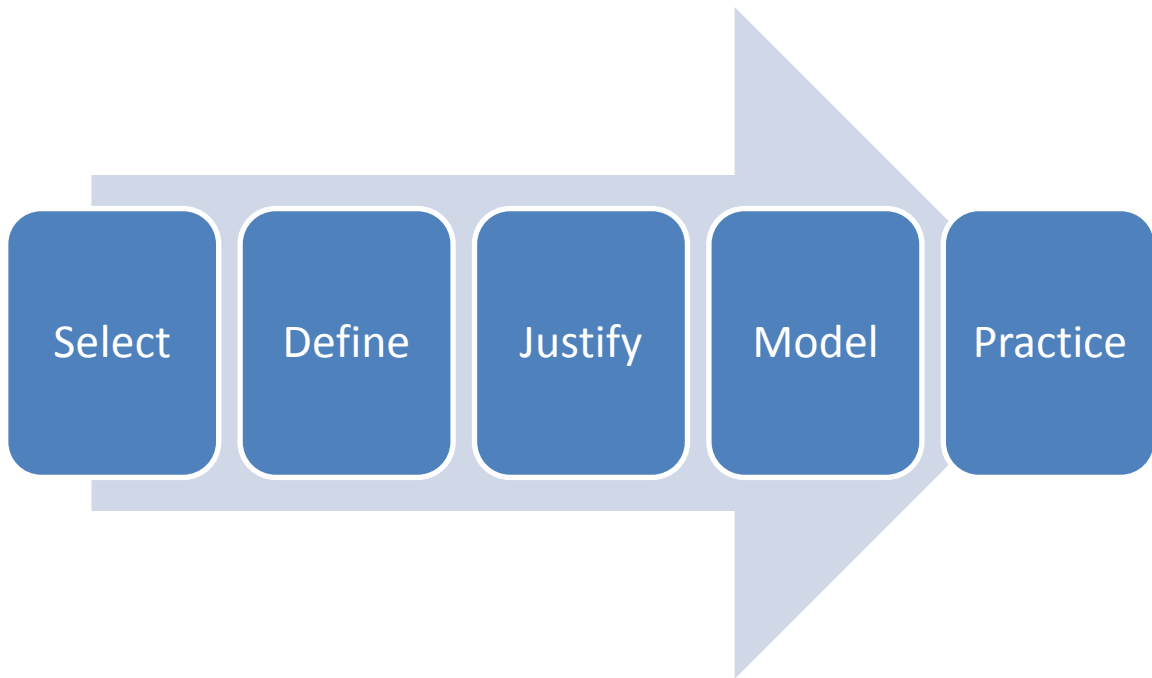
Figure 1. A conceptual framework for a CE lesson in one classroom.

focused on the strength during a single lesson, but, as in Seider's study, the strength could also have been a long-term comprehensive focus in and out of classrooms over one or many school years. As will be detailed in Chapter III, the teachers in this study were encouraged to choose a strength of focus based on students' reported interests, strengths, and values. Defining that character strength for and with the learners—in a concrete, context-specific manner—is the first step of that lesson, a step that Seider (2012) observed across the schools and classrooms he studied. Next, some context-specific rationale for the strength is necessary. Both the definition and the rationale can be buttressed by student or teacher models. Finally, Seider suggested repeating this process of identifying, defining, justifying, and modeling the character strength(s) of choice as part of regular CE opportunities to practice through reading, discussion, and activities, and as an iterative and cyclical approach to building students' strength of character and

academic outcomes. From a single-lesson perspective, after definitions, rationale, and a clear model are presented, students need to have at least one opportunity to practice better understanding and/or building the strength in question.

Some conceptions of CE prescribe very specific and even scripted approaches to helping students to identify and build character strengths. The schools and classrooms in Seider's (2012) study, though, took a wide variety of approaches, and consequently this study's conceptual framework remains open with respect to who (i.e., teachers or students) does the defining, justifying, and modeling, as well as how (i.e., more or less directed, structured experiences) this work happens. Consequently, the teachers in this study were encouraged to think creatively about how best to engage students in the heart of the lesson. In short, the framework suggests that character strengths relevant to teachers and students should be identified, defined, and strengthened in context-specific ways via learning and practice with the twin goals of building students' strength of character and academic outcomes.

Conceptual Framework for Brief, Non-cognitive Interventions

In their review "Social-Psychological Interventions in Education: They're Not Magic," David Yeager and Gregory Walton (2011) defined "small" social-psychological interventions as "typically brief exercises that do not teach academic content but instead target students' thoughts, feelings, and beliefs in and about school" (p. 268). This definition encompasses the dozen or so social-psychological interventions in educational environments that boost student achievement described in Yeager and Walton's review. These interventions were not aimed at helping students acquire new academic content or

understanding, but rather at shifting their beliefs about how, why, or under what conditions people learn and excel in school.

Among many examples of successful interventions, the authors describe Blackwell, Trzniewski, and Dweck's (2007) intervention to teach middle-school students growth mindset (i.e., the belief that one can improve one's intelligence through effort). This simple intervention, which took eight 25-minute advisory periods to implement, with the "treatment" representing just two of those periods, yielded significant academic advantages that lasted through the academic term for the students in the "treatment" group. The crux of the treatment was the two advisory periods that focused on malleable intelligence, introducing students to the idea that "learning changes the brain by forming new connections, and that students are in charge of that process" (p. 254). Students in the control group, on the other hand, had two lessons on memory instead. Something about the "a-ha" moments associated with students' learning that "being smart is a choice you make" (p. 263) had a measurable, positive, and significant impact on students' math grades. Blackwell et al.'s work exemplifies the recent boom in intervention-focused social psychology that does not use the label CE, but has related intents and outcomes.

At the core of the type of work described above—whether that of Blackwell et al. (2007) or the dozen other studies reviewed by Yeager and Walton (2011)—lies a conceptual framework that posits the following: changing someone's thoughts, feelings, and beliefs in and about school through a brief exercise can change his or her academic outcomes for an extended period of time (a semester, a college career, etc.). In other words, one's performance in school is in part predicated upon one's thoughts, feelings, and/or beliefs about school; change the latter, and one can change the former.

Combination Conceptual Framework

Though the school-based CE Seider (2012) and others described differs from the interventions Yeager and Walton (2011) illustrated in terms of designers, duration, and comprehensiveness, the goals, foci, and settings of the two approaches are quite similar. The CE Seider described is often a teacher-designed, years-long experience focused on developing a suite of character strengths in students. The interventions that Yeager and Walton described are researcher-designed, brief (e.g., 15 minutes to a few lessons long), and target one particular school-related feeling or belief (e.g., growth mindset). Both Seider's vision of CE and Yeager and Walton's social-psychological interventions, however, aim to teach school-aged children about non-academic content (e.g., character or growth mindset) to catalyze both social and academic gains among the participants. Both can be set in classrooms and schools (as opposed to laboratories), but differ from typical class time or standard academic instruction. These areas of overlap are notable and suggest that one might be able to reap some of the social and academic benefits for students by combining some aspects of both approaches.

Consequently, the conceptual framework for this study is most closely related to that of the brief, social-psychological interventions, but rests on some foundations of CE (as defined by Peterson and Seligman [2004] and as described by Seider [2012]). In this combination framework, the content of CE is put into the "dosage" size of an intervention (i.e., one lesson), designed and facilitated by teachers but informed by research, with the goal of catalyzing students' academic and character growth. While the study's

intervention shares more with Blackwell et al.'s (2007) approach than with Seider's (2012), the framework seeks to draw upon key elements of both.

As mentioned in the previous section, to distill a comprehensive, school-wide approach to CE into a one-shot lesson is a bridge too far (or too short, as the case may be here). One certainly cannot re-create in one lesson what strong schools do ubiquitously with respect to CE through many facets of their programming over many years. Moreover, one might argue that character is more like reading or math than like growth mindset (or stereotype threat or social belonging, or any of the other beliefs targeted in the interventions in Yeager and Walton's [2011] review). In other words, one might argue that a teacher and her class can no more "do justice" to grit, for example, in one class period than they can do justice to reading in one class period. This may be true. One may need the years-long, comprehensive approach that Seider (2012) described to help students learn about and develop meaningful, measurable strength of character. In which case, the lesson of interest to this study may simply represent the first steps in what could or should be a much longer and more robust endeavor.

That said, the initial reaction to the social-psychological interventions that Yeager and Walton (2011) described was similarly skeptical. The sub-title of their review— "They're Not Magic"—speaks directly to people's disbelief in the power of such brief interventions to have such profound and lasting effects on participants' mindsets and academic achievement. Similarly, writing several years earlier about some of the same interventions, Wilson (2006) began, "Some readers will undoubtedly be surprised, or even incredulous, that a 15-min intervention can reduce the racial achievement gap by 40%. Yet this is precisely what Cohen et al. [2006] report" (p. 1251). The results of

these studies *are* incredible, but they are *not* lacking credibility. A decade later, we now have dozens of examples of brief, social-psychological interventions that had a positive and significant impact on students' academic and social outcomes. In the words of Dweck, Walton, and Cohen (2011), "With greater awareness of non-cognitive factors, educators may be able to do relatively small things in classrooms that can make a big difference in their students' learning" (p. 3).

Though the lesson-level intervention investigated in this study could not be a substitute for a comprehensive, school-wide approach to CE, it may shift students' beliefs about how, why, or under what conditions people learn and excel in school and/or life. The lessons are intended to create "a-ha" moments not dissimilar to those described by Blackwell et al. (2007). For example, if a student has never before considered his daily, 90-minute, 5:30 am subway commute to school as a demonstration of "grit"—a desirable quality related to academic success—a lesson like the ones designed by the teachers in this study may very much change his beliefs about himself and what heretofore unappreciated assets he brings to academic endeavors. As learning that "being smart is a choice one makes" may changes one's academic outcomes for the better, so too may learning that one has notable and transferable grit to bring to bear on academic endeavors.

Much of the literature cited in the previous sections and the literature presented in Chapter II demonstrates that many character strengths are—like the thoughts, feelings, and beliefs targeted by social-psychological interventions—non-cognitive factors related to academic outcomes. Thus, the conceptual framework of this study posits that one can

improve students' understanding and skills related to a particular strength, which in turn can simultaneously propel students' academic achievement.

Researcher Positionality

I am a once and future middle and high school English teacher, and for more than half of my tenure as a teacher, I taught in an urban public charter school. After my first six years of teaching, I paused to study cognitive psychology. Currently, I am the Dean and a teacher educator at the Master of Arts in Teaching (MAT) program that served as the site for this study. Moreover, I was the lead curriculum designer and instructor for the character-related coursework that represents some of the teacher-level "inputs" in this study. All of these roles—teacher, psychology student, curriculum designer, and teacher educator—undoubtedly affect my perspective on the topic and study at hand. I will describe a few of the more salient effects of my positionality, though I am certain that there are others less visible to me, but perhaps no less potent than those I describe here.

As the Dean of and a teacher educator in the MAT program that served as the primary study site, I had direct knowledge of (and in some cases supervision over) many of the teachers who participated in the study. All of these teachers, however, have subsequently graduated from the MAT program and are thus no longer under my direct or indirect supervision. Moreover, I began my data organization and analysis after these teachers had graduated, so that my research did not pose a conflict of interest with respect to my supervisory responsibilities in the MAT program, but at the expense of securing the most timely data possible (i.e., those associated with my current cohorts). Thus, I

have sought to mitigate the role that my professional position might play in conducting this research.

As a teacher, I believed strongly that a teacher's job was manifold, and that primary among those many responsibilities was helping students to build academic skills and strength of character that would help them flourish in school, a career, and life. In short, I came to the work of CE as a believer rather than a skeptic. Likely that belief burgeoned during my "apprenticeship of observation" in my favorite teachers' and mentors' classrooms (Lortie, 1975). Several of my favorite teachers in middle and high school were also athletic coaches, and they brought an attention to grit, optimism, and social intelligence to the classroom in a way that other teachers did not always do. My sixth grade teacher—who inspired and fueled a love of math and reading in so many of us and who created a culture on our tackle football team where I (a girl!) could play with grit and purpose *and* with the respect and support of her male teammates—stands out as a particularly strong example of the many teachers I had who brought both academics and character to life in their practices.

My belief in the connection between the academic and the social was then certainly buttressed by the cognitive psychology research that I read and assisted with as a graduate student. This work—and notably that of Dweck (2000), Peterson and Seligman (2004), and Duckworth et al. (2007)—further convinced me that my "teacher sense" had empirical grounding: non-cognitive factors were related to students' academic and life outcomes. Currently, as a teacher educator, I try to create myriad opportunities for teachers to read, analyze, and find ways to enact the findings of empirical research that support "best practices" in CE. I hope that what I am doing is ultimately helping

teachers to find ways to do their jobs with more efficacy and fulfillment, and that as a result, their young students learn more and flourish.

These hopes and experiences lead me to look for "what works" in CE more than for what does not. In this way, I find myself representative of the larger research community: I am more interested in and inspired by studies that refute the null hypothesis, as opposed to studies that identify variables or interventions that are unrelated to the desired outcomes. I have attempted to guard against this personal (and professional publication) bias in both the literature review and results sections of this study. In the literature review, I have intentionally sought out findings that challenge the hypotheses I have formed as a teacher and a scholar (e.g., the US DOE's [2010] meta-analysis that found no significant effects of CE programs across a range of schools). Thus, while I have grounded my study in Peterson and Seligman's (2004) and Seider's (2012) conceptions of character and CE, I have sought to review a variety of different perspectives and findings.

Regarding the results section of the study, I designed the teacher-level "inputs" (e.g., the graduate level syllabi, curricula, classes, etc.), and I therefore wanted them to "work." I do not think this desire is unique to me. I suspect it is a rare researcher who develops and tests an "intervention" that he or she does *not* want to work. Nonetheless, it seems important to acknowledge that I was aware of this desire and took pains—from independent review of my data, analyses, and findings to a lengthy list of limitations—to guard against its potential to erode the trustworthiness of the study.

Significance of the Study

This study adds to the extant literature in four ways. First, the study describes student-level empirical evidence of character growth. Might some students exhibit different character growth over one year with respect to others? Might one see more growth associated with one strength than another? Previous studies have focused primarily on school-level character growth (e.g., Benninga et al., 2003), making it difficult to determine if there may be grade- or class-level variations in character development. Such a classroom-level understanding of character growth is important, particularly for teachers who work in schools that do not have the resources and/or the school-wide investment needed to launch the kind of successful, whole-school approach to CE seen in studies like Seider's (2012) and Berkowitz and Bier's (2004). In short, the study seeks to understand whether a classroom-level approach to CE might be worth teachers' and students' time with respect to students' social development.

Secondly, this study investigates the relationship between character growth (or lack thereof) and attendant (or lack thereof) academic achievement, as measured in terms of reading growth. While several studies have suggested that particular character strengths, such as self-control, are related to improved academic outcomes, such as GPA (e.g., Duckworth & Seligman, 2005), no research to date has investigated grit or self-control's association (or lack thereof) with reading growth, a key academic outcome in the early years of schooling.

Additionally, this study investigates the lesson-level components of character-focused instruction, in an attempt to see if certain aspects of a CE lesson—from the planning, to the rationale, to the modeling, to the practice—seem to be associated with

greater academic or social growth *or* have more resonance with teachers. Finally, this study explores how teachers conceive of their role in CE. In some approaches to CE, teachers are essentially executing a set of scripted lessons; in other paradigms, teachers play little if any intentional role in character development, aside from correcting students when they misbehave. The CE lessons of interest in this study fell somewhere in the middle of that continuum. They followed a basic framework that requires a definition, rationale, modeling, and practice opportunities within the lesson, but that allowed teachers creative liberty regarding the strength, definition, rationale, and manner of presentation, modeling, and practice. Given these choices, how might teachers' lesson-level decisions and high-level conceptions of CE affect their work with their students during character-focused instruction?

In sum, this study hopes to build upon previous research by investigating one large question—in what ways might a teacher promote students' character growth and academic outcomes through classroom-based CE?—through four, smaller questions related to character growth, character growth and academic achievement, CE curriculum and pedagogy, and teachers' conceptions of all of the above.

II – LITERATURE REVIEW

The literature relevant to this study falls into four basic categories: (1) defining "character" and CE, (2) measuring K-12 reading skills and character strengths, (3) the relationship between CE and students' social and academic growth, and (4) teachers' enactment of CE in the classroom. Broadly speaking, the order represented here is aligned with the study's research questions and design. To undertake this work, I first needed a theoretically sound definition of the key concepts under investigation. Once defined, I proceeded to determine relevant outcome measures, then the relationship between those outcomes, and then models for instruction designed to improve those measures. In short, the literature review presents the literature in the order used to conceive of and enact the study, which in turn is aligned to the order of the research questions.

The first two sub-sections—"Defining Character" and "Defining Character Education"—present the etymology, history, and controversies surrounding the terms "character" and "CE". Because the word "character" encompasses a broad set of meanings and has several limitations, a precise, study-specific definition is presented and several major competing views and limitations are addressed. Similarly, because the term "CE" has different meanings depending on the author or audience, this study needed a clear, concise, and theoretically grounded definition. Following the initial discussion of the language of character, each subsequent section of the literature review is motivated by the related research questions.

Defining Character and Character Education

Defining Character

The language of "character" lacks precision and consensus, and in some cases, engenders enmity. The Oxford English Dictionary (1971) offers 40 distinct definitions of "character." (See Appendix A for more on the language of character.) Given the multitude and precision of character's distinct definitions and sub-definitions, one might think that selecting one would allow for precise, consensus usage across the fields of philosophy, psychology, economics, policy, and education. Alas, not so.

The term is a slippery one, with some scholars defining it one way, while others define it another. Still others say the term does not do justice to the concepts they are attempting to study, so they choose alternatives. "Moral development" (e.g., Kohlberg, 1969), "caring" (e.g., Noddings, 1988), "socio-emotional skills" (e.g., Durlak, Weissberg, Dymnicki, Taylor, and Schellinger, 2011), "noncognitive factors" (e.g., Farrington et al., 2013), and "mindsets" (e.g., Dweck et al., 2011) are just a few of the alternatives to "character" that one finds in the related literature of philosophy, psychology, economics, policy, and education.

Others do not *just* choose alternatives, they *insist* on them. Noddings (2002), for example, objects to associating what she calls "care ethics" with "character." For her, "character" (and particularly CE, as defined by the more conservative voices of her era) has an air of indoctrination from which she seeks to distance herself and the "caring" she is attempting to describe and promote in children, teachers, and schools. In short, research—within and across fields—is far from consensus on the term "character" (or alternatives).

Thus, how is one to choose? One short, wisdom-of-practice answer is this: choose "character" because children, teachers, and parents generally know what one means when one says "character." Another answer is this: the contemporary term of choice for psychologists, economists, and policy-makers is "non-cognitive." The field of education has also begun to adopt this term more recently (e.g., Martin, 2016). Therefore, if three (or four) out of our five fields of interest can agree on the term "non-cognitive," why not use phrases like "non-cognitive skills" or "non-cognitive factors" to describe the concepts of interest in this study? The answer to this question is: (a) not even those fields have reached consensus (e.g., the psychologists Duckworth and Yeager [2015] use "personal qualities") and (b) even those who choose "non-cognitive" as their descriptor acknowledge its lack of precision and problematic nature:

- "We eschew the term 'non-cognitive'" (Borghans, Duckworth, Heckman, & Weel, 2008, p. 5)

- "We find 'non-cognitive' to be an unfortunate word" (Farrington et al., 2012, p. 2)

- "*Non-cognitive* is, of course, a misnomer" (West et al., 2014, p. 1)

Consequently, there is no single term that "best" describes the concept in question from the perspective of everyone writing, researching, and/or practicing with that concept (or closely related ones).

We thus return to the theoretical foundations of the study: the traditions and language of social psychology broadly, and positive psychology more particularly. Many positive psychologists use the term "character" for several reasons. First, the term connects to a long history of respected thinkers. Plato, Aristotle, Horace Mann, and

Martin Luther King, Jr. all wrote and spoke famously and lastingly of the importance of character (Duckworth & Yeager, 2015). Second, as mentioned above, the word has face validity. Children, parents, and teachers generally know what one means when one says "character", whereas terms like "non-cognitive" or "socio-emotional learning" can sound like jargon with a lay audience. Third, Peterson and Seligman (2004) wanted to "reclaim" the term from the hold of "armchair philosophy and political rhetoric" (p. 3). These reasons resonate both philosophically and practically with the perspective and purposes of this study.

  "Character," though, has its limitations related to and extending beyond those mentioned briefly above. First, the term has earned long-standing conservative, moralistic, and religious connotations. "Character" and religion have been closely connected concepts through much of this country's history (McClellan, 1999). For example, William Bennett emphasized the conservative, moralistic, and religious aspects of character when he was Ronald Reagan's Secretary of Education. In a speech during his tenure as Secretary, Bennett (1986) cited several stories from the Bible that he thought should be included as part of CE for all children. He noted, "they teach moral values we all share…they shouldn't be thrown out just because they are in the Bible" (p. 12). Whether one is politically and/or religiously aligned with Bennett or not, the empirical study of the promotion of a politically- and religiously-bound concept in public school classrooms seems problematic. Relatedly and also quite recently, "character" can take on restrictive moralistic connotations. For example, in writing about "tough issues" associated with CE, Lickona (2004) noted that parents should maintain a loving relationship with homosexual children, but that they can do this "without approving of

homosexual activity" (p. 105). One need only read a poll or two associated with LGBT rights to understand that Lickona's concept of character (at least regarding this topic) is both politically and religiously charged. In sum, from the very earliest to the most recent times in this country, "character" and conservatism, moralism, and religiosity have been intertwined in ways that might limit one's ability to study them—let alone promote them—in a public school setting.

This specifically conservative, moralistic, religious vision of character is what Nel Noddings (2002) was rebelling against when she called some conceptions of CE indoctrination and in stark contrast to her care ethics approach. An indoctrinating version of CE stands apart from what Peterson and Seligman (2004) are espousing, as well; rather, they focus on character strengths that they claim are ubiquitous, regardless of political, religious, or cohort affiliation (p. 87). Thus, while some conceptions and enactments of character do attempt to indoctrinate, constrain, or otherwise prescribe how people should act to "be good," Peterson and Seligman's conception of character attempts to honor universally agreed-upon visions of what it means to "be good."[1]

Even Peterson and Seligman's (2004) dogged attempts to describe a vision of character that pervades the boundaries of cultures, though, would likely be problematized by scholars who study the role of culture, particularly as it relates to people's visions of prosocial behavior across race, ethnicity, and other overt markers of culture (e.g., Spivak,

---

[1] While a universally agreed-upon vision of what is good sounds appealing, it is hard, if not impossible, to define. For example, in 1909, James White attempted to identify and illustrate (through biographical examples) the "ideals of right living" (p. i). In his list of 32 "traits of character," we find many that are also in Peterson and Seligman's (2004) list of 24 (e.g., honesty, perseverance, self-control, and courage). Other traits on White's 1909 list seem particular to time and/or place (e.g., obedience and patriotism). Yet others just seem out of place (e.g., kindness to animals). Undoubtedly some will disagree that all of the strengths on Peterson and Seligman's list are universal; likely more people will find such exceptions as time and context change from those of Peterson and Seligman's authoring of the original list.

White, Juvonen & Graham, 2015). Even scholars who do not, epistemologically speaking, consider themselves to be studying culture have problematized some of the implications of different approaches to demonstrating certain strengths, such as when Kidd, Palmeri & Aslin, (2013) put a contextual twist on Mischel, Shoda, and Rodriguez's (1989) self-control findings. I will take up this Kidd et al's (2013) findings and the broader question of the contextual and cultural subjectivity of character strengths later in this chapter, and again when I discuss the study's limitations in Chapter III.

An additional problem with using "character" broadly is its holistic nature. One's "character," in common use, denotes one's general way of being, as opposed to one's particular finite qualities, strengths, or actions. In other words, in more general usage, one is of "good" or "bad" character. Peterson and Seligman (2004) addressed this problem by decomposing character into 24 character strengths (e.g., curiosity, bravery, kindness, fairness, humility, and gratitude). It is these strengths (all positive) that are the unit of interest for their work (and a subset of them that are the units of interest for this study).

One final limitation of "character" lies in its humanistic (at best) or treacly (at worst) connotations that set it apart from empirical epistemologies. Again, Peterson and Seligman (2004) attempted to reclaim the concept for their field by writing, "What distinguishes positive psychology from the humanistic psychology of the 1960's and 1970's and from the positive thinking movement is its reliance on empirical research to understand people and the lives they lead" (p. 4). Thus, while some lay connotations of "character" would discourage investigation through empirical methods, Peterson and Seligman have sought to refine and reclaim the concept for scientific study.

Consequently, "character" broadly and "character strengths" particularly will be the language and concepts of interest in this study. Peterson and Seligman (2004) have both a general definition of character strengths and a set of particular criteria that define their 24 character strengths. The general definition, given above, is that character strengths are the multitude of observable, measurable, and ubiquitous strengths that compose character. The more particular criteria for a strength follow:

- Ubiquitous; widely recognized across cultures;

- Contributes to fulfillments that constitute good lives for oneself and others;

- Has intrinsic value, regardless of associated desirable outcomes;

- Display by one person does not diminish other people;

- Its "opposite" cannot be construed as good;

- Manifests in thoughts, feelings, and actions such that it can be assessed;

- Has a degree of generality and stability across contexts and time for an individual;

- Distinct from other character strengths and cannot be decomposed in them;

- Embodied in well-recognized and agreed-upon paragons;

- Has prodigies;

- Is missing altogether in some people; and

- Institutions and rituals exist for cultivation and proliferation. (pp. 17-28)

Their list contains 24 strengths, from "appreciation of beauty and excellence" to "zest" (see Appendix B for the list of the 24 strengths with brief descriptions, as this study presented it to teachers). Eight strengths served as the strengths of focus for the CE-focused GSE coursework taken by the teachers in this study: curiosity, gratitude, grit,

love, optimism, self-control, social intelligence, and zest. Ultimately, as noted earlier, only two strengths, i.e., grit and self-control, served as the focal points of the study's analysis. The research associated with grit and self-control—as well as the rationale for selecting these two strengths and the implications thereof—will be described and discussed in greater detail later in this chapter, as well as in Chapter III.

Defining Character Education

The study's definition of character—as comprising a multitude of observable, measurable, and ubiquitous strengths, including but not limited to curiosity, gratitude, grit, love, optimism, self-control, social intelligence, and zest—is but one of two foundational definitions needed to proceed with a focused investigation. One also needs a clear, concise, and theoretically grounded definition of CE. This is because even people who agree with and use Peterson and Seligman's (2004) definition of character enact CE in very different ways. Character is to some extent taught implicitly and explicitly and in a variety of ways in every school in this country. How one conceives of, defines, describes, and intends that instruction is another matter.

To stay within the theoretical framework of this study (and as discussed above), Peterson and Seligman (2004) defined CE as "a deliberate intervention with the goal of teaching young people a core set of values" (p. 387). As described in Chapter I, this study uses a more specific version of Peterson and Seligman's (2004) definition of CE to investigate CE planning, enactment, and outcomes: a deliberate, lesson-based intervention with the goal of building a particular character strength in K-12 students by defining, modeling, and increasing understanding and/or practicing the use of the

strength. Peterson and Seligman's definition of CE, either in the original form or as operationalized in this study, importantly does not rule out the type of brief, one-shot approach seen in the social-psychological interventions described above. In fact, through the definition's use of the word "intervention," it seems *intended* to include such forms of CE, despite the fact that many in the field of CE might object to such a connection.

As with "character," there is a huge variety of different words and phrases used to describe educational endeavors related to CE. Depending on the profession, philosophy, or era of the author, we might see any of the following used to describe the same classroom activity: "instilling moral habits," "learning caring approaches," or "building character strengths." Each implies a very different conception of both character and how one might cultivate it. (See Appendix C for more on the language of CE). However, moving from the universe of *possible* phrases used to describe the type of instruction of interest to this study, there is a relatively finite set of terms that contemporary scholars and practitioners use. For example, beyond CE, contemporary scholars use Social and Emotional Learning (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011); Social, Emotional, Ethical, and Academic Education (Cohen, 2006); Social and Character Development (U.S. Department of Education, 2010); and Positive Education (Seligman et al., 2009) as just a few of the alternative terms to CE that seek to describe concepts very similar to CE. Though many of these scholars and practitioners selected these terms to clarify the nuances of the field (as opposed to marking their philosophical separation from or disapproval of others' work), the differences in terminology nonetheless abound.

This study's conception of CE attempts to navigate this complex, interrelated terrain. CE as defined and proposed herein is neither intended as an overt, direct, one-

way "indoctrination," nor as the subtle, "persuasive yet stealthy" approach of many recent social-psychological interventions (Yeager & Walton, 2011, p. 267). Instead, this study attempts to describe and enact a "micro" version of CE as Peterson and Seligman (2004) and Seider (2012) described it. This will be done via the "dosage" of the social-psychological interventions entailing an intentional but transparent and tailored approach to helping students identify and build a particular character strength of meaning and import to them, where both teachers and students are possible facilitators of growth by catalyzing or capitalizing on "a-ha" moments. One simple lesson is the "intervention" of choice.

Measuring Character

With theoretically grounded and workable definitions of both "character" and "character education" in place, I turn now to the literature related to the study's research questions, starting with defining measures of "character." How one might—or whether one even should attempt to—"measure" a concept so difficult to enumerate, so arguably context-dependent at best and immeasurable at worst, has been a matter of debate and endeavor for nearly a century. Attempts to measure character in educational settings began in earnest in the first half of the 20th century, with Hartshorne and May's (1928-1930) "Character Education Inquiry" being the most prominent and oft-cited early empirical examination of school-aged students' conduct. Hartshorne and May's work was part of a CE research boom in the 1930s, which was followed by a dramatic decline in CE-focused research in the middle of the 20th century (Leming, 2008). Lawrence Kohlberg's work in the 1960s and 1970s (e.g., Kohlberg, 1969) was arguably the next

paradigm-shifting attempt to codify and measure character (though he used the term "moral development") in a student population, followed by the more recent work of the educators and social psychologists who have laid the theoretical and practical foundations for this study.

In this section, I will first describe the two primary forms of measuring character strengths in current social-psychological research, i.e., surveys and performance tasks, as well as the advantages and disadvantages of each in the broader literature. I will then explain why I chose to utilize the first method, i.e., surveys, as the primary means of measuring the character strengths of interest in this study, motivating that choice through the literature. Finally, I will report the findings of recent attempts to measure character and character growth, so as to offer points of comparison to this study's findings detailed in Chapter IV.

In a recent review, Duckworth and Yeager (2015) divided current measures of "personal qualities" in school-aged children into three categories: (1) self-report questionnaires, (2) teacher-report (or other-report) questionnaires, and (3) performance tasks. All of these approaches have long histories in the field of psychology. For example, diagnosing psychological struggles like depression has long involved self-report surveys. Using this same method to different ends in the sub-field of positive psychology, Peterson et al. (2005) designed the Values in Action (VIA) survey to measure individuals' character strengths, as based on self-reported answers to a suite of character-related questions. As an education-related example, "other-report" surveys were used as one measure of teacher effectiveness in the Gates Foundation's (2012) Measures of Effective Teaching (MET) study. The MET study found that students'

ratings of their teachers' efficacy were predictive of the students' academic achievement results (as measured by state test gains) under the tutelage of those same teachers. On the performance task front, Hartshorne and May (1928-1930) pioneered the use of performance tasks to measure the traits that they were investigating, including self-control. Moreover, Kohlberg's (1969) famous "moral dilemmas" are a form of performance task, as are Mischel's self-control studies (see Mischel et al. [1989] for a review), utilizing the now-famous "marshmallow" performance tasks[2] as the primary sources of self-control measures.

Distilling these methods of measurement into two categories, i.e., questionnaires and performance tasks, I will now turn to examining each method in more detail before outlining recent findings of studies utilizing questionnaires, representing this study's method of choice.

Questionnaires

Duckworth and Yeager (2015) cited self-report and teacher-report questionnaires as the most prevalent forms of measuring personal qualities. They went on to note the many advantages of these approaches: affordable, quick, reliable, and often predictive of other outcomes of interest (p. 239). Duckworth and Yeager, as well as many others,

---

[2] Mischel's opus of "marshmallow" studies, conducted across several decades, follow a similar (though not identical) methodology. In each study, researchers followed some version of this procedure: A child is escorted by a researcher into a room where a single treat—sometimes a marshmallow—is offered to the child with a caveat. The child is told that he can either eat the treat now (or imminently) *or* wait until the researcher returns. The child is told that if he chooses not to eat the treat until the researcher returns, he will be rewarded with a second treat. In short, one treat before the researcher returns or two treats later. Once "the rules" have been explained, the researcher leaves the child seemingly unattended (though he and/or other researchers are watching through a two-way mirror, so as to be able to observe and time the child's attempts to delay gratification). The duration between the researcher's leaving and returning varies from study to study, depending on the age of the participants and the particular research questions, but is typically around fifteen minutes (Mischel et al., 1989).

however, noted the myriad disadvantages to self- and other-reports as means of measuring personal qualities. For example, reference bias—the relationship between how one rates oneself on such questionnaires and the other people, groups, or contexts that serve as reference points for that self-evaluation (Heine, Lehman, Peng, & Greenholtz, 2002)—is one of several threats to the validity of self- and other-report data.

Seider (2012) and West et al. (2014) both used student self-report surveys to measure a variety of character strengths and non-cognitive skills (e.g., integrity, perseverance, daring, and growth mindset) in middle- and high-school students. Both sets of researchers found that students' self-reported strength of character remained stable or fell over the course of a school year (Seider, 2012) or a middle school career (West et al., 2014). While I will further examine their findings later in this chapter, I raise these studies here because both sets of researchers hypothesize that reference bias affected the validity of their findings about children's perception of their strengths of character over the course of an academic year or middle school career.

Similarly, social desirability bias, i.e., the natural inclination to be viewed in a favorable light, can also affect the validity of self-report questionnaires (Peterson et al., 2005). One can easily imagine a child wanting to be perceived as curious, or grateful, or self-controlled (and so on) by himself and his teacher, and answering self-report items with that perception, rather than accurate self-reflection, as the goal. Social desirability bias can affect the validity of other-report questionnaires as well, especially if the "other" has something to gain, such as a promotion or raise, based upon the subject's reported qualities.

Despite these and other disadvantages (e.g., acquiescence bias, faking, and refusal), self- and teacher-report questionnaires continue to be used regularly and reliably by practitioners and researchers alike. Moreover, some are strongly predictive of other more objectively measured outcomes. For example, people's self-reported answers to the items on Duckworth et al.'s (2007) Grit-O survey predicted a range of positive outcomes, including but not limited to college GPA and retention through the first summer at the U.S. Military Academy (i.e., the "Beast Barracks" of West Point). Such findings have fueled much recent discussion of the possibilities for using simple self-report surveys to investigate all sorts of relationships between non-cognitive and other more objectively measured outcomes. Consequently, despite some well-documented disadvantages, this study uses self- and teacher-report surveys as the primary means of measuring students' character strengths for all of the theoretical, methodological, and practical reasons outlined above.

Performance Tasks

Duckworth and Yeager (2015) found much promise in performance tasks as measures of personal qualities. Most notably, they argued that performance tasks remove the subjective, bias-related disadvantages of questionnaires (p. 242). For example, one could argue that it does not matter whether a child knows that it is socially desirable to exercise self-control and not to eat the marshmallow; the question is simply when (or whether) he eats the marshmallow. This example, though, if one plays it out further, begins to illustrate some of the disadvantages of performance tasks. First, performance tasks need to be very consistently implemented to yield reliable and valid measures. If

the researcher administering the marshmallow task were to come back with a second marshmallow in 13 minutes for one child but wait 21 minutes before returning to another child, the results quickly become meaningless. Second, while the measures themselves are not subject to bias (e.g., the child waited two minutes and forty-five seconds before eating the marshmallow), the interpretation thereof *is* subject to bias and misinterpretation. For example, Kidd et al. (2013) found that when researchers intentionally made themselves appear less trustworthy (by breaking a promise to the child), children were far more likely to eat the first treat (and with less of a wait) than if researchers gave the child no reason to believe they might break their promise of a second treat. In this instance, the measure of time elapsed before eating the treat is still unbiased, but its interpretation as proxy for children's levels of self-control is questionable. The measure may, in fact, be a better proxy for the children's perceptions of the trustworthiness of the researcher or for children's capacity for rational, context-based decision-making.[3]

An additional limitation of performance tasks is that they often require significant resources to design and implement with fidelity. The marshmallow task, though relatively simple to describe, still requires trained personnel, access to a private room with a two-way mirror, enough time to administer the task to children one-by-one, and so

---

[3] Kidd et al.'s (2013) study may call into question the interpretation of the data from Mischel et al.'s (1989) earlier work on self-control. In Kidd et al.'s (2013) study, the researcher's broken promise is just one blip on a child's radar, and this blip has significant effects on the child's choice to wait in hopes of a second treat. Imagine a child who has had myriad experiences with adults who break their promises. This child might, by extension, be less likely to trust the researcher—broken promise or not—and eat the first treat quickly. In doing so, the child would be making an arguably rational and strategic choice based on his prior experiences with less-than-trustworthy adults: eat the first treat quickly because who knows if this adult is going to fulfill his promise of a second. In such a situation, Mischel's (1989) interpretation of the data might conclude that the child had little self-control relative to his peers who waited longer; an alternative interpretation, however, in light of Kidd et al.'s (2013) findings, may be that the child is making a wise choice.

on. In short, one could not feasibly administer this task in the course of a typical school day. Even performance tasks that have been specifically designed for in-school administration, such as the essay-writing tasks used in Cohen et al. (2006) or the monetary choice task used by Duckworth and Seligman (2005), require significant preparation and time to administer with fidelity.

Because of the resource-based limitations of performance tasks described above (and the advantages of questionnaires outlined in the previous section), this study used self- and other-report questionnaires to measure students' character growth. The precise methods of this study's approach will be outlined in Chapter III, and the study's findings regarding students' character growth will be described in Chapter IV and further discussed in Chapter VI.

Measuring Reading

While character growth is the outcome of primary interest in this study, students' reading growth serves as an important comparative (and perhaps related) outcome measure. Unlike the domain of non-cognitive measurement, a field dominated by researchers and not yet widely adopted by practitioners like teachers and school leaders, measuring students' reading ability has been standard practice for both researchers and practitioners for many decades now. Though the measurement tool is often different based on the user (e.g., researchers, schools system officials, or teachers) and/or the reading components of interest (e.g., decoding, reading fluency, reading comprehension, and/or some composite measure), a variety of measurement options are available to

researchers and practitioners interested in measuring some aspect of children's reading ability and/or growth over time (Salinger, 2006).

The two main categories of reading assessment are standardized achievement tests (e.g., the National Assessment of Educational Progress and the Stanford Achievement Test) and classroom-based reading assessments (e.g., the Fountas and Pinnell Benchmark Assessments). In the sections that follow, I will first briefly outline the literature describing both types of assessments. In Chapter III, I will provide more detail regarding the methods and the findings of classroom-based reading assessment, as it is this study's preferred method, a choice I will motivate through the literature. Finally, this study's precise methods for measuring students' reading growth will be outlined in Chapter III, and the study's findings regarding students' reading growth will be described in Chapter IV and further discussed in Chapter VI.

Standardized Achievement Tests

With the relatively recent federal emphasis on accountability for K-12 students' achievement (e.g., NCLB, 2002 and ESSA, 2015), both national- and state-level standardized tests of students' academic achievement have proliferated. It is now possible to find nationally normed tests of reading ability for every grade level from pre-kindergarten through 12th grade. Some of these tests are administered to purposive samples of students across this country (and others) on an annual basis to better understand how American students are faring as compared with students from other nations (e.g., the National Assessment of Educational Progress [NAEP]). Other exams are administered at the request of states, districts, or individual schools to gauge how the

students in those locales are faring as compared with same-age or -grade peers (e.g., the

Stanford Achievement Test [SAT] or the Terra Nova—Comprehensive Test of Basic

Skills). Yet other standardized achievement tests are administered at the mandate of the

federal or state governments, as required measures of how students, teachers, schools,

districts, and/or states are performing as compared with their counterparts on the same

exams. The Partnership for Assessment of Readiness for College and Careers (PARCC)

assessment is a multi-state example of such a mandated exam, while the New York State

English Language Arts (ELA) assessment is a state-specific example.

Such normed, standardized, reading achievement tests present several advantages,

including reliability, validity, and in some cases, detailed results reporting, and large

samples of comparative data (for example, see Appendix D for more on students'

expected grade-to-grade gains in reading). That said, they have several limitations for

practitioners who are trying to measure their students' progress during a given school

year and attempting to use those data to make informed instructional decisions. First,

cost is a prohibitive factor. Schools have to pay significant fees to purchase, administer,

and receive score reports on proprietary tests like the Stanford Achievement Test, Ninth

Edition (SAT-9). Second, test duration is a limiting factor. Whether one is administering

the Terra Nova or the New York State ELA exam, students generally need several hours

(if not days) to complete the test, which is an amount of time that teachers cannot spare

for assessment (as opposed to instruction) on a regular basis. Additionally, such tests

typically do not provide teachers with specific information about students' strengths and

weaknesses as readers; they typically offer only a final, holistic score that informs

instruction in only the broadest terms. Finally, intended test timing is a limiting factor.

Many standardized reading achievement tests are designed to be taken at the end of a given school year (as opposed to the beginning or repeatedly over the course of a year). Consequently, teachers often cannot analyze students' achievement until they have left their class for the year. Such analyses may be helpful for the incoming class, assuming some of the trends in the data are associated with the teacher's instruction as opposed to the particular children tested in her class the previous year. However, the children who actually took that year's test are long gone by the time a teacher knows their results (Salinger, 2006). In short, while standardized tests of students' reading achievement offer the benefits of reliability, validity, and often detailed score reporting, they suffer many shortcomings from an instructional and practical classroom perspective.

Classroom-based Measures of Reading

While nationally normed standardized reading achievement tests offer many advantages, the practical limitations noted above are reasons why many teachers— including all the teachers in this study—use classroom-based assessments of reading to monitor children's progress and make appropriate instructional decisions throughout the school year. Moreover, much professional literature extolls the benefits of using classroom-based reading assessments as a means of following children's reading progress through the academic year and using the data generated by the assessments to inform instructional decisions (e.g., Kerbow & Bryk, 2005). However, though classroom-based reading assessments are commonly used by teachers and thoroughly supported in the professional literature, there is a dearth of peer-reviewed literature describing or evaluating these assessments. This is likely the case for many reasons, ranging from the

commercial nature of the assessments to the difficulty and expense of conducting such a study (Stoelinga, personal communication, November 4, 2016). Thus, while all such assessments used by the teachers in this study have been validated by their publishers, none has been validated in a third-party, peer-reviewed study. Consequently, the information that is available about these tests is best left for Chapter III, where I will briefly outline the procedures, advantages, and limitations of classroom-based assessments of reading.

The Relationship Between Character and Academics

In relation to this study's first research question, I have outlined above the major findings regarding measuring character and reading in school-aged children. I turn now to the study's second research question: what, if any, relationship exists between students' academic and character growth? While this study operationalizes academic growth as reading growth and character growth as student- and teacher-reports of said growth (with more on both of these measures in Chapter III), this section of the literature review will look broadly at recent studies examining the relationship between academics and character in school-aged populations. I will describe this work in two sub-sections: the first focuses on *descriptive studies* and the second focuses on *intervention-based studies*. This study's data can (and will) be used to produce both descriptive and intervention-based findings. Thus, an outline of related recent scholarship will serve to ground the study's methods and findings in the more robust work of other researchers.

Descriptive Studies

While myriad studies look at the relationship between certain strengths (e.g., gratitude) and likely related outcomes (e.g., feelings of positivity and optimism about one's life), this study focuses on grit and self-control, with particular reference to (a) their prevalence and change in a school-aged sample over a single school year, as well as (b) their relationship to one academic outcome: students' reading growth within the same academic year. Consequently, in the sub-sections that follow, I will focus primarily upon studies that identify one particular strength or non-cognitive characteristic, and examine its relationship with academic outcomes. In so doing, I will be ignoring wide swathes of related recent work in the fields of, for example, economics, positive psychology, and public health (i.e., work focused on outcomes beyond the academic). I do so not because I think this work is not valuable, but because my outcomes of interest are much more narrowly defined that those of this broader field of research. Thus, it seems most relevant to review the recent research that is closest to both the conceptual and methodological confines of this study.

Self-control, self-discipline, self-regulation. In many studies over several decades, Walter Mischel and colleagues examined children's ability to exercise self-control. Typically through some version of his now famous "marshmallow" studies (described in more detail on p. 42), these researchers tested children's ability to delay gratification, sometimes following up with later studies to test the relationship between childhood self-control and later-in-life outcomes. In an overview of this research, Mischel et al. (1989) outlined the headlines of their previous findings. The following are

a selection of those findings that connect self-control to academic or intellectual measures:

- Children who prefer delayed rewards tend to be more intelligent and have higher achievement strivings.

- Children who exhibit higher levels of self-control at the age of four are described by their parents ten years later as more academically competent than their peers.

- Children who were able to delay gratification longer in pre-school earned, on average, higher Scholastic Aptitude Test (SAT) test scores upon applying to college. (p. 934)

In short, Mischel and colleagues' research demonstrates that not only are there notable differences in self-control among children of similar ages, but that higher levels of self-control as measured in the "marshmallow" studies have positive and significant relationships to later-in-life academic outcomes.

While Mischel's conception of self-control centered on a child's ability to delay gratification as measured by related performance tasks, more recent investigations of the relationship between self-control and academic outcomes often use self- and other-report surveys to measure self-control. For example, Duckworth and Seligman (2005) implemented a battery of widely used self-control-related surveys (and one performance task) to assess middle school students' self-discipline. They found that this construct was positively and significantly associated with students' GPA, achievement test scores, and acceptance to selective high schools.

In related research, Tsukayama, Duckworth, and Kim (2013) worked with middle-school students to describe a set of behaviors indicating lapses in self-control

(e.g., allowing one's mind to wander or interrupting others [p. 889]). The researchers then used the children's descriptions to develop a set of self-control (or rather, lack thereof) indicators, which they then translated into student-, parent-, and teacher-report surveys. Using these surveys to measure students' self-control, Tsukayama et al. (2013) found that self-control decomposed into two domain-specific forms: school-work related and interpersonal-related. School-work impulsivity (the counterpoint to school-work self-control) was significantly and negatively related to students' GPA, a more refined but expected result given Mischel and colleagues' findings some decades earlier.

In conclusion, whether researchers used performance tasks, widely available self-control surveys, or student-created self-control surveys to measure self-control in school-aged samples, the relationship between self-control and desirable academic outcomes—from GPA to performance on standardized achievement tests to high school acceptance—was consistently positive and significant.

Grit, perseverance, industriousness. If self-control is the ability to regulate what one feels and does from moment to moment, "grit," as Duckworth et al. (2007) call it, is perseverance and passion for long-term goals (p. 1087). The two concepts are highly related, but self-control can be measured in minutes (e.g., the time a child can wait before eating a marshmallow), whereas it may take years of sustained grit to accomplish a bigger goal like graduating from college. Like self-control, however, grit can be reliably and validly assessed using self- and other-report surveys (Duckworth et al., 2007; Duckworth & Quinn, 2009). Moreover, like self-control, grit is positively and significantly associated with desirable academic outcomes, such as educational attainment and undergraduate GPA (Duckworth et al., 2007).

Duckworth et al. (2007) developed and validated a 12-item, 5-point Likert scale self-report Grit-O survey, and Duckworth and Quinn (2009) later developed and validated an 8-item version of the same survey (Grit-S). Duckworth and colleagues have now used these surveys in dozens of studies, showing that people's self-reported measures of grit are associated with everything from high school graduation and remaining married (Eskreis-Winkler et al., 2014), to choosing more challenging forms of academic practice and making it to higher rounds of the Scripps National Spelling Bee (Duckworth, Kirby, Tsukayama, Berstein, & Ericsson, 2011).

Finally, grit, too, must certainly be subject to context and interpretation like self-control. As noted earlier in the chapter, the measurement of self-control as in studies like Mischel et al. (1989) is an objective process, but the implications and interpretations of those measurements are subjective, given the context and experience of both the participants and the researchers (Kidd et al., 2013). While no completely analogous study to Kidd et al.'s work has been published regarding the subjective interpretation of grit as measured by, for example, Duckworth and Quinn (2009), several researchers have written about the context- and culture-dependent aspects of what are deemed prosocial, admirable behavior (e.g., Taylor & Graham, 2007). In short, what may be admirable grit in one context, may be dismissible folly in another, and vice versa.

Growth-mindset, incremental theories of intelligence. Carol Dweck's work of the last several decades has largely focused on students' implicit theories of intelligence, as well as the relationship between those theories and students' academic performance. In Dweck's (2000) framework, students have one of two theories of intelligence: (1) a fixed or "entity" theory of intelligence, in which one believes one's intelligence is fixed

(like eye color) or (2) a malleable or "incremental" theory of intelligence, in which one believes one's intelligence is a malleable quality that can be developed (like a muscle). Though not one of Peterson and Seligman's (2004) 24 character strengths, an incremental theory of intelligence is one part of the family of "non-cognitive" factors, including self-control and grit, that researchers and policy-makers (e.g., Farrington et al., 2012) have connected both positively and significantly with desirable academic outcomes. Thus, while an incremental theory of intelligence does not fit formally within the theoretical framework of this study (as it is not one of Peterson and Seligman's 24 strengths, per se), its relationship to more "fitting" concepts like self-control and grit has become a foregone conclusion in many circles. Moreover, the intervention of choice in many of Dweck and colleagues' later studies (described in the next sub-section of this review) is a classroom-based lesson (i.e., the same format as the intervention in this study). Consequently, it seems relevant to this study to examine the relationship between this non-cognitive factor and academic outcomes.

Dweck (2006) has dubbed a malleable or incremental theory of intelligence a "growth mindset" and finds that students with a growth mindset do better in school than their peers with more fixed mindsets. In further research Dweck and colleagues (e.g., Blackwell et al., 2007) have shown that students with stronger growth mindsets fare academically better across difficult academic transitions, showing higher rates of math test score growth, for example, through the transition to middle school. Though originally measured through a brief battery of survey items, Dweck's laboratory can now measure growth mindset with just two items answered on a 6-point Likert scale (from "strongly disagree" to "strongly agree"):

- You can learn new things, but you can't really change your basic intelligence.

- Your intelligence is something about you that you can't change very much

  (Dweck, 2006).

People's answers to these questions characterize their mindset along a spectrum from fixed to growth.

West et al. (2014) measured growth mindset in 1,000+ middle schools students using three of Dweck's (2006) original survey items. Though the researchers did not report simple descriptive statistics associated with the students' scores on those measures, they did report that (a) growth mindset increased on average by 0.49 SDs for fifth through seventh grade and (b) growth mindset measures were positively related to eighth-grade math and ELA tests scores, with correlations of 0.32 and 0.36, respectively (p. 13). In sum, a very short set of self-report survey items about growth mindset allowed researchers to conclude that this non-cognitive factor is (a) likely to increase over time and (b) associated significantly and positively with desirable academic outcomes.

Intervention-based Studies

As mentioned above, this study will investigate the relationship between academics and character, both descriptively and through the lens of a particular classroom-based "intervention." The studies outlined above demonstrate, from a descriptive perspective, that certain strengths of character and non-cognitive skills—more precisely, self-control, grit, and growth mindset—are consistently, positively, and significantly associated with desirable academic outcomes. In the studies I review in the sub-sections that follow, researchers are investigating the next logical set of questions:

How might one boost students' strength of character and non-cognitive skills through targeted interventions? If one can do that, does one see an attendant rise in academic outcomes, as well?

I will describe the intervention-based research in two sub-sections: (1) *social-psychological interventions* and (2) *education-based interventions*. Though I will focus, in both cases, on studies that were set in K-12 schools, the distinction between the two approaches is important. This study's "intervention", i.e., a character-focused lesson, takes place in a classroom in the format of a more standard lesson (as described in the conceptual framework above): it is an education-based intervention. It takes place during a regular school day, as part of a regular class, with students' regular teacher at the helm. It is not a scripted activity or procedure led by a trained researcher that takes place in school for convenience. Again, I will provide more detail regarding the methods and procedures of this study in Chapter III, but understanding where this study fits within the context of the research presented in the sub-sections that follow may prove helpful to the reader.

Social-psychological interventions. Earlier I described in broad strokes some of the methods (e.g., performance tasks) and findings of Walter Mischel's suite of "marshmallow" studies. As noted in those sections above, Mischel and colleagues (e.g., Mischel et al., 1989) found that children's self-control was positively and significantly associated with academic and social outcomes. Mischel and colleagues also conducted studies that focused on teaching children strategies to increase their self-control (e.g., increase the time they waited to eat the treat). Indeed, children who were taught to focus on the "cool" qualities of the treats (e.g., how marshmallows look like white, fluffy

clouds) were able to delay twice as long as children taught to focus on the treat's "hot" qualities (e.g., how sweet marshmallows taste) (Mischel & Baker, 1975, as cited in Duckworth, Gendler, & Gross, 2014, p. 210). Similarly, children taught how to turn treats into pictures (by putting an imaginary frame around them) delayed more than twice as long than children taught to imagine that pictures of the treat were real (Moore, Mischel, & Zeiss, 1976, as cited in Mischel et al., 1989). In other words, several of Mischel's early studies proved that one could teach children strategies that improved their ability to delay gratification, or in other words, increased their self-control. However, could such an increase in strength of character or an additional non-cognitive strategy be associated with academic outcomes? Mischel's work has yet to report the long-term academic differences between children taught such self-control strategies and those who were not. Nevertheless, recent research has begun to examine similar constructs, e.g., a non-academic intervention with positive academic outcomes, with much to report.

As mentioned in Chapter I, David Yeager and Gregory Walton published a literature review in 2011 describing the burgeoning field of small, social-psychological interventions designed to improve students' performance in school. Yeager and Walton (2011) defined "small" social-psychological interventions as "typically brief exercises that do not teach academic content but instead target students' thoughts, feelings, and beliefs in and about school" (p. 268). Wilson (2006) had written a similar (though much briefer) perspective piece in *Science* several years earlier, outlining the findings of a handful of such studies. Just five years later, Yeager and Walton (2011) were able to cite dozens of studies, each targeting some aspect of students' character, non-cognitive skills, or psychological beliefs—from those intended to build growth mindset (e.g., Blackwell et

al., 2007) to those intended to reduce stereotype threat (e.g., Cohen et al., 2006)—and

each leading to GPA gains for students in the treatment groups ranging from .23 to .80

grade points. While I will describe some of the studies in Yeager and Walton's (2011)

review (as well as others) in more detail below, their headline was staggering: small,

inexpensive interventions were helping students to achieve academic gains far greater

than many more expensive and comprehensive interventions yield. In Yeager and

Walton's words, the findings were "not magic" (p. 267); they were many, robust, highly

consistent, and replicable across different studies, school contexts, and student

populations.

One of the focal studies of both Wilson's (2006) piece and Yeager and Walton's

(2011) review was Cohen et al.'s (2006) study designed to mitigate the effects of

stereotype threat in middle-school students. The study used a deceptively simple

intervention. In a fifteen-minute writing task, students identified a value of importance to

them (from a given list of a dozen or so values) and described why that value had

importance for them (while students in the control group simply wrote about a value that

was important to someone else). The effects were astonishing: for seventh-grade

African-American students in the treatment group, the average treatment effect was .30

grade points, nearly a 40% reduction in the achievement gap between those students and

their white peers over the latter two-thirds of a school year (p. 1308). This study, perhaps

more than any other to date, proved that very brief social-psychological interventions

could have profound and measurable effects on students' academic achievement.

The Yeager and Walton (2011) review also highlighted some of their own

research, including Walton and Cohen's studies (e.g., Walton & Cohen, 2011) about the

relationship between a sense of social belonging and academic performance in college students. Similar to Cohen et al.'s (2006) earlier work, the Walton and Cohen (2011) study featured a brief intervention (a one-hour session) comprising a relatively simple set of tasks. These included reading about how many students felt they did not belong in college early in their time there, but "grew confident in their belonging with time" (p. 1448), writing an essay affirming this trajectory for future students, and then turning that essay into an on-camera speech. This intervention reduced the racial achievement gap (as measured by GPA) between black students and their white peers by 52% over the next three years of college. Additionally, the intervention improved African American students' self-reported health and well-being three years after the intervention. This intervention led to an astounding set of academic and health outcomes, especially when one considers its brevity and simplicity.

At the time of Yeager and Walton's (2011) review, Oettingen and Duckworth and colleagues had just begun publishing findings regarding a self-control-, grit- and motivation-related intervention (e.g., Duckworth, Grant, Loew, Oettingen, & Gollwitzer, 2011). Building on Mischel and others' work on self-control and Duckworth and colleagues' work on grit, Oettingen, Duckworth, and colleagues (e.g., Duckworth et al., 2013) designed a goal-visualization and planning activity which was called, in some iterations, Mental Contrasting with Implementation Intentions (MCII), and in others, Wish-Outcome-Obstacles-Plan (WOOP). The activity was intended to help school-aged children set meaningful personal goals, identify both positive associated outcomes and potential obstacles, and thus set themselves up to use self-control and grit to overcome obstacles and achieve their desired outcomes.

In a series of studies, Oettingen, Duckworth and colleagues demonstrated that students who were led through a MCII/WOOP goal-setting session were able to earn higher scores on vocabulary quizzes (Gollwitzer, Oettingen, Kirby, Duckworth, & Mayer, 2011), complete more practice problems (Duckworth et al., 2011), and improve their GPAs, attendance rates, and conduct grades (Duckworth et al., 2013) relative to peers in control groups.

While I have only described a handful of social-psychological interventions designed to improve participants' academic and social outcomes, scores of such empirically validated interventions now exist. As noted in Chapter I, these interventions both inspire and conceptually undergird the intervention of focus in this study: they are brief, they are simple, they are focused on non-academic aspects of students' lives, and they have strong connections to students' academic and social progress. Moreover, in the three studies described above, the format of the intervention is one brief lesson of writing, reading, speaking, and/or goal-setting. None of the lessons, however, are explicitly "character education." While some do focus on certain character strengths (e.g., Cohen et al.'s (2006) self-identification of values) or constructs related to character (e.g., Oettingen and Duckworth and colleagues' goal-visualization protocol), none of this work is categorized as CE.

In the section that follows, I will describe education-based work focused on CE. The study's conceptual goal, again, is to utilize the conceptual structure of interventions like those described above, and the conceptual content of the CE work described below to create a brief CE intervention designed to improve both academic and social outcomes

for participants. The next section, therefore, provides a review of the research related to the CE-focused portion of this combination framework.

Education-based interventions. In the studies that I outline next, the classroom is the setting, and the curricula, teachers, teaching, and/or schools in question are the "interventions." In other words, there is no "experimental moment" as in Cohen et al.'s (2006) fifteen-minute writing task or Oettingen, Duckworth, and colleagues' teaching students to set goals using MCII/WOOP. The "interventions" are cumulative and generally doled out over the course of the school year. Sometimes they are explicit, as with the US DOE's (2010) examination of the effects of commercial CE programs in various schools. Other times they are more implicit, as with Blazar and Kraft's (2015) investigation of the relationship between teachers and teaching *and* students' cognitive and non-cognitive outcomes.

The simpler studies in this group sought to determine the extent to which CE was present in certain schools and to what extent CE presence was related to school-level academic outcomes. Benninga et al. (2003) undertook this work and found that the presence of CE as a school-wide curricular focus was predictive of school-wide academic achievement, with more prevalent CE predicting higher academic outcomes. However, the sample bias inherent in the study's methods, which included only schools that applied for a competitive recognition process run by the state, calls the generalizability of its findings into question.

The US DOE (2010) did similar but more methodologically rigorous work in a meta-analysis of seven school-level Social and Character Development (SACD) programs and their effects on some 6,000 elementary school students in these programs.

Using an experimental design that paired schools implementing a prescribed SACD program (the treatment group) with schools carrying on their home-grown SACD activities (the control group), the Department found very few meaningful positive effects, either academic or social, of SACD program involvement on student outcomes. However, of the 42-school control group, "standard practice," rather than "no treatment" was the norm (p. xxxv), and over half of the teachers and almost all of the principals at those schools reported school- and classroom-level CE efforts. In other words, at many "control" schools, there was homegrown CE "treatment." Consequently, the study may have said less about the reported non-impact of SACD programs than it did about the un-reported lack of differences between the commercial SACD programs under investigation and the homegrown approaches to SACD happening in the "control" schools. Thus, while there were few and inconsistent differences favoring the former, that may simply mean that the latter was working better than one might expect. In fact, some scholars (e.g., Seider, 2012) would make the case that some particularly tailored and comprehensive homegrown approaches to CE might be more effective than packaged programs of the sort investigated by the US DOE (2010).

Seider (2012) engaged in a version of this investigation in a mixed-methods study of three public charter schools in Boston, each of which had building students' academic skills and strength of character as the central foci of their schools' missions. Though Seider did not use a matched comparison group of schools lacking a character-focused component to their programming, he was able to collect and analyze student-level beginning- and end-of-year character data. All students at all three schools took self-report surveys in September of 2010 and again in May of 2011. The surveys measured

students' integrity, perseverance, and daring (using items from both the Academic

Motivation and Integrity Survey [Stephens, Young, & Calabrese, 2007] and the youth

version of the VIA [Peterson et al., 2005]). In almost all cases, students' perceptions of

their own integrity, perseverance, and daring either stayed the same or fell from the

beginning of the year to the end. While Seider's (2012) findings at first seem to confirm

the null hypothesis that the U.S. DOE (2010) study reported, Seider's (2012) student-

level data allows for a more nuanced examination.

Although students' perceptions of their integrity, perseverance, and daring stayed

the same or fell over the course of the school year, the inter-school differences were

notable. Each school had a particular domain of character on which it was focused. The

students in schools that focused on a particular domain of character maintained a higher

self-conception of and commitment to the character strength most closely aligned with

that domain (relative to their peers at other schools). In short, the school-level, character-

domain-specific interventions *were* associated with students' seeing that strength more

strongly in themselves at the end of the school year (relative to students at similarly

structured schools that focused on other domains of character), despite a near universal

decline in students' perceptions of their strengths.

A year later, Seider, Gilbert, Novick, and Gomez (2013) conducted a follow-up

study that connected Seider's (2012) initial findings regarding students' strength of

character to their academic outcomes. Seider and colleagues (2013) found that students'

self-reports of perseverance were significantly and positively associated with students'

academic achievement (as measured by GPA). In contrast, students' self-reported

measures of integrity were significantly and negatively associated with their academic

achievement, though positively and significantly associated with reports of students'

conduct. In short, the performance character associated with one school's character focus

was associated with better grades, while the moral character associated with another

school's character focus was associated with more positive reports of conduct.

The advantages of Seider and colleagues' work (2012, 2013) over that of, for

example, Benninga et al. (2003) and U.S. DOE (2010) are two-fold. First, from a

qualitative perspective, Seider (2012) provided detailed descriptions of how the schools

he studied arrived at their character domains of focus, as well as how they sought to

develop related strengths among their students through curricula, classes, and traditions.

Benninga et al. (2003) did provide a list of criteria for gauging the extent to which CE is

focused in schools. This list, though, does not provide the depth or richness of Seider's

(2012) descriptions, many of which would allow motivated teachers to incorporate these

very specific practices into their own classrooms. Second, from a quantitative

perspective, Seider provided benchmark measures of student-level character growth over

the course of a given school year. These measures allow other researchers (this one

included) to gauge how other character-focused education-based initiatives compare,

from a quantitative perspective, with those of the three schools Seider studied.

Seider's (2012) notable finding of stability or negative change in students'

perceptions of their strengths of character was replicated in the longitudinal data set

reported in West et al. (2014). In this small portion of West et al.'s work, the researchers

tracked one grade cohort of students in two over-subscribed charter schools (n = 104) and

one open-enrollment district school (n = 65) over three years. West et al. found that

enrollment in the over-subscribed charters was associated with a steady decline in

students' perceptions of their own conscientiousness, self-control, and grit (as assessed by self-report survey items), by 0.65 SD, 0.78 SD, and 0.65 SD for each strength, respectively, over the three years of data collection. These self-reported scores also declined among the students attending the open-enrollment district school, but by notably smaller amounts (p. 22). These findings of self-perceived decline in key character strengths are consistent with West et al.'s broader, point-in-time findings, as well as Seider's results (2012) described above.

As noted previously, Seider (2012) and West et al. (2014) attributed the decline in students' perceptions of their own strength of character to reference bias. In Seider (2012), the baseline character measures were taken at the start of the school year when, as Seider noted, students are "at their most idealistic regarding their own levels of self-discipline and perseverance in carrying out their academic responsibilities" (p. 108). As the school year progresses, students encounter real, grit-testing situations like the choice between completing homework or playing video games, a teacher with exceptionally high standards, and/or classmates whose work ethic goes beyond that of students' previous peers. As students' grit is tested (and probably occasionally fails), their self-perception of their own strength in that domain may decline. This might be especially true in schools with large amounts of homework, many teachers with exceptionally high standards, and many students with strong work ethics. Seider (2012) and West et al. (2014) both inferred that these characteristics are more prevalent in the "no excuses" and/or over-subscribed charter schools where these declines are (most) pronounced in their findings. In sum, though students' self-reported levels of certain strengths of character generally stayed the same or declined across time, both Seider (2012) and West

et al. (2014) concluded that this had less to do with an actual decline and more to do with a heightened sense of what true grit, etc., looks like and requires.

This study's focus on teachers' one-lesson grit and self-control interventions in urban charter schools is targeted to address this self-perceived decline. Seider (2012) and West et al. (2014) documented schools that were raising students' academic outcomes but depressing their self-perception of their non-cognitive strengths. If teachers in these schools focused, even for one lesson, on helping students to identify such strengths and see those strengths in themselves, might one see an attendant rise in self-concept, along with academics?

Conclusion. The U.S. DOE's (2010) conclusions that school-wide CE programs had no meaningful effects on students' academic or social outcomes stand in direct contrast to the findings of Benninga et al. (2003) and others (e.g., Durlak et al., 2011). Seider's (2012) concluding remarks suggest that the Department's findings in support of the null hypothesis may have more to do with the type of CE programs the Department examined, as opposed to the possibilities of CE. He wrote, "'copying and pasting' a character-education program into a school's existing culture and practices is not likely to be successful. Context matters" (p. 220). The Department's study was focused on commercially available, pre-planned, scripted CE programs, which are by definition, "copied and pasted" into schools' existing culture. While Seider's (2012) comments alone do not discount the Department's large-scale, rigorous study, they do suggest that CE may have as much to do with the "how" as the "what," especially if one is looking for attendant academic gains.

In the section that follows, I will review recent findings on both the "what" (i.e., curricular characteristics and foci) and the "how" (i.e., pedagogical practices) of contemporary CE. Because of its rich qualitative detail, Seider's (2012) work will continue to be a focal point, as will Benninga et al.'s research (2003), as each of these studies provides at least some broad criteria for what CE curricula and instruction should look like.

The Curriculum and Pedagogy of Character Education

There are many ways to examine CE. Scholars tend to promote and/or investigate CE of particular types, so one can examine what they are promoting and/or investigating. Practitioners tend to design and/or teach CE, so one can examine what they are designing and/or teaching. The literature representing scholars' visions for and investigations of CE is plentiful. The literature representing practitioners' designs for and enactments of CE in the classroom is far less abundant. Consequently, this section will begin with a brief description of CE as promoted and investigated by scholars over the half-century and end with a briefer description of CE as designed and practiced in the classroom more recently.

The desire to promote students' strength of character in educational settings has a long, broad history in this country (see Appendix E for more on the history of pre-21st century CE in America). CE, as operationalized and analyzed in this study, is but one recent, particular enactment of CE derived in part from a diverse range of related curricular and pedagogical efforts over the last several centuries. To understand how this study's conception and operationalization of CE is both situated within and distinct from

various parts of that history, the next section will describe what CE looks like in some contemporary schools through the lens of what CE has looked like over the last century. For any references to pre-21st century CE that are unfamiliar in the sections that follow, the reader should refer to Appendix E, which comprises a more comprehensive history of CE in America.

Contemporary Character Education

In their white paper "What Works in Character Education," Berkowitz and Bier (2005) listed 33 CE programs "for which there is scientific evidence of effectiveness" (p. 5). While the authors do not state their criteria for "scientific evidence" or "effectiveness," they do offer several broad conclusions about CE based on the 33 programs in question. Some of these conclusions are circular (e.g., "[CE] *does* work" [p. 5]), while others are more elucidating (e.g., peer interaction is the norm, direct instruction is common, and skills are a major focus [p. 7]). It would be difficult, however, to get a clear vision of contemporary CE from Berkowitz and Bier's work alone. Examining the plans and materials associated with the 33 programs they are synthesizing into a handful of conclusions might help to get a better sense of what these programs are asking schools and teachers to do. Yet those plans are merely *plans* for CE. We still would not know if or how leaders and teachers were enacting those programs and plans in their schools and classrooms.

The Character Education Project (CEP) (2010) investigated a similar question— what works in CE?—from the opposite direction, starting with "effective schools" (p. 1) and building out principles of effective CE based on those schools' practices.

Unfortunately, like Berkowitz and Bier's (2005) earlier work, the CEP's report "Eleven Principles of Effective Character Education" also fails to note the criteria they used to define and select the "effective schools" that they studied in order to arrive at their CE principles. They also fail to define "effective CE" and how they know that these principles are, in fact, what leads to said efficacy. Their report does, though, offer a more concrete and enaction-based vision of contemporary CE than Berkowitz and Bier's (2005) work. For example, among the CEP's 11 principles, there are descriptions of how "effective schools" select core values, help staff and students to understand and practice those values, and how teachers might integrate CE into their academic classes. The details are helpful whether one is trying to envision or enact CE in contemporary schools and classrooms.

While Berkowitz and Bier (2005) described what schools need to do to enact effective CE, CEP (2010) described what schools are doing in enacting effective CE. Despite the different approaches to a common question, the two groups of researchers reached several common conclusions. Both papers noted the following actions as key to effective CE:

- Involving the whole school community in selecting the foci of CE

- Explicitly articulating the school's CE foci

- Integrating CE across different aspects of school life

- Providing faculty and staff with CE-focused professional development

- Having teachers serve as role models of the school's values and strengths

- Explaining character-related concepts directly and explicitly to students

- Giving students opportunities to interact in CE-focused instruction (e.g.,

  discussions, cooperative learning, class meetings)

- Giving students opportunities in varied settings to practice character-related skills

The first four of these commonalities in the list above operate at the school level and are therefore beyond the purview of this study's classroom-based single-lesson focus. They are, however, prevalent in Seider's (2012) descriptions of CE at the three school sites he studied. The last four examples are all at the classroom level and are similarly representative of the types of teaching, activities, and experiences Seider noted in the individual classrooms he observed. The fact that Seider (whose work was based in schools that demonstrably helped students to build academic skills and strength of character) documented many of these practices in the schools he studied seems to provide some evidence that there is validity to Berkowitz and Bier's (2005) conclusions and CEP's (2010) principles. What practices—and what aspects of the practices—are most important to effective CE, though, are still questions ripe for investigation.

Recent Character Education and This Study

The CE that Seider (2012) observed in his schools and the CE enacted in this study in many ways exemplify several aspects of these more recent conceptions of CE. At Boston Prep, one of Seider's study sites, all students took an ethics class, in which they engaged in moral reasoning around empathy, compassion, and perspective-taking. Moral reasoning took the form of grappling with ethical dilemmas and debating ethical decisions. One can see the connection to both Kohlberg's (1969) approach to cognitive moral development in the pedagogical choices, as well as Noddings's (2002) approach to

caring, in the foci of those lessons (and again, see Appendix E for more on both

Kohlberg's and Noddings's work). There is also certainly an air of the neo-classical

approach to CE in many of the classrooms Seider (2012) studied. For example, several

of the texts and authors featured in Boston Prep's ethics classes—Aristotle's

Nichomachean Ethics, Emerson's thoughts on friendship John McCain's *Character is*

Destiny, etc.—seem sprung from former Secretary of Education Bennett's suggestions

regarding neo-classical CE (e.g., Bennett, 1986). Finally, as already noted above,

Seider's observations align with many of the visions of strong CE as defined by

Berkowitz and Bier (2005) and CEP (2010). With connections to Kohlberg (1969),

caring, neo-classical CE, and contemporary visions of CE, the CE described in Seider's

(2012) study seems less heir to just one focus or approach to CE and more related to all.

Similarly, many of the above (and more) foci and approaches to CE were

instructional options for the teachers in this study. For example, the "direct" approach

advocated in neo-classical visions of CE was, in some of the lessons examined in this

study, the approach of choice. In other words, teachers defined, offered rationale, and

modeled the strength in question for students, and then gave them some "standard" text

through which to analyze said strength (e.g., a book about Wilma Rudolph as an example

of grit). In other lessons, however, teachers used a more "indirect" approach, offering

students the chance to define a strength and identify its importance in their lives, and then

giving them some more authentic way of flexing that strength as a "practice" opportunity

(e.g., identifying something that was a distraction to completing homework and

developing a plan to surmount that distraction). In short, provided that each lesson

included a definition, rationale, model, and practice opportunity related to one of the

strengths of focus, the parameters of the curriculum and pedagogy were left to the teachers themselves.

Teachers' Character Education Conceptions and Practices

The preceding sections of this literature review focused on scholars' endeavors to measure, describe, and promote character in school-based contexts. The last research question, however, asks how teachers themselves conceptualize character, CE, and the relation of these concepts to the academic aspects of schooling. In other words, how do teachers define and describe character and CE? What choices are they making as they enact CE in their classrooms? Why are they making those choices? What outcomes— social and/or academic—are they striving towards? It is one thing for scholars and others to investigate and explicate these questions; that is the literature I have reviewed thus far. It is quite another thing for teachers themselves to answer these questions; that is the literature I will briefly review here.

This section of the review is brief not because the associated questions lack potency or because teachers lack substantive answers, but rather because so little has been written about what teachers think about these topics. Consequently, I will describe the published findings to date. However, sharing contemporary teachers' perspectives on character, CE, and the relation of those concepts to their students' academic growth may be one of the most novel aspects of this study. The extant related literature falls into two categories: survey-based studies of teachers' sense of efficacy in teaching character and illustrative teacher practices as described in CE-related studies (e.g., Seider, 2012;

Robertson-Kraft & Austin, 2014). I will first summarize the findings of the survey-based studies and then the select teacher practices.

Self-efficacy studies. Milson and colleagues (e.g., Milson, 2003; Milson & Mehlig, 2002) investigated teachers' sense of efficacy with respect to teaching character via a survey-based analysis. Their instrument addressed teachers' beliefs about their own abilities to build, influence, and shape students' character. Milson and colleagues found that teachers' sense of efficacy regarding teaching character is generally positive, with elementary school teachers having a higher sense of efficacy than their secondary school colleagues (Milson, 2003). Additionally, Milson (2003) found that teachers who participated in character-focused professional development (PD) felt more effective than their peers. Similarly, Nucci, Drill, Larson, and Browne (2005) found (using the same survey instrument) that teacher candidates who participated in character-focused coursework felt more effective than peers in a control condition.

These headlines—that teachers feel effective with respect to teaching character *and* that one can increase this sense of self-efficacy via PD and coursework—are promising. That said, Milson and colleagues (Milson, 2003; Milson & Mehlig, 2002) found that when the survey items featured students who were not demonstrating strong character (e.g., "Teachers cannot be blamed for students who are dishonest"), on average teachers indicated that they felt less effective. Similarly, when survey items featured students whose home environment may not support the development of certain character strengths (e.g., "I am able to positively influence the character development of a child who has had little direction from parents"), on average, teachers again indicated that they felt less effective (Milson, 2003, p. 97). In short, Milson and colleagues found that

teachers, especially elementary school teachers, felt positively about their ability to teach character, except in circumstances where the student's disposition or home may not support that teaching.

While the headline here, i.e., teachers feel positively about their ability to teach character, is still promising, the sub-findings are disconcerting on two levels. First, teachers seem to feel less effective with the students who may need their support the most. Second, survey prompts like those quoted above contain problematic assumptions about the definition of character and the role of teachers and schools in developing it. The first prompt above (regarding the "dishonest" children) implies that some children are honest, while others are dishonest. This seems a false dichotomy and a problematic one for teachers to hold. If someone either is or is not honest or gritty or grateful, it leaves little room for the type of gradual learning and/or contextual nuances that mark the development and enactment of character strengths as defined and measured in this study and many others. Moreover, the second prompt above, regarding the child with unhelpful parents, presumes that a teacher knows what each child's parents do to support (or not) the child's character development. Teachers often do know a fair amount about a student's home context, but no teacher can know everything, and whether or not (and how) a child's parents are giving "direction" regarding character would be very difficult to ascertain with certainty. Thus, while Milson and colleagues' major findings are encouraging, one wonders how teachers interpreted some of the survey items given their limiting assumptions. Additionally, Milson and colleagues' survey did not ask teachers the broader questions that are of interest in this study (e.g., How would you define character? How would you describe CE generally and in your classroom? What choices

are you making as you plan and enact CE?). Consequently, there is much left to explore regarding teachers' conceptions of character, CE, and the relationship of those phenomena to academic endeavors.

Teachers' practices. Milson and colleagues' work found that teachers generally feel positively about teaching character, but does not report how teachers would define, describe, or enact that work in their own words. Nucci et al. (2005) found that teacher candidates' sense of efficacy for teaching character can be improved via targeted coursework, but again, this does not tell us what teachers learned or how teachers conceptualized or enacted character or CE based on that learning. In fact, teachers' "in-their-own-words" thoughts and "in-their-own-classrooms" practices are largely absent from most of the studies reviewed thus far; only Seider (2012) and Robertson-Kraft and Austin (2015) have provided much insight on this front.

Though he wrote largely about school-wide structures like community meetings, honor codes, and awards, Seider (2012) also chronicled ethics classes at Boston Prep, advisory lessons at Roxbury Prep, and civics and CE classes at Pacific Rim. These classes are the three schools' dedicated CE times, and Seider's descriptions often capture what teachers are thinking and doing regarding CE in their classrooms. Though each class takes on a different focus, depending on each school's CE focus, the curricular structures and pedagogical methods are fairly similar across the schools and classes. Students read character-related texts, participate in discussions, and write and reflect about their learning, typically with the teacher at the helm.

In Boston Prep's ethics classes, the texts might be more classical (e.g., a Platonic dialogue) or philosophical (e.g., an excerpt from Gandhi's memoir) (p. 57). At Roxbury

Prep and Pacific Rim, the texts might be more current-events focused (e.g., a video about the aftermath of Katrina or Zuckerberg's donation to Newark public schools) (p. 121 and p. 162). Regardless, texts usually serve as the point of grounding and departure in these classes. With the text as foundation, teachers help students to both define and describe the character strengths in question and, when appropriate, debate and discuss the relevant issues. While there are many different structures for student participation—from Socratic circles to "four corners" activities—the lessons are generally teacher-constructed and teacher-led. The lessons typically conclude with written reflection. Only Pacific Rim seems to offer a more "real world" practice-based approach to CE, giving students the chance to participate in a civics-focused exchange program with a Chinese school and requiring all seniors to complete a community engagement project. These, however, are school-wide structures, as opposed to classroom-based experiences.

The schools Seider (2012) studied all have particular character strengths at the heart of their missions and cultures. Consequently, the leaders and teachers who work there are—by virtue of seeking and retaining employment—either explicitly or implicitly defining and enacting character and CE as the school does. In other words, Seider did not report a wide range of perspectives on the definition of character or a teacher's role in helping students to develop it. The teachers in his study presumably supported their schools' efforts and saw teaching character, as defined and structured by the school, as part of their jobs.

Robertson-Kraft and Austin (2015),[4] however, examined the CE lessons of over

100 teachers at dozens of different schools, some with explicit character-driven missions

and some without. They found that, given the choice, teachers were overwhelmingly

planning lessons focused on achievement character (i.e., lessons focusing on grit and self-

control), as opposed to social character (e.g., optimism and gratitude) or intellectual

character (e.g., curiosity). Like the teachers in Seider's (2012) schools, all the teachers in

Robertson-Kraft and Austin's (2015) study were required to teach CE (though only one

lesson was the focus of investigation). These teachers, however, had far more freedom

than those at Seider's (2012) schools to design their own approach to CE within the

lesson of interest. Perhaps because of this freedom or perhaps because of a lack of

school-curated texts and materials, the teachers in Robertson-Kraft and Austin's (2014)

study often chose to introduce character strengths to their students via general language

and personal anecdotes, as opposed to specific definitions and texts. Yet like the teachers

at Roxbury Prep, many teachers in Robertson-Kraft and Austin's study chose to connect

grit and self-control to academic performance outcomes (e.g., end-of-year reading levels

or success in college) in presenting the strength to students. Because these lessons were

less likely to be text-based, they often skipped straight from the teacher's general

definition of the strength to student reflection (e.g., when did you demonstrate this

strength?) and discussion (e.g., how might we demonstrate this strength in class?). In

some cases, students were given opportunities to practice strategies or actions related to

---

[4] During the data collection period for this study, Austin was an Assistant Professor at the graduate school that served as the study site. Consequently, the lesson plans that she and Robertson-Kraft examined represent a subset of the plans for the lessons enacted by participants in this study. In other words, Austin and Robertson-Kraft described in their study a subset of the lessons that I investigate in this study. Their findings versus those of this study will be discussed in greater detail in later sections of this work.

the strength (pp. 122-123). The teachers in this study were also generally teaching just a one-shot CE lesson, without apparent connection to academic topics or to ongoing CE in the classroom or school (p. 120). Again, this is quite a contrast to the wrap-around approach to CE that Seider (2012) described in his schools.

If the findings of Milson and colleagues (Milson 2003; Milson & Mehlig, 2002) and Nucci et al. (2005) indicated that teachers generally feel confident teaching character, and especially so with targeted training, Seider (2012) and Robertson-Kraft and Austin (2015) suggested that CE looks quite different from teacher to teacher depending on their contexts and resources. In other words, teachers think they are good at CE, but this looks different across schools and classrooms. These differences may partially explain recent mixed findings regarding the outcomes of CE (e.g., Benninga et al., 2003 versus US DOE, 2010). Relatedly, these findings leave many more questions to answer. Are teachers choosing their approach to CE based on their school's materials as studied by Seider (2012), or on commercial curriculum as studied by the US DOE (2010)? Do they use research-based interventions? Do teachers try to enact these curricula with fidelity? Or are teachers basing their CE on their own best guesses in true, home-grown style? Regardless of the approach, do teachers believe these curricula and practices will help students to measurably improve their strength of character? What, if any, relationship do teachers see between CE and academic growth? These are some of the many questions left unanswered by the extant research that this study will attempt to investigate via teacher interviews.

Conclusion

This study defines character as comprising a plurality of observable, measurable, and ubiquitous strengths, from "appreciation of beauty and excellence" to "zest" (Peterson & Seligman, 2004). CE is defined broadly then—by Peterson and Seligman and this study—as "a deliberate intervention with the goal of teaching young people a core set of values" (p. 387). In particular, this study operationalizes this broad definition of CE in a deliberate, classroom lesson with the goal of building a particular character strength in K-12 students by defining, modeling, and increasing understanding and/or practicing the use of that strength.

With these definitions in place, the literature review above has outlined major recent findings related to measuring character and, relevant for this study, reading growth, as well as the relationship between character and academics. While many different types and instruments for measuring character exist, this study (like many) uses self-report and teacher-report survey items as the primary means of measuring students' character. Similarly, while there are a variety of ways to measure students' reading abilities, this study (like many classrooms) uses classroom-based reading growth assessments.

The relationship between certain character strengths and academic outcomes is often significant and positive, as shown in many recent studies. While none of these studies used students' reading growth as an outcome measure, cognate measures with positive and significant relationships to character abound (e.g., GPA, educational attainment, math test gains, SAT-9 scores, etc.) Moreover, there is robust and growing evidence that helping students to change or improve certain non-academic beliefs,

thoughts, or actions can have positive effects on both academic and social outcomes. This study will examine both the relationship between strength of character and academic outcomes, as well as whether an attempt to build the former is associated with positive effects on the latter.

The literature review above also contextualized contemporary approaches to CE within a century of CE in America, connecting both contemporary CE practices and the history thereof to the particular vision of CE enacted and examined in this study. This study will examine current teachers' beliefs about, goals for, and practice of CE in light of both the long tradition and the more recent trends in related social-psychological research.

From this foundational set of definitions and findings, the study will now proceed to outline its methods, drawing largely from the methods used in the scholarship described above.

III – DATA AND METHODOLOGY

Overview

This study investigated four questions:

- How do changes in grit or self-control among the students in the study sample compare to those in students in other studies using similar measures?

- How does reading growth among the students in the study sample compare to that of students in other studies using similar measures?

- What, if any, relationship exists between students' academic and character skills and growth?

- What aspects of CE curriculum, planning, pedagogy, and teachers' choices might be associated with students' academic and character growth?

To investigate these questions, this study employed a mixed-methods design (Calfee & Sperling, 2010). By combining quantitative methods (i.e., descriptive statistics, correlations, and multi-variable linear regression) and qualitative methods (i.e., video observations and interviews with a sample of the teacher participants), this study sought to present a more complete vision of enactments of CE and associated outcomes. More particularly, the study examined several key data points and sources across an entire single school year: Where were students, both academically and socially, at the beginning of the year? What did teachers do to help students develop on both fronts over the course of the year? Where did students finish the year, both academically and socially? These questions were the broad parameters of this study, but each has clearly

defined particulars (described in the sections that follow), allowing for meaningful comparisons across students, teachers, classrooms, grades, and schools.

At its most basic, this study required its teacher participants to record their students' beginning-of-school-year reading and character strength levels, as measured by classroom-based reading assessments and self- and teacher-report surveys, respectively. These served as the baseline measures. Teacher participants then taught one, structured-but-not-scripted, character-focused lesson. This was the "intervention" of interest in the study. Teacher participants subsequently measured students' reading and character levels at least three other times over the course of the school year, including the end of the year. What did reading and character change look like? Did the quality of teachers' character-focused lesson predict anything about that change? These are just two of the questions that this study examined. In the sections that follow, I describe the particulars of the data relevant to investigating the study's research questions, the procedures I followed to analyze those data and investigate those questions, and finally the limitations inherent in both the data and the methods.

Data

Participants

This study followed three cohorts of elementary school teachers[1] (n = 88) who were part-time graduate students, earning their Master of Arts in Teaching (MAT) at the graduate school of education (GSE) that served as the study site. The teachers worked with over 2,100 kindergarten through fourth grade students in over 30 different urban public charter schools in New York City (NYC). In terms of race and ethnicity, 52% (n = 46) of the teachers identified as white or Asian; 41% (n = 36) identified as black or Latinx; and 7% (n = 6) chose not to identify. The teachers in the quantitative study sample (n = 88) represent only those who met the following criteria:

- Were members of the Classes of 2014, 2015, or 2016 at the GSE

- Taught kindergarten, first, second, third, or fourth grade in a public charter school in NYC during their second year of their MAT program at the GSE

- Measured and reported (to the GSE) their students' reading growth in their classrooms, using one of two reading assessments: the Strategic Teaching and Evaluation of Progress (STEP) or the Fountas and Pinnell (F&P) Benchmark Assessment System

- Measured and reported (to the GSE) their students' grit or self-control growth in their classrooms

---

[1] In a study that takes place in a graduate school of education (GSE), the terms "professor," "teacher," and "student" are ambiguous. For the purposes of this study, a "professor" is any one of the full-time, professional teacher educators who serves as an instructor at the GSE where part of the study took place. A "teacher" is any one of the full-time K-12 teachers enrolled as part-time graduate students in the MAT program at said GSE. A "student" is any one of the elementary school students in the public school classrooms led by the teachers in the study.

Additionally, the elementary school students (n = 2144) of these teachers were only included in the quantitative study sample if their teacher met the above criteria *and* collected and reported *both* reading and character growth data for the students. The students in the sample are further described in Table 1.

Table 1

Student Sample Sizes by Grade, Character Strength of Focus, and Reading Assessment

Grade	Students	Character Focus		Reading Assessment	
		Grit	*Self-control*	*F&P*	*STEP*
K	646	446	200	213	433
1	591	423	168	240	351
2	524	439	85	214	310
3	280	231	49	120	160
4	103	103	0	80	23
All	2144	1642	502	867	1277

As can be seen in the table, the vast majority of the students in the sample were lower elementary school students (n = 1761), and most were learning from teachers who chose to focus students on grit (n = 1642) instead of self-control (n = 502) in their CE lesson. The STEP reading assessment was the assessment of record for more students (n = 1277) than the F&P reading assessment (n = 867). Not represented in the table, but represented in the data, the students were enrolled across more than 30 different schools, in geographically disparate parts of NYC. The data set does not contain student demographic data, such as gender, race, or socio-economic status. However, based on publicly available school-level data, one can assume that the vast majority of students identify as African American and/or Latinx and qualify for free or reduced-price lunch.

That the teachers disproportionately identify as black or Latinx—as compared with those in charter schools, NYC, or New York State broadly—is not coincidental.

The GSE that served as the study site has an explicit commitment to recruiting and retaining black and Latinx teachers into the profession for many reasons, not the least of which being the well-documented benefits of black and Latinx teachers to their students (Education Trust—New York, 2017).

Additionally, that all the schools in the study are public charter schools (as compared with public district schools or independent schools), is in part reflective of the composition of the student body of the study site at the time of the study, in part reflective of NYC's hiring practices, and in part reflective of methodological choices. The GSE that served as the study site enrolled more charter teachers than district teachers at the time of the study. Though the enrollment balance at the GSE has shifted closer to half-and-half in the years subsequent to the study, NYC then and now hires predominantly only middle and high school teachers through alternative certification programs (such as those offered at the study site). Consequently, the vast majority of the teachers working in district schools who were enrolled at the GSE during the study were teaching in secondary grades, and thus not measuring students' reading growth. (See the section on measuring reading growth, below, for more on why secondary teachers typically do not assess it.) While there were a handful of teachers working in elementary classrooms in district schools enrolled at the GSE at the time of the study, creating statistical controls and otherwise accounting for district-versus-charter differences for such a small number of teachers was deemed more complicated than necessary, and thus neither those handful of teachers nor their students are represented in the study.

Quantitative Data and Analytic Approach[2]

  Outcomes. This study focused on two outcome variables: a measure of students'

character growth and a measure of students' reading growth. As a measure of students'

character growth, the study used the difference between each child's beginning-of-year

and end-of-year ratings on survey items associated with a particular character strength.

As a measure of students' reading growth, the study used the difference between each

child's beginning-of-year and end-of-year grade-level equivalency scores in reading.

These outcomes are further described below.

  Character growth. Each teacher enrolled at the GSE at the time of the study

picked one of eight sets of survey items, each associated with a different character

strength (many as described and validated in Park, Tsukayama, Goodwin, Patrick, &

Duckworth, 2017). As mentioned in Chapter I, of these eight strengths, the study focused

only on two, i.e., grit and self-control, in the analysis. This choice was motivated by two

factors: (1) of the eight strengths, grit and self-control have the most validated measures

and are the most widely cited in the research literature and (2) over half of the teachers in

the potential study sample chose to focus on either grit or self-control with their students.

The survey items used a 5-point Likert scale ranging from 1 = *very much unlike me* to 5 =

very much like me, and the items associated with grit and self-control are below (along

with all the items in Appendix F):

  Grit
- Finishes whatever he or she begins
- Tries very hard even after experiencing failure
- Works independently with focus

---

[2] As part of my pilot work, I collected and organized the data described in the sub-sections that follow.
This quantitative data set was also used as the basis for the analyses in this study, and in turn drove my
selection of participants for the qualitative portion of the study.

Self-Control
- Allows others to speak without interruption
- Is polite to adults and peers
- Keeps his/her temper in check

Teachers administered these assessments to their students four times over the course of an academic year: once before the intervention and three times after the intervention, with the final administration occurring at or near the end of the school year. A character growth measure for each student was calculated by subtracting beginning-of-year self-ratings from end-of-year self-ratings. Teachers also rated each student at these same times on these same items. Consequently, the study examined character scores and growth derived from both student self-ratings and teachers' ratings of students.

Reading growth. As described briefly in Chapter II, the teachers in the study used classroom-based reading assessments to assess their students' reading abilities at least four times over the course of the academic year. Classroom-based reading assessments are relatively brief assessments of children's reading ability, designed to be administered by teachers to individual children in a typical school setting and structured to provide the teacher with several types of information regarding the child's current reading level, as well as particular strengths and weaknesses. The assessments comprise items associated with the components of reading relevant for the particular reading level in question. For example, assessments of reading levels typically seen in kindergarten and first-grade students often contain items related to concepts about print (e.g., looking from left to right across pages) and phonemic awareness. Assessments geared toward higher reading levels (like those typically seen in second and third grade), however, drop these foundational components and focus on reading accuracy, fluency, and comprehension (Kerbow & Bryk, 2005). Teachers are trained to record students'

performance on these tasks in terms of accuracy, rate, or other criteria depending on the task. These performances are then turned into a score on that component of the test; these component scores are then aggregated to determine the child's reading level and corresponding grade-level equivalency (GLE). As measured by these normed assessments, a child's expected reading growth each year would be one GLE.

Each teacher at the GSE selected one of five validated reading growth assessments, each of which uses GLE scores as a unit of measure: F&P, STEP, Developmental Reading Assessment, Second Edition (DRA2); Reading Curriculum-Based Measurement (RCBM); and Reading A to Z (RA2Z). Of these five options, in the analysis the study focused on only two: F&P and STEP. This culling was motivated by two factors: (1) F&P and STEP met the study site's highest bar in terms of reliability and validity (Study Site Graduate School of Education, 2014) and (2) over 90% of teachers in the potential study sample chose to use either F&P or STEP. While each of these tests uses slightly different tasks, sequences, and/or scoring systems, all of these assessments can be used to calculate a child's reading GLE, allowing for comparisons across assessments, classrooms, and schools. Teachers administered these assessments to their students at least four times over the course of an academic year: once before the intervention and three times after the intervention, including one measure at or near the end of the school year. A growth measure for each student was calculated by subtracting the beginning-of-year GLE score from the end-of-year GLE score.

Predictors: Teacher Instruction. This study examines one key predictor variable: teachers' instruction within a CE lesson. This measure is described below.

During the winter of teachers' second year in the MAT program at the GSE, members of the faculty evaluated each teacher in two separate lessons, one academic-focused and one character-focused, using a specific observation rubric for each lesson (see Appendix G for the academic lesson rubric and Appendix H for the CE lesson rubric). Each rubric measured several teaching skills. Each skill is represented by a separate strand of the rubric using a 5-point scale ranging from 0 = *lacking* (when there is no evidence that the teacher attempted the skill) to 4 = *exemplary* (when the teacher used the skill in particularly advanced and/or effective manner). In addition to skill-by-skill scores, each rubric has a "switch row" at the end, in which the faculty member gives the teacher a holistic score (again, from 0 to 4) summarizing overall performance with respect to the skills represented in the rubric. Consequently, for each rubric, teachers earned skill-by-skill scores and an overall score.

These measures of particular aspects of instruction were included as predictor variables because my conceptual framework and literature review suggest that there are certain elements of CE instruction which are crucial to leading effective CE lessons (e.g., clear and explicit definitions of the strength in question; presentation of clear models of the strength; opportunities for students to better understand the strength via reading, writing, or discussion; and opportunities for students to practice skills associated with the strength). Consequently, the CE lesson rubric focuses on these aspects of instruction and gives each teacher a score on each of these components. Thus, one can use these data to examine whether or not there is a relationship between the presence and nature of these aspects of instruction and students' character and academic gains. The academic lesson

rubric did not contain these CE-specific elements, but served as an important comparative predictor.

      Covariates. Along with the outcomes and predictor described above, the data set also includes school identification data, as well as teachers' professional data (e.g., grade taught) and performance data (i.e., GPA). Many of these variables have been associated with student outcomes and/or teacher's perceived self-efficacy in previous studies, and thus may explain part of the variance of the outcomes. For example, school effects, like the extent to which a school is engaged in school-wide CE (beyond the one-lesson intervention in this study) may contribute to students' academic and social growth (Benninga et al., 2003; Berkowitz & Bier, 2005; CEP, 2010). Unfortunately, the school-level data in the data set were insufficient to include as covariates in regression models, as the data set contained only school names and, where applicable, the name of the broader system to which the school belongs (i.e., the parent charter management organization). Therefore, while these data were helpful in (a) examining school-based trends in academic and character outcomes and (b) identifying a qualitative sample, school-level data were not included in any models as covariates.

      The grade level data in the data set were simple: the grade each teacher taught and the grade that each student was in. Students' grade levels, of course, are associated with their reading levels (with younger students, on average, reading at lower GLEs than older students), but also with their reading growth. Students in lower elementary classrooms, on average, demonstrate more growth in their reading skills each year than their upper elementary peers (Hill, Bloom, Black, & Lipsey, 2008). Consequently, I created within-

grade standardized reading measures for all correlational analyses. I did not ultimately use students' reading levels in the regression models.

Some evidence suggests that grade level may also explain part of students' character growth. For example, Milson (2003) found that teachers' perceived self-efficacy in leading CE was related to the grade that they taught, with teachers of younger students, on average, feeling more efficacious than their colleagues teaching older students. However, when I examined the descriptive statistics associated with character growth by grade in the sample, no clear pattern emerged. Consequently, I did not create standardized-within-grade character measures, nor did I use grade level as a covariate in the regression models.

Turning from school and grade to the teachers themselves, teachers' academic performance has been shown to be associated with their students' academic growth (Kukla-Acevedo, 2009), and thus I created a standardized measure of academic performance (GPA) in the data set. As suggested by earlier studies, teachers' GPA was associated with their students' reading growth in the study sample; it was not, however, associated with students' character growth. Consequently, when reading growth was the outcome, I included GPA as a covariate in some pilot models; when character growth was the outcome, I did not include GPA as a covariate in any models.

Finally, several studies have shown that students' initial reading levels are associated with their reading growth (e.g., Ready, 2013; Pfost, Hattie, Dörfler, & Artelt, 2014). The same proved true in this data set, and thus, when reading growth was the outcome, I included initial reading level as a covariate in some pilot models. By analogy, I hypothesized and confirmed that students' initial character levels would be associated

with their year-long character growth. Thus, initial character levels served as covariates in the models with character growth as the outcome.

Analytic approach. The analytic approach I employed with the quantitative data described above was somewhat different for each research question and is described briefly below.

How do changes in grit or self-control among students in the sample compare to those students in other studies using similar measures?

To examine this first research question, I calculated basic descriptive statistics associated with the beginning-of-year character and character growth for all (n = 2144) students in the study. I compared these data with descriptive statistics from studies examining student participants and similar measures (e.g., Seider, 2012; West et al., 2014). Additionally, I conducted correlational analyses to examine the relationships between initial character status and character growth. Finally, I examined the relationship between students' perceptions of their own character growth and levels *and* teachers' perceptions thereof.

How does reading growth among the students in the study sample compare to that of students in other studies using similar measures?

To examine this second research question, I calculated basic descriptive statistics associated with the initial reading status and reading growth for all (n = 2144) students in the study. I compared these data with the expected outcomes of the normed, classroom-based reading assessments. Additionally, I conducted correlational analyses to examine the relationship between initial reading level and reading growth.

What, if any, relationship exists between students' academic and character skills and growth?

To examine this third research question, I conducted correlational analyses to examine the relationships between reading status and growth *and* character status and growth measures across the students (n = 2144) in the study.

What aspects of CE (i.e., curriculum, pedagogy, or teachers' conceptions) might be associated with students' academic and character growth?

To examine this fourth and final research question, I first calculated basic descriptive statistics associated with teachers' performance on the academic and CE lesson rubrics. I then conducted correlational analyses to examine the relationship between teachers' performance on the CE lesson rubric and the academic lesson rubric. Next I conducted additional correlational analyses to examine the relationship between teachers' performance on these rubrics and their students' academic and character growth and end-of-year outcomes. Finally, I built several multi-variable linear regression models to examine how these relationships varied by CE lesson component.

Because this is a mixed-methods study, the quantitative data and analytic approach described above represented just a part of the data and analysis involved in this study. I conducted the quantitative analyses first, and those analyses suggested additional avenues to explore in my qualitative work. For example, because the quantitative analyses indicated that students whose teachers focused on grit (instead of self-control) had greater year-long character growth, I ensured that my qualitative sample allowed for the investigation of the differences in CE instruction in classrooms focused on grit versus those focused on self-control. In the sections that follow, I describe my qualitative data and analytic approach.

Qualitative Data and Analytic Approach

While the study's first three research questions were intended to be purely quantitative, the final and broadest question has intentional qualitative elements: *what aspects of CE curriculum, planning, pedagogy, and teachers' choices might be associated with students' academic and character growth?* To investigate qualitative aspects of this question, I collected and analyzed two types of qualitative data: observations of a sample of videos of teachers' CE lessons, followed by interviews with a subset those teachers. Below I describe both the participants for the qualitative portion of the study, as well as the methods for video review and interview protocols.

Participants. I anticipated four distant-from-the-mean categories of teachers within the full quantitative sample: teachers whose classes had (1) above average reading and character growth, (2) above average reading growth, but below average character growth, (3) below average reading growth, but above average character growth, or (4) below average reading and character growth. To ensure a notable qualitative sample, I determined that I would seek out teachers whose class average reading growth was at least .5 SDs above or below the within-grade sample mean, and whose class average character growth was at least .5 SDs above or below the sample mean.

These parameters left me with 53 teachers divided into the four groups described above. Using these four groups, I tried to create stratified subsets of each to compose the final study sample. Several factors determined which teachers were removed and which teachers remained in the final sample. First, I wanted to examine both grit and self-control lessons within each sub-group. Second, I wanted to examine, whenever possible,

a range of grades and schools within each sub-group. And while teacher race and ethnicity were not related with the major quantitative outcomes, I wanted to have a representative qualitative sample in that respect, as well. Finally, there were several unexpected factors that caused me to remove some teachers from the final qualitative sample (e.g., one teacher had been my student when she was in eighth grade, one teacher had not ultimately graduated from the program, etc.). With these selection factors in place, the final qualitative sample (n = 14) emerged as three-fold:

- **Group I:** Teachers who led their students to above-average reading and character growth (n = 7)

- **Group II:** Teachers who led their students to above-average reading growth, but below-average character growth (n = 4)

- **Group III:** Teachers who led their students to below-average reading and character growth (n = 3)

The sample is described further in the Table 2.

Table 2

Sample of Teachers (n = 14) for Qualitative Analysis

Group	Teachers	Grades	Focus Strength	
			Grit	*Self-control*
I	7	K (n = 4), 1, 2 3	5	2
II	4	K (n = 3), 2	1	3
III	3	1, 3 (n = 2)	2	1

As noted above, I intended to have four groups of teachers in the qualitative sample, but the above-average reading and below-average character group was the smallest identified from the original quantitative sample, and each of its members fell

into one of the "miscellaneous" categories for culling. Thus, there was no teacher who led his students to below average reading growth, but above average character growth in the final qualitative sample.

Additionally, my intent to have a somewhat stratified sample—with respect to grade, strength, school, and teacher race and ethnicity—within each of the purposive groups was generally realized within two of these groups. As shown in Table 2, each group comprises at least two grade levels *and* classrooms focused on both grit and self-control. Not depicted in the table is the fact that the seven teachers in Group I work at six different schools (and all six are run by different charter management organizations [CMOs]). The four teachers in Group II work at three different schools, again all run by different CMOs. The three teachers in Group III, however, work at two different schools, both run by the same CMO. In other words, Groups I and II represent a relatively stratified sample of schools, and thus are more likely to be generalizable to the larger quantitative sample. Group III, however, represents a relatively homogenous sample of schools, and thus is less likely to be generalizable to the broader sample. In sum, the qualitative sample as constructed generally met the criteria of the sample as proposed and conceived, with the exception of Group III's lack of stratification with respect to school.

The participant selection process entailed two steps: first, selecting the final sample for CE lesson video review (as described above) and then, recruiting these teachers for interviews. The intent was to view and analyze these 14 teachers' CE lesson videos using the protocol described in Appendix L, and then interview those same teachers using the protocol described in Appendix N. While I reviewed and analyzed all 14 CE lesson videos, I was only able to recruit six of these teachers for interviews. Of

those six, four are in Group I (i.e., above average reading and character growth) and two

are in Group II (i.e., above average reading and below average character growth). No

teachers from Group III (i.e., below average reading and character growth) responded to

my requests for interviews. Consequently, the video findings presented below represent

the entire intended sample, while the interview findings represent a subset of that sample,

and a non-representative sub-set at that.

All interviewees provided written permission in advance of my interviewing

them. Participation was entirely optional, and all prospective participants had ample

opportunity to discuss the research with me in advance of reviewing a participant's rights

document and deciding whether or not to sign a letter of consent (see Appendix K) and

participate in the research. Below I briefly describe the procedure for collecting and

analyzing classroom videos and for interviewing each participating teacher.

Videos. In the late fall of their second year in the MAT program at the GSE that

served as the study site, the teachers took coursework entitled "Teaching Character

Strengths," introducing research-based approaches to CE and social-psychological

interventions. (See Appendix I for more context and details regarding the GSE

coursework associated with the study.) As part of this coursework, teachers read sample

CE lesson objectives, along with selections from Seligman et al. (2009) presenting

headlines from social-psychological interventions; the entirety of Wilson's (2006) brief

description of some social-psychological interventions that improved students' academic

outcomes, and a piece from Tough (2012) outlining Oettingen's findings regarding her

goal-visualization protocol. In the class, the teachers also viewed classroom video of a

lesson focused on love in a middle-school classroom, in which students identified loved

ones and wrote them letters. They also participated in a mock lesson (taught by their professor) focused on self-control, in which the professor defined the strength, presented a rationale for building it (i.e., improved social and cognitive outcomes), and gave students a chance to practice by leaving their cell phones out, but untouched for the duration of the class night. Teachers also practiced teaching (and received feedback on) the definitions and rationales they intended to use in their classrooms with their students. In short, the teachers' introduction to CE in their graduate coursework was a hybrid of the social-psychological literature and "home-grown" CE lessons.

As outlined in Chapter I, the coursework presented a framework for these one-shot "home-grown" CE lessons, predicated on the literature graduate students read, as well as other frameworks for CE (notably, Seider, 2012). Teachers were instructed to present students with (or elicit from students) a clear definition of the strength in question, as well as a compelling rationale for having, using, and/or developing that strength. Additionally, teachers were instructed to model (or have students model) what that strength looked in action. Modeling, in general, and examples and enactments, in particular—as described in the CE coursework and lesson rubric—were intended to give students a clear vision of what the strength looks like in action. Finally, teachers were instructed to give students practice opportunities to better understand the strength and/or use it real-time in the classroom. This framework was presented as both part of the coursework and the evaluation for the teachers' final assessment: a CE lesson.

For the CE lesson assessment, each teacher planned, taught, and filmed a character-focused lesson in his classroom. Professors then reviewed these tapes and recorded each graduate student's score on a common, 5-row, 5-point rubric (see

Appendix H). As noted above, teachers' scores on these rubric criteria served as

quantitative predictors in this study.

I reviewed the CE lesson videos of the teachers in my sample (n = 14). The

substance of those videos served as qualitative data demonstrating what the teachers (and

their students) were doing when CE was the instructional focus of their classrooms. In

particular, I investigated to what extent the teachers were employing the structures and

strategies recommended by Seider (2012) and contemporary CE scholars, as well as what

(if any) aspects of recent social-psychological intervention work teachers integrated into

their CE lessons. The coding protocol, detailed in Appendix L, involved three reviews of

the footage in question. In the first review, I simply watched the footage of the CE

lesson. In the second, I kept a running record of the lesson. In the third, I used an every-

minute pause to focus on identifying and coding the CE-related instructional structures

and strategies of the previous minute of instruction, such as a teacher or student

presenting an explanation of a character strength, or students discussing some character-

related aspect of a text.

Interviews. I conducted interviews with a subset of the sample of teachers whose

CE lesson videos I reviewed. The purpose of the interview was to better understand how

and why the teachers made the choices that they did in their CE lessons, as well as to

better understand their broader conceptions of character, CE, and the relationship

between those phenomena and academic learning. The interviews were semi-structured,

but followed a standard set of questions (see Appendix N for all questions).

Representative questions included "What does 'character education' look like in your

classroom?", "Why did you choose to focus on grit [or self-control] in your classroom?",

"How would you determine if CE is 'working' in your classroom?", and "Why did you choose the type and number of practice opportunities that you did [for students within the CE lesson we watched on video]?".

Procedurally, after teachers agreed to be interviewed, I sent them their CE videos and lesson reflections in advance of our interviews, so that they could refresh their memories regarding the lesson in question prior to our interview conversation. I also sent each interviewee a brief pre-interview survey. The pre-interview surveys served two purposes: (1) to gather some current information about each interviewee (e.g., tenure-to-date as teachers, current place/nature of employment, etc.) and (2) to gather interviewees' preliminary thoughts about the efficacy of CE. This latter part of the survey comprised 12 items adapted from Milson and Mehlig's (2001) Character Education Efficacy Belief Instrument (CEEBI). The CEEBI consists of 24 5-point Likert-scale items focusing on personal teacher efficacy (PTE) and general teacher efficacy (GTE). Of these 24 items, I chose and adapted 12 for the pre-interview survey: 6 PTE and 6 GTE items. Three additional items were simply quantitative versions of the questions I asked in the interview (e.g., "CE is working in my classroom."). The entire pre-interview survey can be found in Appendix M.

I calculated basic descriptive statistics from the pre-interview survey and transcribed all of the interviews. I used the survey data and the transcripts to describe the patterns across the sample, as well as the connections (and lack thereof) between the teachers' conceptions of character, CE, and their classroom practices and the relevant published literature. The analysis of the pre-interview surveys and interviews served as an additional examination of how teachers' pedagogical vision, instructional choices, and

conceptual understandings might be associated with students' social and academic growth.

<center>**Limitations**</center>

The limitations in this study fall into two broad categories: those of positionality (e.g., the author's preferred epistemology, relationship to the participants, etc.) and those of data, measures, and methods. Because I addressed many of the former limitations in the "Researcher Positionality" section in Chapter I, here I will focus exclusively on the limitations associated with data, measures, and methods.

Data and Measures

The limitations associated with the data and methods accounted herein comprise both those foreseen before the study commenced and those identified while conducting the work, analyzing the data, or reviewing the findings with colleagues. I will do my best to give a thorough accounting of all such limitations, but there are likely also limitations that I have not yet identified. Consequently, what follows is my best attempt to identify the limitations of the data and methods in question; it is, undoubtedly, only a partial accounting.

Both philosophically and methodologically, the culling of the data set from eight character strengths of interest to only two (i.e., grit and self-control) is a major limitation of the study. This narrowing took place, as noted previously, as a result of (a) the majority of teachers at the GSE choosing to focus on either grit or self-control in their classrooms and (b) that those strengths had the most validated measures and comparable

findings of the eight strengths. That said, I adopted Peterson and Seligman's (2004) pluralistic view of character (encompassing 24 strengths in their model) as the theoretical foundation for this study in part because it mirrored the many strengths I see in humanity. That my own work would necessitate culling this great many to only two for methodological reasons was a great disappointment, a paltry representation of how I envision character, and a major limitation of the study. I was not able to truly investigate how teachers "cultivate character" in elementary school classrooms; I was only able to investigate how teachers cultivate grit and self-control.

As mentioned several times heretofore, the context in which schooling, and CE in particular, takes place undoubtedly influences both the education that students received and the related outcomes. For example, there were likely school-based factors that influenced the quantitative findings in this study, but I was not be able to determine this as the data set contained no school-level predictors. Additionally, because all the data in this study are associated with teachers working in urban public charter schools, they are less generalizable to other school settings. The same may be said of the fact that all the teachers in the study were attending the same GSE and were all teaching elementary school students during the time of the quantitative data collection. In short, several contextual factors—from urban public charter school teaching placements in primary grades to formal teacher preparation—may limit the generalizability of the findings.

Missing data is often a limitation in education research (Peugh and Enders, 2004), as was also the case in this study. While the student sample was composed to mitigate this factor in some respects (i.e., only students with both reading and character measures were included in the sample), both the student and the teacher-level data had some gaps.

As one example, while most of the teachers had self-identified race/ethnicity data, none of the students in the data set were identified in this way, nor were either teachers or students identified by gender. As an additional example, only a subset of the teachers (n = 55/88) had scores for the academic lesson. This is because only two of the three cohorts represented in the sample (i.e., the Classes of 2015 and 2016) used the same academic lesson rubric (see Appendix G). The other cohort (the Class of 2014) used a different measure, and consequently, their academic lesson scores are not included in the data set. While this was an unforeseen limitation, it did not have much bearing on the analyses, as (a) the majority of the teachers in the sample did have academic lesson scores and (b) the academic lesson data were only included as a comparative predictor. Additionally, the final data set included school identification data for most, but not all of the students (n = 1533/2144). The GSE did not have school placement data for the other 611 students in the quantitative sample.

The measures used in the study also have limitations. The character growth survey (Park et al., 2017) used to produce both the student- and teacher-report character data for the students in the study is subject to all the limitations of surveys described in Chapter II (and in more detail in Duckworth and Yeager's [2015] review of the means of measuring personal qualities). From reference bias, to social desirability bias, to faking, self- and other-report surveys have all these limitations and more. While these limitations are not insignificant, the advantages of using self- and other-report surveys to measure students' strength of character are well documented in Chapter II and outweigh the limitations noted here.

Additionally, the actual survey items used by the teachers and students in this study are not an exact match with the validated versions in Park et al. (2017). As a minor example, sometimes the vocabulary or phrasing of the items was changed by an individual teacher to make them more accessible to an elementary school student audience. Moreover, Park et al.'s items are framed in the past tense, while the items in this study use a present tense structure. Comprehensively, Tables 3 and 4 demonstrate the differences between the student-facing versions of the relevant items in Park et al. and those administered by and to the teachers in this study.

Table 3

Grit Survey Items

Item	Park et al. (2017) Version	Study Version
1	I finished whatever I began	I finish whatever I begin
2	I tried very hard even after experiencing failure	I try very hard even after experiencing failure
3	I stayed committed to my goals even if they took a long time to complete	I work independently with focus
4	I kept working hard even when I felt like quitting	

Table 4

Self-control Survey Items

Item	Park et al. (2017) Version	Study Version
1	I remained calm even when criticized or provoked	I remain calm even when criticized or provoked
2	I allowed others to speak without interruption	I allow others to speak without interruption
3	I was polite to adults and peers	I am polite to adults and peers
4	I kept my temper in check	I keep my temper in check

The primary reason for the differences between Park et al.'s (2017) items and the items used by the teachers in the study is the existence of a pilot set of items that Duckworth's lab (of which Park and colleagues were members) was using from 2011–2014, prior to the published validation of those measures by Park et al. Because the character growth measurement conducted by the teachers in this study began in 2013, the teachers used the pilot survey items, as opposed to the final items described above.

This lag between the availability of pilot measures in 2011 and the publication of the validated measures in 2017 also caused another limitation: the absence of comparative descriptive data. Park et al. (2017) did not report descriptive statistics at the item level. Instead, they combined their related items into factors (e.g., the self-control items become part, but not all, of an "interpersonal" character factor), and reported the descriptive statistics for those factors. Moreover, they did not report point-in-time or growth measures for those factors; instead, they averaged initial and end-of-year measures together. Finally, their sample represented only fifth through eighth grade students, as compared with this study's sample of kindergarten through fourth grade students.[3] Consequently, it is fairly meaningless to compare Park et al.'s findings regarding students' character levels and growth to those reported in this study.

Fortunately, there are two published studies that used and reported similar grit measures for school-aged children (Seider, 2012; West et al., 2014). There are unfortunately no studies that used *and* report similar self-control measures. Several studies implemented similar self-control measures (e.g., Tsukayama et al., 2013), but

---

[3] Initially this study had a kindergarten through 12th grade sample, but the decision to limit it to kindergarten through fourth grade in order to maximize the reliability and validity of the academic outcome measures resulted in the absence of perfectly comparable character growth data.

none reported basic descriptive statistics that could be used as points of comparison with this data set. Additionally, no published study reported grit or self-control data for children as young as those in my study.[4] Thus, I could only contextualize the character skill and growth findings in broad terms, based on relatively comparable findings from Seider's (2012) and West et al.'s (2014) middle school samples.

One additional limitation of the character measures used in the study was the potential for ceiling effects, especially in the context of repeated measures across the course of a single school year. As the survey items utilize a five-point Likert scale, if a student rates himself a 4 on one or multiple measures at the beginning of the year, there is not much room to measure growth. I will address this limitation further in Chapter IV.

Classroom-based reading assessments also have well-established limitations. For example, whereas standardized achievement tests of reading produce reliable measures of children's point-in-time reading abilities, classroom-based reading tests are less reliable (as they are teacher-administered and teacher-scored). The GSE that served as the study site had a periodic spot-checking and final auditing process to guard against the most extreme limitations associated with teacher-reported student achievement results (e.g., missing data, inaccurate data, or false data) and reported no concerns among the cohorts in the study. These processes, while reassuring, do not preclude the possibility that some of these extreme limitations are factors in the final data set. Moreover, these precautions fail to guard against some of the other limitations associated with teacher-reported data (e.g., selective, subtle "rounding up," selective rostering, etc.)

---

[4] While on the one hand, this lack of same-age comparison group is a limitation, on the other hand, this data set and the findings therein may be a helpful contribution to the literature.

Methods

Both the data collection and analysis conducted in this study have limitations. With respect to data collection, the biggest limitations are two-fold: (1) teachers themselves collected the data during prescribed (but relatively wide) collection windows and (2) teachers were incentivized to help their students reach certain growth or absolute achievement outcomes. When teachers collect and report data about their own students, they are likely susceptible to the same types of reference bias, social desirability bias, and faking that are limitations of self- and other-report surveys. Moreover, when there are incentives for their students to grow a certain amount or reach a certain goal, those limitations are likely heightened. For example, the teachers in this study had to help their students reach an average of one year's worth of reading growth over the course of one school year in order to graduate from the MAT program. In other words, teachers were asked to submit accurate and up-to-date records of students' reading data over the course of a single academic year as benchmarks of student progress. At the end of the year, those data had to show that a teacher's students had demonstrated an average of at least one year's worth of reading growth during that time in order for that teacher to graduate from the MAT program in question.

Similarly, although the stakes were not as high, teachers were also challenged to help their students meet an end-of-year character average of 4.0 (on the 5-point Likert scale). There was no graduation requirement or academic incentive around character growth, but the simple suggestion of a class-wide, end-of-year average of 4.0 as a meaningful goal may have been perceived as an incentive by goal-driven teachers. These requirements and goals may have lent themselves to some teachers consciously or

unconsciously "rounding up" (or outright misrepresenting) some students' performance on either the reading assessments or the character growth surveys. The GSE reports that no teacher in these cohorts was identified as having misrepresented students' data or outcomes, after auditing a subset of teachers' results every year, but this does not preclude it as a possibility.

Additionally, there was undoubtedly variation in when and how teachers administered both the reading assessments and character growth surveys to their students. For example, the GSE required teachers to administer these assessments and report these data within approximately one-month windows, and all teachers in a given cohort had the same due date for submitting each round of data. In a school year that is only nine months long, however, there can be quite a difference between a student's beginning-of-year reading status as measured in early September versus early October. While one could somewhat mitigate the potential implications of these test administration date ranges by controlling for the number of days between the initial and end-of-year measures, this proved practically difficult because the format of the dates in the data set varied widely, making it cumbersome to calculate for all 2144 students. Consequently, I did not control for the range of test administration dates, relying on their relative proximity to minimize potential related differences in outcome.

Relatedly, there were likely differences in the manner of administering the character growth surveys. While classroom-based reading assessments generally include guidelines for how to present and administer the assessment to students, the character growth surveys came without such protocols. Consequently, some teachers may have administered the first round of surveys after a rousing pep talk about the class's great grit,

for example, perhaps inflating beginning-of-year scores and thus deflating overall growth. Other teachers may have administered the first round of surveys after giving students a stern talk about the importance of building grit, perhaps deflating beginning-of-year scores and thus inflating overall growth. These are all, of course, only hypothetical situations, but all represent possible limitations of a survey instrument that is administered by practitioners as they wish, as opposed to researchers using a scripted protocol in a controlled environment.

In the qualitative portion of the study, aside from the limitations of the sample itself described above, the pre-interview survey instrument that I used did not allow for a like-to-like comparison with its validated parent survey (i.e., Milson and Mehlig, 2001). This was due to two reasons: (1) I shortened the instrument (as described above) and (2) I adjusted language within the items to reflect a focus on just grit and self-control. The Milson and Mehlig portion of the pre-interview survey, though, was not intended as a valid quantitative comparison instrument, but rather as a roughly comparative means of collecting some CE self-efficacy data from the teachers in the interview sample.

While the data and methods had relatively robust limitations, the analytical limitations were comparatively modest. The limited nature of the school-level data (i.e., that, as described above, they were useful for pattern identification but not appropriate for inclusion as covariates in the regression models) was one analytical hindrance. Additionally, these analyses do not account for the clustering of students within teachers, or teachers within schools. Future studies might explore these questions using multi-level methods both to address the independence of error term issue, but also to examine how classroom and school contexts influence the student-level associations reported here.

There were also analytical limitations imposed by using only one year's worth of data for each teacher. As Rockoff (2004) noted regarding his analysis of teachers' effects on students' academic achievement, "Teacher effects cannot be separated from other classroom-specific factors . . . because teachers were only observed with one class of students" (p. 247). This study, too, only examined teachers' work with one group of students over one school year, and thus classroom effects could not be meaningfully separated from teacher effects.

Finally, I did not foresee the analytical limitations of the edited CE lesson video footage. When viewing teachers' videos, the footage would often cut from one classroom moment to the next, leaving some unquantifiable amount of time and teaching "off camera," sometimes seemingly only a second and other times seemingly many minutes of instruction. Teachers did this to ensure that they were highlighting the most salient parts of the lesson. However, it meant that one could not discern *all* that transpired, but only what portions the teachers selected for submission.

In this section, I have described the limitations associated with the quantitative aspects of the study, as well as with the qualitative data. As I present the findings in the next two chapters, I will also identify when and how these limitations may have had meaningful effects on the study's results.

IV—QUANTITATIVE FINDINGS AND ANALYSIS

In this chapter, I first present my quantitative findings by research question. I also present salient quantitative findings relevant to—but beyond the most limited conception of—my research questions. Whenever possible, I present my quantitative results along with contextualizing data and/or findings from previously published literature. However, I leave the most in-depth connections between my findings and the extant literature to the discussion in Chapter VI.

Character Growth

Research Question #1: How do changes in grit or self-control among the students in the study sample compare to those in students in other studies using similar measures?

On average, students in the study (n = 2144) saw themselves as starting the school year with higher levels of grit (M = 3.58, SD = 0.98) or self-control (M = 3.66, SD = 0.94) than their teachers (n = 88) reported seeing in them (M = 2.93, SD = 1.05 for grit; M = 3.34, SD = 1.14 for self-control). By the end of the school year, however, students' and teachers' perceptions, on average, were relatively similar, falling within a 0.11-point range on the 5-point instrument ($4.01 \leq M \leq 4.12$). The descriptive statistics associated with students' self-reported and teacher-reported character scores at each survey round are summarized by strength in Tables 5 and 6.

Table 5

Students' Average Self-reported Character Scores by Survey Round

Strength (n)	R1	R2	R3	R4
	(SD)	(SD)	(SD)	(SD)
Grit (1642)	3.58	3.78	3.95	4.10
	(0.98)	(0.92)	(0.81)	(0.75)
Self-control (502)	3.66	3.75	3.83	4.04
	(0.94)	(0.88)	(0.80)	(0.77)

Table 6

Students' Average Teacher-reported Character Scores by Survey Round

Strength (n)	R1	R2	R3	R4
	(SD)	(SD)	(SD)	(SD)
Grit (1642)	2.93	3.37	3.70	4.01
	(1.05)	(0.97)	(0.95)	(0.88)
Self-control (502)	3.34	3.58	3.82	4.12
	(1.14)	(1.04)	(0.90)	(0.87)

The tables above indicate point-in-time character scores, though character *growth* was the focus of this research question.[1] The next set of tables describes those findings: students' year-long character growth (as calculated by the difference between R4 scores and R1 scores) was larger when using the teacher-report measures (grit M = 1.07, SD = 0.91; self-control M = 0.78, SD = 0.95) than when using the student-report measures (grit M = 0.51, SD = 1.02; self-control M = 0.37, SD, 0.95). Moreover, both student-reported and teacher-reported calculated growth was greater in classrooms focused on grit than in those focused on self-control. Growth from one survey round to the next was

---

[1] That both students' and teachers' perceptions of students' character status—in terms of both grit and self-control—tend to converge at the end of the year just above four (with the smallest SDs of the year) on the five-point Likert scale is an indication that the measures were subject to ceiling effects.

inconsistent by round, strength, and source. For example, students focused on grit

reported less (calculated) growth with each successive survey round, while students

focused on self-control reported over double the (calculated) growth in the final survey

interim (M = 0.21, SD = 0.62) than the first (M = 0.09, SD = 0.79). The descriptive

statistics associated with students' self-reported and teacher-reported character change

from one survey round to the next (i.e., R1 → R2), and from beginning-of-year to end

(R1 → R4), are summarized by strength and source in Tables 7 and 8.

Table 7

Students' Average Self-reported Character Change

Strength (n)	R1→ R2 (SD)	R2→ R3 (SD)	R3→ R4 (SD)	R1→ R4 (SD)
Grit (1642)	0.20 (0.85)	0.17 (0.72)	0.15 (0.71)	0.51 (1.02)
Self-control (502)	0.09 (0.79)	0.08 (0.66)	0.21 (0.62)	0.37 (0.93)

Table 8

Students' Average Teacher-reported Character Change

Strength (n)	R1→ R2 (SD)	R2→ R3 (SD)	R3→ R4 (SD)	R1→ R4 (SD)
Grit (1642)	0.43 (0.69)	0.34 (0.63)	0.31 (0.61)	1.07 (0.91)
Self-control (502)	0.24 (0.74)	0.24 (0.61)	0.30 (0.61)	0.78 (0.95)

The data above describe students' character growth relative only to other students

in this same study. The research question, however, asks how this growth compares to

students in other studies using similar measures. As noted in the previous chapter, there

are two published studies (Seider, 2012; West et al., 2014) that used and reported similar

grit measures for school-aged children—also in urban public charter schools serving predominantly African American and Latino children of low socio-economic status—but no studies that used *and* reported similar self-control measures.

Broadly speaking, both Seider (2012) and West et al. (2014) found that middle school students' perceptions of their grit declined from the beginning of the school year to the end. West et al. found a similar year-long decline for self-control, albeit with a different survey instrument than that used in this study. Consequently, the fact that the students in this study (n = 2144)—from their own and their teachers' perspectives—grew with respect to grit and self-control may be a notable finding unto itself.[2] Because Seider (2012) and West et al. (2014) reported similar findings, and because Seider (2012) reported his findings in a more comparable fashion, I will use his grit results as the more detailed point of reference here. Seider found that the middle school students across his sample reported beginning-of-school-year levels of perseverance just under 4.0 (along a 5-point Likert scale composed of similar items to this study's instrument) (p. 108).[3] By the end of the year, however, the students in Seider's study were averaging self-reported perseverance levels around 3.7 (p. 109). In short, while Seider's middle school sample began the year reporting themselves grittier (M = ~4.0) than the elementary school students in this study (M = 3.58), Seider's sample's perceptions of their grittiness declined through the year, whereas this study sample's perceptions increased. (And the

---

[2] That said, Seider (2012) and West et al. (2014) were examining such changes in middle schoolers. Middle school is a time of great developmental change, and thus the difference between this study's findings and those of Seider and West et al., may be more attributable to the age of the participants as opposed to other "real" differences in students' experiences and growth trajectories.

[3] I am reporting Seider's findings in somewhat imprecise terms because he displayed these descriptive statistics in figures (i.e., bar graphs) rather than tables. Therefore I am approximating the exact numbers from the data displayed in the figures, rather than the precise, two-decimal-point data that one might find in a table.

same comparative pattern holds with the West et al.'s [2014] findings regarding grit and self-control). By the end of the year, the study sample's perceptions of their grittiness (M = 4.10) surpassed Seider's sample's beginning-of-year highs. In other words, the two groups of students' trajectories were almost the exact opposite, with the middle schoolers in Seider's study starting near 4.0 and ending near 3.7, and the elementary schoolers in this study starting near 3.6 and ending just above 4.0.

To summarize, on average, the elementary school students in this study (n = 2144) grew with respect to both grit and self-control over the course of the school year, from both their own and their teachers' perspectives. Depending on the strength (grit or self-control) and the source (self- or teacher-report), the students, on average, began the year at different starting points and grew at different rates, but ended the school year (across perspectives and strengths) close to the top of the five-point Likert scale, perhaps indicating ceiling effects associated with the measures (as noted in the previous chapter). This stands in contrast to Seider's (2012) and West et al.'s (2014) findings that middle school students steadily declined on both these outcomes over the course of a school year.

I presented descriptive statistics for point-in-time character levels and then character growth (as calculated from students' and teachers' point-in-time survey ratings); some research suggests examining the relationship between those two types of measures. As noted in Chapter III, several studies (e.g., Ready, 2013; Pfost, Hattie, Dörfler, & Artelt, 2014) have shown that students' initial reading levels are associated with their year-long reading growth. Thus, I hypothesized that we might see similar relationships between students' initial character status and year-long character growth.

Indeed, Tables 9 and 10 illustrate strong negative relationships between students' initial perceptions of their strength of character and their year-long growth for both grit ($r = -0.72$, $p < 0.001$) and self-control ($r = -0.66$, $p < 0.001$). There is a similar strong

Table 9

Correlations Between Initial and Year-long Grit Growth

	Initial Grit (Student P)	Initial Grit (Teacher P)	Grit Growth (Student P)	Grit Growth (Teacher P)
Initial grit (Student P)	1.00			
Initial grit (Teacher P)	0.45***	1.00		
Grit growth (Student P)	-0.72***	-0.21***	1.00	
Grit growth (Teacher P)	-0.19***	-0.61***	0.30***	1.00

Note. P = perception
*$p < .05$. **$p < .01$. ***$p < .001$

Table 10

Correlations Between Initial and Year-long Self-control Growth

	Initial Self-control (Student P)	Initial Self-control (Teacher P)	Self-control Growth (Student P)	Self-control Growth (Teacher P)
Initial Self-control (Student P)	1.00			
Initial Self-control (Teacher P)	0.54***	1.00		
Self-control Growth (Student P)	-0.66***	-0.06	1.00	
Self-control Growth (Teacher P)	-0.37***	-0.67***	0.36***	1.00

Note. P = perception
*$p < .05$. **$p < .01$. ***$p < .001$

negative relationship between teachers' initial perceptions of students' strength of character and their year-long growth, for both grit ($r = -0.61$, $p < 0.001$) and self-control ($r = -0.67$, $p < 0.001$). In sum, students with higher initial character levels demonstrate less year-long character growth than their peers with lower initial character levels.

One finding that appears as an artifact of the correlational analysis in the tables above is the moderately strong relationship between students' and teachers' initial perceptions of students' character levels for both grit ($r = 0.45$, $p < 0.001$) and self-control ($r = 0.54$, $p < 0.001$). This finding prompted me to examine the relationship between students' and teachers' perceptions of students' strength of character at each point-in-time measure (i.e., R1, R2, R3, and R4). Though I had not considered this relationship explicitly in the research questions or the literature review, it struck me as salient to both. If CE were "working" in these classrooms, this relationship should be strong and strengthening through the school year. As displayed in Table 11, this relationship was strong even at the beginning of the school year and did strengthen over the course of the year for both grit (with the exception of a dip in R4) and self-control.

Table 11

Correlations Between Students' and Teachers' Perceptions of Students' Strength of Character

Survey Round	Grit	Self-control
R1	0.45***	0.54***
R2	0.52***	0.66***
R3	0.57***	0.72***
R4	0.53***	0.74***

Note. R1 = initial survey administration. R2 = second survey. R3 = third survey. R4 = final survey.
*$p < .05$. **$p < .01$. ***$p < .001$

The strong relationship between students' and teachers' perceptions at the beginning of the year may be attributed to strong initial shared conceptions of the strengths in question. That the relationship continues to strengthen over the course of the school year might be interpreted in a couple ways. One reading of this relationship is a convergence of teachers' and students' "vision" of the strength in question. Another interpretation, however, is that it is an artifact of the ceiling effects that resulted from using a five-point Likert scale to measure growth over a school year, leaving little room in the latter rounds of survey administration for a difference of perspective between teachers and students. I will take up the latter possibility in the concluding chapter.

Reading Growth

Research Question #2: How does reading growth among the students in the study sample compare to that of students in other studies using similar measures?

As noted in the previous chapter, teachers used normed, classroom-based reading assessments to assess students' reading status and growth in terms of reading grade-level equivalency (GLE). By definition, a child's expected year-long reading growth is one GLE. The GSE that served as the study site provided historical data indicating that their graduate students had seen, on average, higher-than-expected reading growth among their K-12 students, stemming in part from students starting the year, on average, with below-expected GLEs and growing to at or slightly-above-expected GLEs by the end of the year. I hypothesized that I would find similar outcomes in the study sample.

Indeed, on average, students in the study began the school year at or slightly below the expected GLE in reading. Table 12 shows that the kindergartners (n = 646) in

the study, for example, had a mean first-round (R1) reading GLE of -0.25 (SD = 0.34),

meaning that they started the year 0.25 GLEs below kindergarten reading level. Also as

expected, by the end of the school year (EOY),[4] all but the third graders (n = 280), on

average, were reading more than one GLE beyond baseline grade level. For example,

first graders (n = 591) were, on average, ending the year at a 2.25 GLE (SD = 0.65), or

1.25 GLEs beyond a baseline first-grade GLE. The descriptive statistics associated with

students' reading GLE by round are summarized by grade in the Table 12.

Table 12

Students' Average Reading Scores by Assessment Round and Grade

Grade (n)	Reading GLE by Round (SD)			
	R1	R2	R3	EOY
K (646)	-0.25	0.23	0.56	1.18
	(0.34)	(0.40)	(0.41)	(0.44)
1st (591)	1.01	1.36	1.69	2.25
	(0.56)	(0.56)	(0.59)	(0.65)
2nd (524)	2.00	2.22	2.53	3.09
	(0.68)	(0.72)	(0.74)	(0.71)
3rd (280)	2.73	2.99	3.24	3.82
	(0.73)	(0.75)	(0.76)	(0.74)
4th (103)	3.97	4.16	4.45	5.07
	(1.04)	(1.10)	(1.07)	(1.05)

As predicted, across the entire sample (n = 2144) and within each of the grades,

the students, on average, grew more than 1.00 GLE. For example, the kindergarteners (n

---

[4] Some classrooms tested students' reading growth four times from R1 to EOY, some five times, and some six times. In other words, for some classrooms EOY = R4, for others EOY = R5, and for others EOY = R6. For ease of calculation and comparison, I am only displaying R1, R2, R3, and EOY rounds because I am calculating reading growth from R1 to EOY, regardless of how many rounds of data (two, three, or four) fall between those two poles.

= 646) grew 1.43 GLEs (SD = 0.43), while the fourth graders (n = 103) grew 1.09 GLEs

(SD = 0.57). Generally speaking, the younger students grew more than the older

students. The descriptive statistics associated with students' R1, R2, R3, and year-long

growth are displayed in Table 13.

Table 13

Students' Average Round-to-round Reading Growth and Overall Growth by Grade

	Reading GLE Growth by Round (SD)			
Grade (n)	R1→ R2	R2→ R3	R3→ EOY[5]	R1→EOY
K (646)	0.47	0.33	0.62	1.43
	(0.35)	(0.26)	(0.32)	(0.43)
1st (591)	0.35	0.34	0.55	1.24
	(0.35)	(0.29)	(0.38)	(0.48)
2nd (524)	0.22	0.31	0.57	1.09
	(0.28)	(0.28)	(0.34)	(0.45)
3rd (280)	0.27	0.25	0.59	1.10
	(0.32)	(0.26)	(0.36)	(0.48)
4th (103)	0.19	0.29	0.61	1.09
	(0.33)	(0.36)	(0.47)	(0.57)
All (2144)	.33	0.32	0.59	1.24
	(0.34)	(0.28)	(0.36)	(0.48)

To summarize, students in the study sample (n = 2144), on average, demonstrated

more reading growth (M = 1.24, SD = 0.48) than expected (M = 1.00), with younger

students, on average, demonstrating even more notable growth than older students.

As with the negative relationship between initial character levels and year-long

character growth described above, we also see a negative relationship between initial

GLE in reading and year-long reading growth (r = -0.29, p < 0.001) in the study sample.

---

[5] Again, the R3 → EOY growth looks larger than the preceding two rounds because some students had three more rounds of assessment after R3, while others had two, and others had only one.

These findings suggest that students in the sample who begin the year with higher reading levels—relative to students in their same grade—grow less with respect to reading than their peers with lower initial reading levels. This finding represents the well-documented "fan-close" pattern of reading development, whereby initial reading status is negatively related to year-long reading gains (Ready, 2013), which I will discuss in greater detail in the concluding chapter of this study.

Academic and Character Growth

Research Question #3: What, if any, relationship exists between students' academic and character skills and growth?

As described in Chapter II, the literature relevant to this question had mixed findings, with some studies reporting a positive relationship between academics and character (e.g., Benninga et al., 2003) and others suggesting little relationship (e.g., US DOE, 2010). Because this study attempted to harness the potency of social-psychological interventions (e.g., Yeager & Walton, 2011) and a more "home-grown" approach to CE (e.g., Seider, 2012), I hypothesized that there would be a significant positive association between academics and character.

Turning first to the relationship between reading and character growth, my findings suggest a trivial positive correlation between students' reading growth and the character growth calculated based on students' perceptions ($r = 0.06$, $p < 0.01$), but no

relationship between students' reading growth and character growth as calculated based on teachers' perceptions. Table 14 presents these findings.[6]

Table 14

Correlations Between Standardized Reading Growth and Character Growth

	Rdg. Growth (Std. w/in Gr.)	Char. Growth (Student P)	Char. Growth (Teacher P)
Reading growth (Standardized)	1.00		
Character growth (Student P)	0.06**	1.00	
Character growth (Teacher P)	0.01	0.32***	1.00

Note. P = perception
*$p < .05$. **$p < .01$. ***$p < .001$

When examining this relationship by character strength, we see that the relationship between students' reading growth and character growth as calculated based on students' perceptions is driven by students in classrooms focused on self-control (n = 502). Whereas there is a weak correlation between students' reading growth and their self-control growth as calculated based on their perceptions (r = 0.16, p < 0.001), there is no relationship between students' reading growth and their own grit growth as calculated based on their perceptions. These findings are displayed in Tables 15 and 16. In sum, students who grew more in reading also grew more with respect to self-control (as calculated based on their perceptions). No such relationship exists for students focused on developing grit.

---

[6] Notably, this is a moderately strong correlation between students' and teachers' perceptions of students' character growth (r = 0.32, p < 0.001). While this relationship was not the focus of the study's research findings, I will return to it later in the chapter as a happenstantial finding of interest.

Table 15

Correlations Between Standardized Reading Growth and Self-control Growth

	Rdg. Growth (Std. w/in Gr.)	Self-c. Growth (Student P)	Self-c. Growth (Teacher P)
Reading growth (Standardized w/in grade)	1.00		
Self-cont. growth (Student perception)	0.16***	1.00	
Self-cont. growth (Teacher perception)	0.07	0.36***	1.00

Note. P = perception
*p < .05. **p < .01. ***p < .001

Table 16

Correlations Between Standardized Reading Growth and Grit Growth

	Rdg. Growth (Std. w/in Gr.)	Grit Growth (Student P)	Grit Growth (Teacher P)
Reading growth (Standardized w/in grade)	1.00		
Grit growth (Student perception)	0.03	1.00	
Grit growth (Teacher perception)	-0.01	0.30***	1.00

Note. P = perception
*p < .05. **p < .01. ***p < .001

The above results detail the relationship between students' reading and character *growth*, as that was the focus of the main research question. In explicating this question in Chapter I, though, I posed a related sub-question: Do students' initial character *levels* have any relationship with their reading gain (and vice versa)? As noted in Chapters I and II, some studies have demonstrated a relationship between point-in-time measures of

grit or self-control and improved academic outcomes (e.g., Duckworth et al., 2007;

Duckworth & Seligman, 2005, respectively). Thus, I next turn to examining the

relationships between point-in-time measures of character and reading, against growth on

those same outcomes.

When examining initial measures of grit against reading growth, there is a weak

positive correlation between teachers' initial perceptions of students' grit and students'

year-long reading growth ($r = 0.16$, $p < 0.001$). There is no relationship, however,

between students' initial perceptions of their grit and their year-long reading growth.

When looking past the initial grit measures into students' and teachers' perceptions of

students' grit in the second (R2), third (R3) and fourth (R4) quarters, we see additional

significant relationships. For example, in the second half of the school year, there is also

a weak positive correlation between students' point-in-time perceptions of their grit and

their year-long reading growth ($0.08 \leq r \leq 0.10$, $p < 0.001$). In sum, greater year-long

reading growth is generally associated with higher point-in-time measures of grit. These

findings are displayed in Table 17.

Table 17

Correlations Between Students' and Teachers' Round-by-round Perceptions of Students'
Grit and Students' Standardized Year-long Reading Growth

	R1	R2	R3	R4
Grit— Student P	0.03	0.04	0.10***	0.08***
Grit— Teacher P	0.16***	0.20***	0.20***	0.18***

Note. P = perception
*p < .05. **p < .01. ***p < .001

Conversely, when examining initial measures of self-control, there is a modest negative correlation between students' initial perceptions of their self-control and their year-long reading growth ($r = -0.20$, $p < 0.001$). This relationship between students' perceptions of their self-control and their year-long reading growth weakens by R2 ($r = -0.11$, $p < 0.05$) and then disappears entirely. There is no relationship between any point-in-time measure of teachers' perceptions of students' self-control and students' year-long reading growth. In short, greater year-long reading growth is associated with lower student-reported self-control ratings during the first half of the school year. These findings are displayed in Table 18.

Table 18

Correlations Between Students' and Teachers' Round-by-round Perceptions of Students' Self-control and Students' Standardized Year-long Reading Growth

	R1	R2	R3	R4
Self-control—Student P	-0.20***	-0.11*	-0.05	-0.05
Self-control—Teacher P	-0.01	-0.03	0.02	0.06

Note. P = perception
*$p < .05$. **$p < .01$. ***$p < .001$

When turning to examine point-in-time reading measures and their relationship with character growth, we see that students' reading levels have a significant and negative association with students' year-long grit growth (from both their own and their teachers' perspectives) at every point-in-time measure of reading GLE (as displayed in Table 19). In other words, greater year-long grit growth is associated with lower point-in-time measures of reading, throughout the school year and regardless of perspective (student- vs. teacher-report).

Table 19

Correlations Between Students' Round-by-round Standardized Reading Measures and Students' Year-long Grit Growth

	Reading R1	Reading R2	Reading R3	Reading Final
Grit growth— Student P	-0.10***	-0.06*	-0.05*	-0.06*
Grit growth— Teacher P	-0.15***	-0.15***	-0.15***	-0.13***

Note. P = perception
*p < .05. **p < .01. ***p < .001

While point-in-time reading measures are always associated with year-long grit growth, when examining self-control, we find that only final reading levels are associated with year-long self-control growth, and only from teachers' perspectives. In other words, greater year-long self-control growth (from the teachers' perspective) is only associated with higher final reading levels. These findings are displayed in Table 20.

Table 20

Correlations Between Students' Round-by-round Standardized Reading Measures and Students' Year-long Self-control Growth

	Reading R1	Reading R2	Reading R3	Reading EOY
SC growth— Student P	-0.03	0.08	0.09	0.12**
SC growth— Teacher P	-0.09	-0.04	-0.03	-0.03

Note. P = perception
*p < .05. **p < .01. ***p < .001

To recap, the research question asked, "What, if any, relationship exists between students' academic and character growth?" Because character growth was measured from both students' and teachers' perspectives, I examined this question through those

two dimensions and found that students who grow more in reading also see themselves as growing more with respect to self-control. There were no relationships between students' reading growth and teachers' perceptions of students' self-control growth. Similarly, there were no relationships between students' reading growth and grit growth, from either students' or teachers' perspectives.

Moving from growth-only to growth and point-in-time measures, the analysis yielded the following findings:

- Greater year-long reading growth is generally associated with higher point-in-time measures of grit.

- Greater year-long reading growth is associated with lower initial student self-ratings of self-control.

- Greater year-long grit growth is associated with lower point-in-time measures of reading.

- Greater year-long self-control growth (from the students' perspectives) is associated with higher final reading levels.

Character Education—Lessons and Beyond

Research Question #4: What aspects of CE curriculum, planning, pedagogy, and teachers' choices might be associated with students' academic and character growth?

Descriptive Statistics

The teachers' performance in the CE lesson that served as the "intervention" and predictor variable in this study is described Table 21. Faculty members at the GSE that served as the study site scored each teacher's lesson on a rubric with a 5-point scale (from

4/"exemplary" to 0/"lacking"). The rubric had six criterion-based rows, each describing

a particular component of a strong CE lesson, as well as a final seventh row, representing

the faculty member's overall evaluation of the lesson. I created two additional measures:

(1) an "average" measure, which was the mean of the five teaching-based rows (i.e.,

objective, definition, rationale, model and practice) and (2) a "total" measure, which was

the sum of all six criterion-based rows. The data describing teachers' (n = 88)

performance are shown in Table 21.

Table 21

Teachers' Performance on the CE Lesson Rubric

Rubric Row	Avg. Score (SD)
Objective	2.87 (0.84)
Definition	3.26 (0.73)
Rationale	2.93 (0.89)
Model	3.12 (0.58)
Practice	3.14 (0.65)
Reflection	3.26 (0.63)
Overall	3.17 (0.55)
Average	3.06 (0.48)
Total	18.7 (2.71)

As point of comparison, a subset of the teachers' (n = 55/88) performance in the academic lesson observed a month or two later displayed the values reported in Table 22.[7]

Table 22

Teachers' Performance on the Academic Lesson Rubric

Rubric Row	Avg. Score (SD)
Classroom culture	3.29 (0.59)
Teaching cycle	3.16 (0.50)
Content	3.05 (0.61)
Self and other people	3.10 (0.35)
Overall	3.16 (0.44)
Average	3.15 (0.38)
Total	12.6 (1.52)

In sum, in the CE and academic lessons, teachers performed, on average, slightly above the GSE's stated performance expectations for their graduate students (i.e., an overall score of 3.00—"proficient"). They performed somewhat better, on average, in the academic lesson (M = 3.15, SD = 0.38) than on the teaching criteria in the CE lesson (M = 3.06, SD = 0.48). Both sets of outcomes, however, surpassed the GES's stated expectation.

---

[7] The full quantitative sample represents three cohorts of teachers (n = 88): the Classes of 2014, 2015, and 2016 at the GSE that served as the study site. All of these teachers taught the CE lesson and were evaluated with the same rubric. The Class of 2014, however, was evaluated differently for the academic lesson. Thus, the academic lesson data only exist for the Classes of 2015 and 2016 (n = 55).

Additionally, the relationship between the two measures of teaching, i.e., CE lesson performance and academic lesson performance, was also significant. Whether we examine the correlation between the constructed "average" measures (r = 0.11, p < 0.001), the constructed "total" measures (r = 0.15, p < 0.001), or the given "overall" measures (r = 0.19, p < 0.001), the relationship between teacher's academic-focused instruction and their character-focused instruction was positive and significant, though only weak.

Correlational Analyses

Teachers' CE lesson performance (a z-scored measure [M = 0, SD = 1] of their "total" score on the rubric) was negatively but trivially related to students' perception of their year-long character growth ($r = -0.07$, $p < 0.01$). CE lesson performance was unrelated to teachers' perceptions of students' year-long character growth ($r = -0.02$, $p < 0.05$) (and as reported above, students' and teachers' perceptions of students' year-long character growth were moderately related [$r = 0.32$, $p < 0.001$]). These findings are displayed in Table 23.

Table 23

Correlations Between CE Lesson Performance and Year-long Character Growth (CG)

	CE Lesson	CG--Student Perception	CG--Teacher Perception
CE lesson	1.00		
Year-long CG--student perception	-0.07**	1.00	
Year-long CG--teacher perception	-0.02	0.32***	1.00

*p < .05. **p < .01. ***p < .001

When examining the same relationships between grit and self-control, respectively, and character growth, we see that the trivial relationship displayed above is driven entirely by the classrooms focused on self-control. In short, there is no relationship between teachers' CE lesson performance and students' year-long grit growth, as displayed in Table 24 below. There is, however, a weak negative relationship between teachers' CE lesson performance and students' year-long self-control growth, from both students' ($r = -0.21$, $p < 0.001$) and teachers' ($r = -0.18$, $r < 0.001$) perspectives, as displayed in Table 24.

Table 24

Correlations Between CE Lesson Performance and Year-long Character Growth (CG)

	Grit CE Lesson	Self-control CE Lesson
Year-long CG--student perception	-0.04	-0.21***
Year-long CG--teacher perception	0.01	-0.18***

*$p < .05$. **$p < .01$. ***$p < .001$

Similar patterns held when examining the relationship between teachers' CE lesson performance and character growth as calculated from students' and teachers' reports during the interval in which the CE lesson took place. In other words, teachers' CE lesson performance was generally negatively, but still only trivially or weakly associated with character growth from before the CE lesson "intervention" (R1) to the measurement round just after the CE lesson "intervention" (R2).[8] These findings are displayed in Tables 25 and 26.

---

[8] The actual time between the first round of character measurement (i.e., R1), the CE lesson "intervention," and the second round of character measurement (i.e., R2) differed somewhat from classroom to classroom and cohort to cohort. Generally, R1 occurred sometime in October; the lesson occurred in November, and R2 occurred in January.

Table 25

Correlations Between CE Lesson Performance and Proximate Character Growth (CG)

	Grit CE Lesson	Self-control CE Lesson
R1 → R2 CG--student perception	-0.10***	-0.23***
R1 → R2 CG--teacher perception	-0.04	-0.09*

*p < .05. **p < .01. ***p < .001

In sum, teachers' CE lesson performance was generally negatively, though only weakly, associated with students' character growth, both proximate to the lesson and over the school year.

For comparison, when turning to the relationship between teachers' performance on the academic lesson rubric and students' reading growth, we see a significant, positive relationship. This positive year-long relationship is weak (r = 0.11, p < 0.001), however, and oddly does not hold when examining proximate reading growth. These findings are displayed in Table 26.

Table 26

Correlations Between Academic Lesson Performance and Standardized Proximate and Year-long Reading Growth (RG)

	Academic Lesson
Academic lesson	1.00
R1 → R2 RG	0.05
Year-Long RG	0.11***

*p < .05. **p < .01. ***p < .001

Regression Analyses

The final portion of the quantitative analyses entailed linear regression models to further examine the relationship between the CE lesson and students' character growth. The original intent of the regression analyses was to examine the relationship between the measured teaching components of the CE lessons from the rubrics (i.e., objectives, definition, rationale, modeling, and practice) to determine which (if any) of these components were driving the hypothesized positive relationship between the lesson and students' character growth. Because the lesson was (generally) negatively associated with students' character growth, however, the regression analyses served to examine which components of the lesson were associated with this overall negative relationship and which ran counter to that trend.

In building the models, there were several dimensions to consider, namely, the differences noted in prior analyses between (1) students' and teachers' perceptions of students' character growth, (2) grit growth and self-control growth, and (3) proximate versus year-long growth. Those dimensions alone call for eight different iterations of the model, displayed in Table 27.

Table 27

Regression Models

Model #	Source	Strength	Duration
1	Student	Grit	Proximate
2	Student	Self-control	Proximate
3	Student	Grit	Year-long
4	Student	Self-control	Year-long
5	Teacher	Grit	Proximate
6	Teacher	Self-control	Proximate
7	Teacher	Grit	Year-long
8	Teacher	Self-control	Year-long

Additionally, given the modest-to-strong relationships between students' and teachers' initial perceptions of students' strength of character and their year-long character growth, I included the initial character levels as covariates in the models. The trends are described below and the significant findings are displayed in Table 28.

The most robust model (Model 3) accounts for 50% of the variance in the outcome (year-long self-control growth from the students' perspective). Most of the explanatory power in this model and the others, however, comes not from the CE lesson components, but rather from the initial character measures. In pilot analyses, the CE lesson component measures—unadjusted for initial character status—explained no more than 7% of the variance in character growth. In short, regardless of the model or the controls, the CE lesson components were only inconsistently and weakly associated with the character growth measures. The trends across the adjusted models are presented in Table 28.

While no single CE lesson component was associated with students' character growth (or lack thereof) across all eight models, several components displayed similar patterns across the models. For example, in Models 2 and 3, the definition of the

character strength presented in the lesson was negatively associated with students' perception of their proximate self-control growth (ES = -0.14; $p < 0.01$) and students' perception of their year-long grit growth (ES = -0.06; $p < 0.05$). Conversely, in Models 1, 4, and 6, the rationale for building the strength presented in the lesson was positively, though again only trivially, associated with students' character growth. Finally, in Models 1, 5, and 8, the opportunities for students to practice within the lessons were negatively associated with students' character growth. These ranged from a trivial association in the proximate grit growth models (ES = -0.06, $p < 0.05$) to a weak association in the teachers' perceptions of students' year-long self-control growth model (ES = -0.22, $p < 0.001$). In sum, the definitions and practice opportunities in the CE lessons were repeatedly negatively associated with students' character growth, while the rationales presented in the lessons were often positively associated with students' character growth.

In sum, students' initial character status explained far more of their character growth than did either the CE lessons broadly or the specific CE lesson components. However, adjusting for initial character status, the definitions, rationale, and practice opportunities in the lessons demonstrated patterns in the models where they were significant predictors. More specifically, stronger definitions and practice opportunities were associated with decreased character growth, while stronger rationales were associated with increased character growth.

Table 28

Predictors of Students' Proximate and Year-long Character Growth

Model #:	1	2	3	4	5	6	7	8
Perspective:	Student	Student	Student	Student	Teacher	Teacher	Teacher	Teacher
Strength:	Grit	Self-Control	Grit	Self-Control	Grit	Self-Control	Grit	Self-Control
Growth Int:	Proximate	Proximate	Year	Year	Proximate	Proximate	Year	Year
Objective		-0.09*		0.09*	-0.05*	-0.10*	-0.05*	
Definition		-0.14**	-0.06*					
Rationale	0.08**			0.07**		0.09**		
Model	-0.06**						0.07**	
Practice	-0.06*				-0.06*			-0.22***
Initial S	-0.52***	-0.45***	-0.72***	-0.62***				
Initial T					-0.38***	-0.48***	-0.61***	-0.64***
Constant		-0.12**						
R2	0.28***	0.26***	0.50***	0.46***	0.15***	0.25***	0.36***	0.48***

Note. Int = Interval; Proximate = R1 → R2; Year = R1 → R4
*p < .05. **p < .01. ***p < .001

Summary

I will summarize the above findings by research question and then turn to the implications for the qualitative portion of the study.

Research Question #1: How do changes in grit or self-control among the students in the study sample compare to those in students in other studies using similar measures?

Students in the study sample demonstrated notable year-long growth in grit and self-control, from both their own and their teachers' perspectives. This stands in contrast to the findings of Seider (2012) and West et al. (2014), where middle school students (of comparable race, socio-economic status, and school setting to the students in this study) reported a year-long decline in both grit and self-control. Additionally, across character strengths and classrooms, teachers saw more character growth—as calculated by their teacher reports—in students than students saw in themselves (as calculated by their self-reports). While teachers' reports indicated more growth than students' reports, students' and teachers' perspectives on students' *point-in-time* character status were strongly and positively related, and this relationship generally strengthened with each subsequent round of measurement, though this strengthening may have been in part an artifact of ceiling effects. Finally, both students' and teachers' reports indicated more character growth in classrooms focused on grit, as compared with those focused on self-control. As with initial reading status and reading growth (Ready, 2013), students with higher initial character status demonstrated less growth, on average, than their peers with lower

initial character status. This was true from both students' and teachers' perspectives, though again, this may have been in part an artifact of ceiling effects.

Both the larger amount of character growth demonstrated by students focused on grit (as compared with those focused on self-control) and the strong relationship between students' and teachers' perspectives on students' point-in-time character levels were unanticipated, interesting findings, though again, the latter may have been an artifact of ceiling effects. Consequently, I adjusted my qualitative focus to further explore and explain these findings. I will detail both my hypotheses related to these unexpected findings and my methods for investigating them in the next chapter.

Research Question #2: How does reading growth among the students in the study sample compare to that of students in other studies using similar measures?

As expected, students in the study sample began the school year slightly below their expected grade-level equivalency (GLE) in reading, on average growing over one GLE to end the school year above the initial expected GLE for the subsequent grade. As predicted, students demonstrated, on average, a "fan-close" pattern of reading growth (Ready, 2013), with lower initial reading status associated with higher year-long growth.

Research Question #3: What, if any, relationship exists between students' academic and character skills and growth?

In contrast to what I theorized in my conceptual framework, there was only a weak positive relationship between reading and character growth, and this relationship existed only from the students' perspective in the classrooms focused on self-control.

Looking at reading growth through the lens of point-in-time character status yielded the following findings:

- Greater year-long reading growth was generally associated with higher point-in-time measures of grit.

- Greater year-long reading growth was associated with lower initial self-ratings of self-control.

These two findings, while unexpected, were consistent with the first. There was not a relationship between reading growth and grit growth, but there was a relationship between reading growth and point-in-time grit status. In other words, higher levels of grit are associated with greater reading growth, but reading and grit are not growing in relation to one another. Conversely, reading growth and self-control are positively (if weakly) related from students' perspectives, while reading growth was negatively (if weakly) associated with students' initial self-control ratings.

Reversing the lens and viewing character growth through point-in-time reading achievement yielded the following findings:

- Greater year-long grit growth was associated with lower point-in-time measures of reading.

- Greater year-long self-control growth (from the students' perspectives) was associated with higher final reading levels.

These findings are the expected inverse of the pair above: reverse the lens, reverse the findings statements. In the tersest terms, more reading growth, higher grit status; more grit growth, lower reading status. More reading growth, lower initial self-control status; more self-control growth, higher final reading status.

While one can explain the findings in logical relation to one another, explaining them in the larger context of the study again seemed ripe for qualitative investigation, which I will detail in the subsequent chapter.

Research Question #4: What aspects of CE curriculum, planning, pedagogy, and teachers' choices might be associated with students' academic and character growth?

Teachers demonstrated above-expected performance in the CE lesson (and a comparably timed academic lesson). I hypothesized that teachers' performance in the CE lesson would have a positive relationship with students' proximate and year-long character growth. I found, however, that CE lesson quality was negatively (though weakly) associated with both proximate character growth (for grit and self-control) and year-long character growth (for self-control). This stood in contrast to both my conceptual framework *and* the comparative findings regarding the positive relationship between the academic lessons and year-long reading growth. In other words, I found that the CE lessons *were* generally associated with students' character growth, just not in the direction I had hypothesized. Consequently, I needed to better understand what was happening in the lessons that might explain this surprising finding.

Regression models designed to examine the particular CE lesson components and students' character growth yielded inconsistent and weak associations only, with the exception of students' initial character status, which was a significant covariate in all models. Across several models, the lesson definitions, rationales, and practice opportunities maintained the valence of their associations with students' character growth. Specifically, stronger definitions and practice opportunities were associated with

decreased character growth, while stronger rationales were associated with increased character growth.

If these components, i.e., definitions and practice, were repeatedly negatively associated with students' proximate and year-long character growth, what exactly were teachers doing in these parts of their lessons such that students sometimes grew less because of it? Additionally, given that the rationales were often weakly positively associated with students' character growth, what was happening in these parts of the lessons that was modestly boosting some students' growth?

To summarize, I had unexpected findings across three of my four research questions. Several of these findings led me to hypothesize that students and/or teachers were conceiving of and presenting grit and self-control quite differently, such that the two strengths had different growth trajectories and different relationships with reading growth and achievement. Some of these findings led me to posit that CE opportunities beyond the single lesson of interest were at work in teachers' classrooms and schools, such that students' and teachers' perspectives on students' character status would be so strongly correlated throughout the year. Finally, my major unexpected finding—that the CE lessons had a suppressive (if weak) relationship with students' character growth—led me to posit that the lessons re-positioned the strengths for students and teachers, such that their subsequent estimations of grit or self-control were adjusted relative to the vision of that strength presented in the lesson.

V—QUALITATIVE FINDINGS AND ANALYSIS

The original intent of the qualitative portion of the study was to add depth and nuance to quantitative findings, in particular those associated with my fourth research question: *What aspects of CE curriculum, planning, pedagogy, and teachers' choices might be associated with students' academic and character growth?* However, as described in the summary of the preceding chapter, there were several unexpected findings across three of my four research questions that compelled further exploration and explanation via qualitative methods. Consequently, my first, third, and fourth research questions became fodder for both quantitative *and* qualitative examination.[1] In this chapter, I first outline the connection between the relevant quantitative findings and my qualitative approach, building off of the summary in the previous chapter. I will then describe the qualitative findings, putting them in the context of the related quantitative findings whenever possible.

Quantitative Findings and Qualitative Approach

The first research question—*How do changes in grit or self-control among the students in the study sample compare to those in students in other studies using similar measures?*—was intended to be purely quantitative, but as detailed in the previous

---

[1] My second research question focused on reading growth. The findings associated with that question were largely as hypothesized, and moreover, reading growth was the secondary outcome of interest in the study. Consequently, I did not investigate reading status or growth alone in the qualitative portion of the study.

chapter, two quantitative findings seemed ripe for further qualitative analysis: (1) the larger amount of character growth demonstrated by students focused on grit as compared with those focused on self-control and (2) the strong relationship between students' and teachers' perspectives on students' point-in-time character levels. With respect to the first finding, I hypothesized that one would see self-control defined, modeled, and/or practiced such that students could find it "harder" to demonstrate than grit. With respect to the second finding, I hypothesized that a strong, shared vision of grit or self-control, such that students' and teachers' perceptions were strongly related throughout the year, was likely reinforced by classroom and school-based CE opportunities beyond this single lesson, even accounting for the possible implications of ceiling effects.

Viewing the CE lesson videos would shed light on both findings. For example, with respect to the differential growth finding, were there elements of the CE lessons that might lead students to think of self-control as "harder" to demonstrate than grit? With respect to the finding regarding the strong relationship between teachers' and students' point-in-time perceptions, were these strong correlations grounded in similar teacher and student definitions and descriptions of the strengths, as demonstrated in the CE lessons? The interviews, too, would further elucidate this latter finding: were students' and teachers' shared conceptions of grit and self-control being reinforced and strengthened through classroom and school-wide CE beyond this single lesson?

The third research question—*What, if any, relationship exists between students' academic and character growth?*—was also intended to be purely quantitative, but two findings here compelled further qualitative examination. Most prominently, I was curious to better understand what might underlie the positive correlation between point-

in-time perceptions of grit and reading growth and, relatedly, the inverse relationship between students' point-in-time perceptions of their self-control and their reading growth. I hypothesized that the positive relationship between grit and reading growth was related to teachers and students connecting grit and academics in the CE lessons (and possibly at other times). Moreover, I posited that the negative relationship between self-control and reading growth was related to teachers' and students' focus on interpersonal (as opposed to academic) forms of self-control (Tsukayama et al., 2013). Both hypotheses proved investigable through the CE lesson videos. Were teachers' and students' definitions and descriptions of grit connected to academics in general, or reading in particular, and if so, how? Were teachers' and students' definitions and descriptions of self-control *not* (or less) connected to academics? If so, how *were* teachers and students envisioning self-control? Furthermore, if point-in-time perceptions of character were related—even if in different directions for grit and self-control—to students' reading growth, were teachers articulating those connections in the lessons and their interviews, and, if so, how? Additionally, I posited that teachers would describe, both in their lessons and their interviews, the broad connections between character and academics differently, depending on whether they were focused on grit or self-control in their lessons. To investigate these hypotheses, I focused on teachers' articulations of the relationship between strength of character and academic outcomes in both the lessons and their interviews.

Finally, the fourth research question—*What aspects of CE curriculum, planning, pedagogy, and teachers' choices might be associated with students' academic and character growth?*—was intended for qualitative examination. First, given that teachers'

performance in the CE lessons was above the stated expectation of the MAT program, did teachers also see themselves as efficacious in teaching character? I hypothesized that teachers felt, albeit in retrospect, a strong sense of CE teaching efficacy. To examine this question, I utilized both the pre-interview survey (designed to focus on teachers' sense of CE efficacy) and teachers' reflections on their instruction in their interviews.

I was most interested, though, in why these CE lessons, broadly speaking, were negatively (if weakly) related to students' character growth. What was happening (or not) in these lessons such that they were negatively associated students' notable character growth? Were teachers using the techniques they had read about, observed, and practiced in their coursework? Again, the CE lesson videos would reveal what was happening in the quantitative black box.

Moreover, given the negative association between the definitions and practice opportunities in the lessons and students' character growth, I hypothesized that either (a) these parts of the lessons were rife with unchecked inaccuracies and/or teaching foibles, such that they were associated with the opposite of what I had theorized *or* (b) the lessons were relatively strong, giving students a higher bar for grit or self-control, and thereby deflating students' (and in some cases, teachers') sense of their own character status. If one of those hypotheses might explain the negative relationship between definitions/practice and character growth, how can we explain the *positive* relationship between the rationales in the lessons and students character growth? I posited that this trend of positive association between rationales and students' character growth might suggest that teachers' rationales were particularly compelling, such that students were striving (and succeeding) to demonstrate character growth. To investigate these

questions, I relied heavily on my video-coding protocol, analyzing the trends associated with research-based practices, the defining of strengths, the framing of rationales, and practice opportunities within the lessons. As secondary evidence, I examined teachers' conceptions of the impact of these lessons as we discussed them in the interviews.

In sum, quantitative findings associated with the first, third, and fourth research questions catalyzed further qualitative investigation of qualitative data composed from videos, surveys, and interviews. The salient quantitative findings and associated qualitative questions and sources are summarized in Table 29.

Because the questions and investigative directions described above emerged from the quantitative analysis (described in the preceding chapter), I was able to select a qualitative sample that would maximize my opportunity to investigate accordingly, as described in Chapter III. Also as previously described, I was able to focus on the most relevant components of my CE lesson video review protocol and my interview protocol to explore these emergent questions in particular.

Because the CE lesson video review preceded the survey and interview (for both the researcher and the participants) and because the CE lesson videos represent a more comprehensive data set, I will begin by describing the CE lesson video findings and then turn to the surveys and interviews. Whenever meaningful, I will connect the findings across these methodologies.

Table 29

Quantitative Findings and Associated Qualitative Questions and Sources

RQ	Quantitative Finding	Qualitative Question	Qualitative Source
1	a. More character growth for students focused on grit than those focused on self-control b. Strong correlation between students' and teachers' point-in-time perceptions of students' strength of character	a. Differences in how teachers and students are presenting/conceiving of grit vs. self-control? Is self-control "harder"? b. i. Do students and teachers present/share similar definitions? ii. Are there other school-based opportunities to build shared visions?	a. Videos b. i. Videos ii. Interviews
3	Positive correlation between point-in-time perceptions of grit and reading growth (and negative relationship between students' point-in-time perceptions of their self-control and reading growth)	a. Were teachers and students defining/ conceiving of grit as more academic than self-control? b. How were teachers articulating the connection between academics and character?	a. Videos b. Interviews
4	a. Teachers performed above expectation in their CE lessons b. CE lessons were only weakly negatively related to students' character growth c. Definitions and practice, in particular, within the CE lessons were generally negatively related to character growth d. Rationales within the CE lessons were generally positively related to character growth.	a. Did the teachers see/describe themselves as effective teachers of CE? b. Were teachers using research-based CE strategies? To what effect? c. How were teachers and students defining grit and self-control? What did the practice opportunities look like? To what effect? d. How were teachers and students articulating the rationale for building these strengths? To what effect?	a. Pre-interview surveys and interviews b. Videos and interviews c. Videos and interviews d. Videos and interviews

Character Education Lesson Videos

As noted in Chapter III, in the late fall of their second year in the MAT program at the GSE that served as the study site, the teachers took coursework entitled "Teaching Character Strengths," introducing research-based approaches to CE and social-psychological interventions. (See Appendix I for my context and details of the related GSE coursework.) I will first describe the related qualitative findings via a component-by-component examination of the videos of the character lessons of each of the 14 teachers in the sample. I will then conclude this section with some more holistic video-based findings before turning to the surveys and interviews in the next section.

Definitions. Generally within the first few minutes of every lesson, the teacher, or occasionally a student or two, provided the class with one or several definitions of grit or self-control.[2] How teachers and students were defining the strengths seemed salient to several of the quantitative findings detailed above. I will first present the qualitative findings regarding the definitions presented in the CE lessons, and then connect them to my hypotheses.

Typically, even when students were the first to proffer the class a definition, the teacher later revealed a definition (or several) that became the standard for the rest of the lesson. Turning first to the grit definitions, they were found to be (a) relatively similar, classroom to classroom and (b) largely true to Duckworth et al.'s (2007) definition: "perseverance and passion for long-term goals" (p. 1087). Given the relatively young age of the students, most teachers made this concept more "student friendly." Continuing

---

[2] Many of the 14 lessons began with the teachers and/or the students reading aloud an objective for the lesson. The lesson objectives, however, were not an in-depth focus of the study because they were (a) only associated with weak and inconsistent findings in the regression analyses and (b) did not present further fodder upon qualitative examination. (See Appendix O for sample CE lesson objectives.)

to try and never giving up, for example, were especially popular re-conceptions of the original definition for an audience of younger students. In general, the grit definitions were broad and applicable to many situations. The teachers' and, when applicable, students' definitions of grit from each lesson are displayed in Table 30.

Table 30

Definitions of Grit

Teacher[3]	Grade	Teacher Definition(s)	Student Definition(s)
Ib	4	"A combination of persistence and resilience."	
Id	K	"Don't give up, even when times are tough."	"It's when you never ever ever give up."
Ie	K	"Try try again"	
If	K	"You can't give up, you have to keep trying." "Passion and perseverance for big goals." "Try and try."	
Ig	2	"Never stop learning."	"Trying until you achieve." "Don't give up."
IIa	2	"Push yourself, reach for goals, and do brave things."	"Something hard and you push yourself to do it." "Even if the work is hard, you keep trying."
IIIa	3	"Persistence and passion for long-term goals." "You keep on going when the going gets tough."	"Never give up; always try your best."
IIIb	1	"You finish what you start, even when it's tough."	

---

[3] Within the qualitative portion of the study, I refer to teachers in a de-identified fashion, using a code that allows the reader to determine each teacher's group status. In short, I refer to each teacher by their group number and alphabetical order within group. For example, the teacher in Group I whose last name is first (in that group) alphabetically is Teacher Ia, while the teacher in Group II whose last name is third (in that group) alphabetically is Teacher IIc.

As with the grit definitions, the self-control definitions were relatively similar from classroom to classroom. Rather than the broad and generalizable definitions of grit, though, teachers and students presented more particular definitions of self-control, often focusing on intra- or inter-personal self-control (like controlling one's feelings or waiting patiently), as opposed to more academic forms of self-control (like staying focused on the lesson instead of daydreaming). The teachers' and, when applicable, students' definitions of grit from each lesson are displayed in Table 31.

Table 31

Definitions of Self-control

Teacher	Grade	Teacher Definition(s)	Student Definition(s)
Ia	1	"Control over your body and feelings." "Ability to wait for something you really want."	
Ic	3		"Control your anger." "Control yourself." "Calm down." "Keeping our anger and frustration in check."
IIb	0	"The act of controlling your own behavior."	
IIc	0	"We use safe bodies, wait our turn to talk and walk in line, always stay in our seats, and worry about ourselves."[4]	
IId	0	"Stay calm even when excited or frustrated."	
IIIc	3	"The ability to control your feelings and behavior, especially when it's difficult."	

---

[4] These were the salient lyrics to a class "self-control chant," as opposed to a formal "definition" of self-control. The teacher and students performed this chant in call-and-response style at the beginning of the lesson.

With respect to these definitions' relationship to the quantitative findings and my related hypotheses, the connections are somewhat mixed. The fact that students, on average, demonstrated more growth with respect to grit than self-control seems perplexing given the definitions above. The grit definitions were broad and manifold. To "finish what you start, even when it's tough" and "never ever ever give up" seem daunting and applicable to every challenging task. The conceptions of self-control proffered in these lessons, though, i.e., to "wait for something you really want" or "calm down," seem particular to specific and therefore perhaps more manageable situations. While this differentiation may be only a matter of perspective, it is difficult to see the conceptions of self-control presented in these definitions as more difficult to attain than the conceptions of grit, as I had hypothesized. Consequently, the finding that students, on average, grew more with respect to grit than self-control seems unexplained by the differences in definitions of the strengths offered in these lessons.

It is possible to see in these definitions, though, the divergence between grit and self-control that can be seen in the quantitative finding associated with the contrasting relationships between these strengths and reading growth. Whereas perceptions of grit were positively related to reading growth, students' perceptions of their self-control were negatively related thereto. I had hypothesized that teachers and students would present definitions of grit more closely related to academics, and most of the definitions of grit apply to academic challenge. In some cases, teachers and students are explicit about the relationship between grit and academics. For example, Teacher Ig defines grit as "never stop learning," and a student in Teacher IIa's classroom says, "Even if the work is hard, you keep trying," with reference to schoolwork and grit. For these young students, still

very much in the learning-to-read stage of reading development, this connection could be related to what can be seen in the quantitative results: more grit is associated with more reading growth.

Turning next to the self-control definitions, none are related to academic endeavors directly, again, as I had hypothesized. In some cases, the definitions include school-based factors, as when Teacher IIc says, "…we wait our turn to talk and walk in line, [and] always stay in our seats…," thereby defining self-control in relation to some classroom and school behavior norms. But no teacher or student mentions academic work explicitly. While this does not explain the negative relationship between students' perceptions of self-control and their reading growth, it does perhaps partially explain the difference between this relationship and that of grit and reading growth.

Both the negative (though weak) relationship between the definitions and students' character growth *and* the negative (though weak) relationship between the character lessons more broadly and students' character growth seem difficult to explain based on these findings. I turn next to the models, i.e., examples and enactments that teachers and students presented during the lessons.

Models: examples and enactment. To build on the definitions presented in the CE lessons, all 14 teachers and, in many cases, their students cited and/or enacted examples of both grit and self-control. These examples, as extensions of teachers' and students' definitions, were of interest in relation to the same quantitative findings as the definitions. As with the definitions above, I will first describe the video-based findings associated with the examples, and then connect those findings to the related quantitative results.

Examples. Generally, in-lesson examples of grit and self-control took two forms: people and actions. When a person was identified as a paragon of grit or self-control, generally his actions or accomplishments were the evidence. When an action was noted as an example, sometimes that action was simply described, or more exceptionally a teacher or student(s) enacted that action as an intentional model. Tables 32 and 33 present a sampling of the examples (but not enactments) of grit and self-control that teachers and students generated in these 14 lessons. Within the tables, examples are highlighted (or not) by category:

- No highlighting = example with a non-academic, non-school-based context
- Yellow highlighting = example of current-state intellectual endeavor (e.g., homework)
- Blue highlighting = example of classroom behavioral norm (e.g., sitting still)
- Pink highlighting = example from college/grad school context (e.g., completing college applications)

I discuss the trends associated with these highlights in the section following the tables.

Table 32

A Sampling of Teachers' and Students' Examples of Grit

Teacher	Grade	Teacher-generated Example(s)	Student-generated Example(s)
Ib	4	• Kids: reading all the way through an independent reading packet • Teacher: going to grad school	• Overcoming skill weaknesses in basketball • Improving as a writer with feedback
Id	K	• Teacher's cousin: learning to ride a bike • Teachers: taking hard classes • Kids: persistence on tests • Kids: sitting with good posture	• Learning to ride a scooter • Doing a tough problem • Writing numbers • Putting on lace-up shoes in the morning
Ie	K	• Barack Obama: trying very hard and becoming President • Kids: learning how to read	• Learning how to read • Sitting with good posture in class all day
If	K	• Learning to ride a bike • Teacher: climbing a mountain	• Getting back on bike after falling • Climbing a huge tree house
Ig	2	• Teacher: doing school work instead of just eating dinner and going to bed. • Teacher: training for a marathon. • Learning from mistakes on test and trying a different way.	
IIa	2	• Wilma Rudolph: persisting through polio, winning the Olympics	• Wilma Rudolph • Spelling a really hard word
IIIa	3	• Teacher: learning English as a kindergartener, going to college. • Albert Einstein and Oprah	
IIIb	1	• Persisting with a tough math problem • Teacher: learning to write lessons	• Doing a tricky story problem • Persevering in a football game

Table 33

A Sampling of Teachers' and Students' Examples of Self-control

Teacher	Grade	Teacher-generated Example(s)	Student-generated Example(s)
Ia	1	• Cookie Monster: waiting for a cookie • Teacher: studying for tests in college instead of having fun with friends. Not eating pumpkin pie to lose weight for a wedding • Dog: holding his own leash • Kids: using words when upset, completing classwork, homework, or reading when you want to play	• Not crying when you miss your mom • Being a leader and not following bad students • Staying calm • Waiting your turn for the teacher to call on you • Saying sorry when a classmate bumps you in line • Not calling out and staying in your seat
Ic	3	• Keeping your emotions in check • Not shouting out • Kid: choosing to color	• Controlling your anger • Calming down • Keeping anger and frustration in check
IIb	0	• Not eating dinner before it's done • Kids: sitting on the rug	• Waiting patiently to eat your birthday cake • Waiting patiently for a snack
IIc	0	• Staying calm • Coloring nice and slowly • Waiting for a second Hershey Kiss before eating first • Kid: slowing down his coloring	
IId	0	• Class: not calling out the answer to questions during reading • Bottoms staying in the square (i.e., in assigned seats on rug)	Kid: staying in his square
IIIc	3	• Teacher: going to the gym after school (instead of watching TV) • Class: not dancing in class when a song is playing • Completing the college application process	• Behaving and taking a deep breath if you're angry • Waking up in the morning and walking to school • Writing elaborations • Keeping calm when your classmate messes with you

Several trends emerge from these examples. First, the examples are (a) largely accurate and (b) remarkably similar across classrooms and grades, regardless of source. Some of this accuracy and similarity, of course, comes from younger students parroting or deviating only slightly from teachers' initial examples. For example, in both Teacher Id's and Teacher If's kindergarten classes, the students re-used learning to ride a bike as an example of grit after the teachers said it first. This happens less frequently and/or less obviously, though, in the lessons with the older students (i.e., second-, third-, and fourth-graders).[5] Regardless of the original sources of the examples, whether teacher or student, several are salient across classrooms, grades, and strengths:

- Grit: learning to ride a bike; completing challenging school work

- Self-control: keeping calm; not calling out the answer to a question

- Grit and self-control: sitting (with good posture) in school, going to/succeeding in college

There are, of course, outliers in terms of both accuracy (e.g., Teacher Ia's picture of a dog holding its own leash does not seem a paragon of self-control) and similarity (e.g., putting on lace-up shoes seems an example of grit perhaps unique to kindergarten). By and large, though, students' and teachers' sound examples of grit and self-control were relatively similar.

Building on the similarity of examples that we see across classrooms and grades, in both the grit and the self-control lessons, many examples connect the strength to

---

[5] While a third-grader in Teacher IIIc's classroom does note getting up early to walk to school as his example of self-control (after his teacher says the same), this is the only example of an older student using the same example as the teacher. In Teacher Ib's third-grade classroom, it may appear that the students are citing similar examples as the teacher, but the students shared their examples *before* the teacher.

school-based endeavors (all highlighted in ==yellow==, ==blue==, or ==pink== in the tables above). This stands in contrast to the definitions for these strengths, where grit definitions were often focused on academic endeavors, and self-control definitions were typically focused on intra- or interpersonal skills. Here, with the examples, we see both grit and self-control associated with school-based endeavors. Teachers cited school-based examples—from doing tricky math problems to completing the college application process—of grit and self-control 23 times across these 14 lessons (11 for grit and 12 for self-control). Students also generated original school-focused examples—like writing numbers and not calling out—15 times (seven for grit and eight for self-control). It was clear that both teachers and students saw these strengths as highly applicable to and, in many examples, crucial for school.

These school-focused examples fell into two broad categories: (1) engaging in *intellectual endeavors* (e.g., doing tricky math problems, writing numbers, etc.) and (2) adhering to *classroom behavioral norms* (e.g., sitting with good posture on the rug, not calling out, etc.).[6] In the grit lessons, both the teachers' and the students' examples were far more focused on the former category, i.e., current-state *intellectual endeavors*. Thirteen of the 18 school-based grit examples focused there. In the self-control lessons, again teachers and students shared a focus, but in these lessons, it was *classroom behavioral norms*, with 16 of the 20 school-based examples focused there. In sum, in these 14 lessons, grit was doubly associated with the academic and intellectual aspects of schools: first in the academically focused definitions and then in the examples. Conversely, self-control was doubly associated with non-academic (though still

---

[6] Four of the 14 teachers also made mention of college-focused challenges, e.g., completing the college application process, taking hard college math classes, highlighted in ==pink== in the tables.

sometimes school-based) endeavors: first in the intra- and interpersonally-focused definitions, and then in the classroom behavior-focused examples. Returning to the fundamental tension/symbiosis of schooling that I posited in Chapter I—that balance (or lack thereof) between the academic and social foci of schooling—we can see that grit often positioned in these lessons as the "academic" strength, while self-control is frequently positioned as the "social" strength.

This trend may partially explain the positive association between point-in-time measures of grit and students' reading growth. In these examples, several teachers and students cited a direct connection between grit and reading. Teacher 1b notes reading all the way through an independent reading packet as an example of grit, and a student in Teacher Ie's class says that learning how to read requires grit. If teachers and students are thinking this way early in the school year, the positive association between grit and reading growth seen in the quantitative results is not so surprising. Again, though, this pattern does not readily explain the negative (though weak) association between students' perceptions of their self-control and their reading growth. While teachers and students are not focused on intellectual endeavors in describing self-control, they are not contradicting the relationship between those two phenomena, and thus the cause of the inverse relationship is unclear. In short, the connection between grit and reading growth seen in the quantitative analyses is present in these qualitative findings as well. The negative relationship between students' perceptions of self-control and reading growth is difficult to explain even partially from what we see of the definitions and examples in these 14 lessons.

Enactments. The examples described above represent one form of model presented to students. Another form of modeling presented in these lessons is an intentional enactment of the strength by the teacher or student(s). Real-time enactment, where the teacher or student(s) acted out the strength in skits, role-plays, or other live-action scenarios, was less common than the purely descriptive examples in these lessons. Only four of the 14 teachers used real-time enactment in their lessons, as summarized in Table 34.

Table 34

Real-time Modeling in CE Lessons

Class	Grade	Strength	Description of Modeling
Ic	3	Self-control	Teacher models both going to her "happy place" and "exhaling" when presented with challenging encounters, from both a student- and self-perspective (i.e., getting accidentally kicked on the rug and getting told to do the dishes, respectively). Student models counting to ten in a role-play where another student (playing his parent) is demanding that he do his homework.
Ie	K	Grit	Teacher models using a picture and metacognitive strategies to read a new word. Student models (with teacher assistance) same as above.
IIc	K	Self-control	Teacher models staying calm and coloring slowly and neatly on the board.
IIIb	1	Grit	Teacher models (as if student) persevering through a long math problem.

Even in these few moments of enactment, we see the same trends as described with the examples above. Grit modeling was focused on academic endeavors (i.e., reading and math) and self-control modeling was focused on maintaining behavioral norms (i.e., staying calm and coloring neatly). In short, though enactments were few,

they reinforced the same connections between grit and academics, and self-control and social behavior, that are seen in the definitions and examples described above.

   Rationale. The rationales presented by teachers and students for building grit or self-control were the only CE lesson element that was positively associated (though again, only weakly, and not in all of the models) with students' character growth. Once outlined from the 14 lessons, the rationales for building grit or self-control presented by teachers and students displayed an interesting qualitative pattern: teachers and students often focused on the positive outcomes associated with grit, but dwelled on the negative repercussions of *not* demonstrating self-control. For example, going to college, doing well on tests, and general goal attainment were all rationales for building and using grit. However, *not* saying something mean, *not* going to the Dean's office, *not* shouting out, *not* eating undone food, *not* hurting a friend, etc., were all examples of what could happen if one does *not* demonstrate self-control. While there were also positive rationales for demonstrating self-control (e.g., being successful and going to college), the negatives were more plentiful, as displayed in Table 35.

   Regardless of the valence of the rationales, students may have found something compelling in them, as they were positively associated with students' character growth in both a regression model using grit growth as an outcome (Model 1) and two models using self-control growth as an outcome (Models 4 and 6). I will return to these rationales again in the discussion, when I posit a connection between the "presence of grit = positive" and the "absence of self-control = negative" trend seen here. For now, though, I turn to the final element of the lesson: practice opportunities.

Table 35

Sample Rationales for Building Grit or Self-control

Class	Grade	Strength	Rationale	Student or Teacher?
Ia	1	Self-control	"…give us the ability to regulate or control and also be successful."	T
			"…really important for college…"	T
Ic	3	Self-control	"…I feel so much happier…"	T
			"…you could get suspended…say something you don't mean…go to the Dean's office…hurt yourself and your body."	T
			"…you might explode…"	S
			"…you could even get a headache…"	S
Id	K	Grit	"…so we can always do our best…so we can go to college…"	S
IIb	K	Self-control	"…important to wait for something or it might not be as good."	T
			"You have to wait patiently so you understand the whole question."	S
IIc	K	Self-control	"We should always stay calm because we don't want to hurt our friend."	T
			"[If you color too fast, your picture]…will get messed up."	S
IIIa	3	Grit	"Scholars with the most grit tend to do the best on state exams, will help you not give up on hard math problems…you will start to feel smarter and better about yourself…it will help you get one step closer to college…. Life rewards people who do not give up during hard times."	T
IIIb	1	Grit	"…shows that you are determined to achieve your goal."	T

Practice. In designing practice opportunities within their CE lessons, teachers

could choose to focus on building students' conceptual understanding of the strength

and/or to push students to use the strength real-time through practice opportunities. Most

teachers (68%) in the quantitative sample (n = 88) chose to do one or the other form of practice, as opposed to both (25%) or none (7%).

Within the qualitative data set, all 14 teachers gave students opportunities to deepen their intellectual understanding of the strength, and these opportunities looked very similar across classrooms and grades (e.g., talking with a classmate about a time you/they demonstrated grit or self-control). Only five of the 14 lessons, though, gave students an explicit, intentional opportunity to practice grit or self-control. Table 36 describes those application-based practice opportunities:

Table 36

Application-based Practice in CE Lessons

Class	Grade	Strength	Application-based Practice: Students practiced…
Ic	3	Self-control	…using strategies (e.g., counting to 10, going to their "happy place," exhaling, etc.) when confronted with challenging situations of their own devising in role-plays.
IIb	K	Self-control	…waiting patiently until the teacher asked the entire question before putting a hand up to answer.
IIc	K	Self-control	…(a) coloring slowly and neatly and (b) not eating their first Hershey Kiss before the teacher gave them a second Kiss.
IId	K	Self-control	…staying seated in their square on the rug while popping bubbles with their classmates.
IIIc	3	Self-control	…not dancing while completing classwork, despite a fun song playing.

The obvious trend in these data is that five out of the six teachers who taught lessons on self-control provided at least one opportunity for students to engage in application-based practice. Conversely, none of the eight teachers who focused on grit provided an application-based practice opportunity. On the one hand, this makes sense given that self-control is a strength often measured in minutes (if not seconds), and grit is by definition about persevering on a longer-term basis. One lesson does not lend itself to

practice that must be sustained over weeks, months, or years as with grit, but it does lend itself to practicing a skill that one can demonstrate (or fail to) in an instant as with self-control. On the other hand, this may simply be a quirk of the small sample. In the larger quantitative sample, comparable percentages of grit and self-control lessons provided students with both concept-building *and* application-based practice (i.e., 24% in grit lessons vs. 28% in self-control lessons).

The final trend worth noting about the practice opportunities within the lessons was the relatively small amount of the lesson that was spent in student practice, whether concept-building or application-based. As noted in Chapter III, while sometimes this was simply a limitation of video editing (e.g., a teacher would turn students to their writing and then the video would cut to the next teacher-led portion of the lesson), more often it seemed representative of the lesson's "true" teacher-talk-to-student-work ratio. While the video editing and other limitations of retro-active review make it impossible to say what percentage of the lessons were actually spent in "teacher talk" as opposed to "student work," the balance definitely fell to the former.

In sum, among the 14 videos in the video sample, only five teachers provided students with application-based practice opportunities, and these were limited to classrooms focused on self-control. Moreover, across the 14 lessons, relatively little time was seemingly allotted for students' practice, whether intellectual or application-based practice. In the next section, I will synthesize these findings with those described above to present an explanation for the negative association between student practice and students' character growth.

Surveys and Interviews

Only six teachers completed pre-interview surveys and interviews: four teachers from video Group I (the high reading, high character group) who were all focused on grit in their classrooms, and two teachers from video Group II (the high reading, low character group), with one who was focused on grit and the other on self-control.

Pre-interview surveys. The six interviewees on average provided survey answers that were very similar to those of the teachers in Milson's (2003) much larger sample (n = 920+). In contextualizing their quantitative findings, Milson and Mehlig (2001) deemed mean item scores between 1.00 and 2.99 as "negatively efficacious," those between 3.00 and 3.99 as "neither positive nor negative," and those above 4.00 as "positively efficacious."

Appendix P presents both Milson's (2003) and the present study's findings on the relevant survey items. Notably, the study sample was, on average, positively efficacious (M = 4.19, SD = 0.68) in Milson and Mehlig's terms. Moreover, the study sample viewed themselves (PTE) and teachers broadly (GTE) as more efficacious than those in the Milson (2003) sample. In only one instance (Item 4) did the interviewees, on average, see themselves as notably less efficacious than the other sample: "I am usually at a loss as to how to help a student be more persistent" (study M = 3.83; Milson M = 4.23). In the study sample, only Teacher IIa's average score (M = 3.50) on the survey fell in the "neither positive nor negative" effectiveness category; the rest of the teachers all averaged above 4.00 across the items. In short, the pre-interview survey demonstrated that the six teachers in the sample had generally positive views of their own efficacy as teachers of CE.

That said, aside from the notably different sample sizes, the two data sets are not a like-to-like comparison for two additional reasons, as detailed in Chapter III: (1) the shortened instrument and (2) the adjusted language within the items. The pre-interview survey's main purpose was, as noted above, simply to gather interviewees' preliminary thoughts about the efficacy of CE, relative to some broad-strokes trends in the profession. Thus, these limitations were not of major concern. The interviews, detailed in the next section, were intended as the more in-depth methodology in this phase of the research.

Interviews. Across the six interviews, three major themes related to the quantitative findings emerged:

(1) *Talk vs. walk*: Interviewees described the insufficiency of just "talking to" students about character or as a form of CE. They described a preference for CE that was "genuine," "authentic," and that included in-the-moment modeling and meta-cognitive narration of strength of character in action.

(2) *Connections between academics and character*: Interviewees all affirmed their belief in the connection between academics and character. All of their answers to a question explicitly focused on the relationship (or lack thereof) between academics and character affirmed this belief. Additionally, many interviewees organically connected the two phenomena in other parts of their interviews.

(3) *Enabling school environments*: The Group I interviewees all had regular classroom-based and school-wide CE opportunities, and several described the connections between this lesson and broader classroom and school initiatives. While interviewees' characterizations of their school's CE environments ran the gamut (with the Group II teachers among the most critical or questioning of the

efficacy or sincerity of their schools' approaches), the presence of some form of school-wide CE was universal.

In the sections that follow, I will add depth and nuance to the trends described above, connecting them to both the quantitative and the CE lesson video findings, and drawing on interviewees' actual words whenever possible.

Talk vs. walk. Four teachers across Groups I and II were critical of "talk-heavy" approaches to CE, often advocating for more relatable or palpable approaches. This trend had three major strands: (1) criticism of a talk-heavy approach to CE, (2) a desire for more "genuine" forms of CE, and (3) advocacy for the importance of modeling, and role-modeling in particular, in CE. All three strands will be illuminated through the interviewees' words below, but one quote from Teacher Ig synthesizes the three strands succinctly:

> …[CE] is about modeling things for the kids. I think more than anything else, you can sit there and tell them what perseverance is… but it doesn't actually get at the heart of what it is for them on a day-to-day basis… it really has to be about showing them different ways that you can persevere and doing it in a very honest and genuine way.

Most of the interviewees were critical of CE that emphasized talking to (or "at") students about character. Teacher Ib put this succinctly, saying, "[CE] is not just defining traits and talking at [students] about traits." In this instance, Teacher Ib was speaking critically of her own lesson, in which she felt she had spent too much time simply talking about grit. In other instances, teachers decried a talk-heavy approach in their school's CE. Teacher IIa described her school's CE-focused assemblies in this manner: "We would be like, OK, this person's student of the week because blah blah blah…. So [the school meeting] was very brief, and…my kids were kind of just over it." Teacher Ib described

her school's Community Circles similarly: "…it was kind of droning and it was pretty boring." Beyond individual lessons or specific school-wide events, Teacher Ig expanded on his colleagues' criticisms, talking about teachers' approach to CE in general: "I think we focus a lot on the repetition of specific words, and I don't think that's what is most important… I think it is the embodiment of those traits." This focus on the embodiment or enactment of character strengths was the general preferred approach across many of the teachers. In the words of Teacher Ie, "I think it's one thing to tell students how they should act or what character looks like, but I think it's showing them, as well."

While the interviewees were in agreement that "showing" students character was preferable to talking about it, only two teachers from Group I said that they used modeling for CE (and only one of the two used modeling in the lesson I observed). Teacher Ig was bullish on the import of modeling strength of character. He noted it at least a half-dozen times in his interview, with reference to both his role in CE— "[Character] meant trying to be an excellent role model for [my students] and embody all of those values at the same time…"—and to his students' roles in CE—"…giving the kids opportunities in the classroom to model it. Be like, 'Here's your chance. You can show us right now how to persevere,' and highlighting that." In fact, he cited modeling as the most crucial aspect of CE: "…the most important thing is showing [the students] those values yourself." Similarly, Teacher Ie said, "[CE] requires looking at people who exemplify certain characteristics. I think it is modeling the way…. Role modeling…." It was not clear, however, whether Teacher Ie saw this as an integral part of more formal, structured CE—of the type exemplified in the teachers' lessons in this study—or if he was referring to the more implicit CE that pervades a school year with students.

While modeling was offered as one counter-point to a talk-heavy approach to CE, interviewees across Groups I and II also advocated strongly for more "authentic" or "genuine" approaches to CE, but gave different examples of what such authenticity might look like. For example, in explaining why he chose to have students write their definitions of grit before crafting a class definition, Teacher Ig noted, "I wanted to give them some sort of ownership over the definition and make it something that was genuine to them." For Teacher Ig, authenticity stemmed from students owning the language and vision of grit. Teacher Ig was the only interviewee, though, who noted genuineness as a strength of part of his lesson. Other teachers described it in more aspirational, retrospective terms. For example, many described a desire to have infused their lessons with more authentic or genuine examples or activities. Teacher IIa expressed some reservation about having chosen Wilma Rudolph as an exemplar of grit because, as she said, "I think it would be nice to actually bring people in that are…showing grit currently…because…a lot of times, my kids are like, 'Oh, is that person still alive?'" For her, authenticity stemmed from students' ability to see grit enacted by a real, living person. Similarly, Teacher Id noted that having more real-life examples of grit in action would have improved her lesson. Finally, Teacher Ib, noted, "Instead of talking about [a palpable challenge], I talked about grad school…. The kids were like, 'What are you talking about?' I felt inauthentic." For Teacher Ib, authenticity was about offering examples that kids could see and relate to, as opposed to something that was salient for her only. In sum, the interviewees wanted to make grit real for kids, felt talk-focused CE fell short of that aspiration, and were proponents of both modeling and more authentic examples and activities as a more efficacious approach to CE.

Connection between academics and character. As noted above, the interviewees
all saw connections between academics and character. Their affirmation of these
connections took two forms: (1) when prompted with a specific question about the
connection (or lack thereof) between academics and character and (2) the many examples
they generated—unprompted—during other portions of their interviews.

Early in each interview, before talking about interviewees' particular lessons, I
posed the following questions: "What, if any, is the relationship between CE and
academic growth? How did/do you see this in your classroom?" Each of the six
interviewees described a strong connection between the two. In some cases, teachers felt
as if the two were so related that they were difficult to pull apart, as when Teacher Id
responded, "…the lines are almost blurred and you wouldn't be able to tell the difference
between character and academics." More succinctly, Teacher If responded without
hesitation, "Definitely, they are intertwined." In other cases, teachers were thinking
about grit, in particular, and cited its "direct correlation" with academics, as when
Teacher Ib described ideal student meta-cognition, saying, "…how is [this challenge
requiring grit] affecting you as a person, not just math class?" Regardless of their precise
words, the six teachers were all adamant in their belief that character education and
academic growth were integrally related when asked directly.

Aside from answering the one explicit question about the relationship between
character and academics, teachers across Groups I and II also organically connected these
two phenomena in other portions of their interviews. For example, in describing how she
might know if CE was "working" in her classroom, Teacher Ib noted, "There are the
obvious markers [of grit-focused CE working] like homework completion, study

habits…." Other teachers articulated lay hypotheses about these connections, as when

Teacher Id, given a chance to share her final thoughts, said, "I can speculate that teachers

that have strong character development [in their classrooms]…their [students'] academic

growth would also be just as high, if not higher, than classrooms that are lacking it."

Teacher Ig expressed a similar opinion in explaining why he chose to focus on this

strength in his class and thinking about students at the opposite end of the academic

spectrum, saying,

> It was very easy for some of the lower readers…to give up and feel like,
> 'Oh, I'm just not as smart as everyone else. I really can't do this,' and then shut
> down. And so I felt like by working at that, it really does lend itself
> to…achievement…

Even the teachers with the students demonstrating lower-than-average character growth

saw this connection. Teacher IIa, for example, chose grit as her classroom's focus

strength because of this, articulating her choice this way: "They want to go to college, so

let me pick a character trait to teach that emphasizes that or that they'll need when they

go to college."

Finally, several teachers alluded to the under-emphasized connection between

character and academics in their schools. Interestingly, all of these teachers were in

Group I, where both class-level reading and character growth were notably above

average. Despite this, these teachers still cited a need for a more robust approach to CE.

Teacher Ib, for example, said of her students, "I'm realizing more and more that for them

to be successful, it's not necessarily going to be that math lesson…but so much more…."

In support of Teacher Ib's hypothesis about the longer term import of CE in addition to

strong academics, Teacher Ie, who now leads residential life for an honors college

dormitory, said of his current college students,

> I see students who haven't gotten that CE…. They are extremely book smart… But when I think about [them] dealing with roommate conflicts and learning resilience and learning about awareness of others or awareness of who they are, they don't have that. And that's why I think this work that I do now is so important so that they can be successful in and outside of the classroom.

In short, all teachers articulated a connection between academics and character when asked explicitly, and many articulated this connection organically in other parts of the interview. Some saw the connection as so essential that they were advocates for ensuring a balanced approach to both aspects of schooling for the sake of their students.

Enabling school environment. As described above, all six of the interviewees worked at schools with at least an occasional focus on CE at regular school-wide gatherings. In particular, across both groups, several noted similarities between the values of focus at their school, and the character strength they chose to focus on in their own classroom. Three named grit, in particular, as identical or similar to a value of focus at their school.

All the Group I teachers described having weekly time set aside in their classrooms for CE-focused instruction, and in some cases, named a commercial CE curriculum (e.g., Second Step) that their school provided and that they used during this time. Neither teacher in Group II had CE-specific time in her classroom, nor did either have access to a CE curriculum. Teacher IIa, in fact, noted that her school started using Second Step (a commercial CE curriculum) at the beginning of that school year, but then abandoned the curriculum due to some leadership staffing changes.

At the school level, all the teachers named some example of school-wide gatherings that focused on character, such as Community Circles, all-school assemblies, and breakfasts. The teachers in Group I described robust, consistent versions of these

school-wide structures. Teacher Id described a character award—the Spirit Stick—given to students at the weekly school meeting for exemplifying a school value. Teacher Ig described a daily morning meeting regarding core values, reinforced by a weekly community meeting, focusing each time on a particular value. However, the teachers in Group II again described a different experience. Both teachers in Group II were uncertain and/or skeptical of the efficacy of their schools' approaches. Teacher IIa described a rather rote approach to this work in her school's monthly assemblies, noting, "We never really talked about what those letters [in the school's core values acronym] stood for…. [The students] just kind of memorized it." Teacher IIb described focusing on CE in school-wide structures only during "test prep time," when teachers and students would, during school breakfast, describe and model strengths associated with stronger academic performance.

While the CE resources and structures, as described by teachers, varied predictably across the two groups, with the teachers in Group I reporting more robust, predictable, and well-resourced CE in their classrooms and schools, teachers had a broad range of characterizations of their schools' CE environments. Across the six interviews, two teachers (both in Group I) described their schools' CE positively, one (in Group I) with mixed reviews, one (in Group II) with a neutral perspective, and two (Group I and Group II) in critical terms. In the positive descriptions, teachers noted investment and support from both school leaders and the broader school community as important factors in making their school environments "positive" (Teacher Ie) and "strong" (Teacher Ig) with respect to CE. Teacher Id saw both the strengths—"those character development lessons on Fridays"—and the areas of growth—"building interpersonal skills"—in her

school's approach to CE. The teachers who were critical saw a dearth or a deficiency in the CE in their schools, in both cases stemming from a lack of investment at the school leadership level. These two teachers (at different schools) described an over-emphasis on students' academic performance to the detriment of students' character (and academics). In the words of Teacher IIa, "…teachers valued [CE.] They could see its value in the classroom…. [The administration] felt like it wasn't necessary since they had to make sure that [the students] are moving reading levels, that they know how to take the test."

In sum, while all the teachers had classrooms and/or schools that presented CE opportunities beyond the CE lessons of interest in this study, only the Group I (high reading growth, high character growth) teachers had regular, in-classroom time set aside for CE, oftentimes having received curriculum for those lessons. The Group II teachers described (a) little classroom time spent on CE outside of these lessons and (b) somewhat ineffectual and/or limited school-wide CE opportunities. Teachers across both groups had mixed reviews of their schools' environment with respect to CE, with several pointing to school leadership as the arbiter—for better or worse—of the presence and efficacy of school-wide CE.

Summary

In this chapter, I have presented the qualitative findings in order of discovery, starting with a component-by-component examination of the CE lessons, and then turning to the pre-interview surveys and interviews. I will summarize these findings here, synthesizing them as they relate to each research question.

Research Question #1: How do changes in grit or self-control among the students in the study sample compare to those in students in other studies using similar measures?

Given that students in classrooms focused on grit grew more (with respect to strength of character) than students in classrooms focused on self-control, I hypothesized that teachers and students would define, model, and practice self-control in such ways that might make it seem "harder" to demonstrate than grit. The definitions, models, and practice in the 14 CE lessons did not bear this out. If anything, the definitions and examples of self-control that teachers and students presented—focused as they were on context-specific, relatively finite examples of self-control (e.g., not calling out)—made self-control seem more attainable than grit, often presented in the context of long-term and lofty ambitions (e.g., college graduation).

However, the rationales in these lessons perhaps explained some of the strength-based growth differential: the rationales presented in the self-control lessons were generally focused on the disadvantages of not demonstrating self-control. In contrast, the rationales presented in the grit lessons presented the advantages of using and building grit. This avoidance-based view of self-control may have positioned it as more difficult to attain than the approach-based view of grit (Lyubomirksy, 2007). I will explore this idea in more depth in the concluding chapter.

Additionally, several interviewees mentioned grit as a focus of their school's approach to CE, while none mentioned self-control as a school-wide focus or topic. Thus, students may have had more exposure to grit as a quality worthy of pursuit, perhaps seeing more models, getting more practice, and/or seeking more incentives in and beyond the classroom and throughout the school year. These additional "at-bats"

may have facilitated greater growth with respect to this strength, as compared with self-control.

Research Question #3: What, if any, relationship exists between students' academic and character skills and growth?

I hypothesized that the positive correlations between point-in-time measures of grit and reading growth were related to teachers and students connecting grit and academics in the CE lessons. Moreover, I posited that the negative correlation between point-in-time measures of self-control and reading growth were related to teachers' and students' focus on interpersonal (as opposed to academic) forms of self-control (Tsukayama et al., 2013). Both of these hypotheses were confirmed by the videos. Teachers and students defined and described grit in far more academic terms than self-control, occasionally making direct connections between grit and reading, in particular, while generally connecting self-control to inter-personal aspects of conduct.

In the interviews, teachers agreed unanimously that CE and academic growth were closely related. While this unfailing belief does little to explain the negative relationship between point-in-time self-control and reading growth, it may explain part of the positive associations between grit status and reading growth. If teachers were articulating these connections both in these lessons and in their regular exchanges with students, as they were in their interviews with me, students may have internalized the positive association between grit and academic growth. That the Group I teachers, in particular, generally found such messaging both prevalent and purposeful at their school may make this explanation even more likely.

Research Question #4: What aspects of CE curriculum, planning, pedagogy, and teachers' choices might be associated with students' academic and character growth?

First, teachers' above-expectation performance on the CE lessons may have been buttressed by their "positively efficacious" outlook as measured—albeit retrospectively—on the CEEBI survey. This strong performance may have also been supported by both the regular CE opportunities and resources at their schools, particularly as described and experienced by the Group I teachers.

Given the negative association between the definitions and practice opportunities in the lessons and students' character growth, I hypothesized that either these parts of the lessons were full of unchecked foibles *or* they were relatively strong, thus giving students a higher bar for grit or self-control, and thereby deflating students' (and in some cases, teachers') sense of their own character status. The definitions and practice opportunities were not flawed. On the contrary, they were, on the whole, solid introductions to grit and self-control. They likely did lead to students improving their understanding of the strengths, and therefore, potentially increasing their standards for themselves in their subsequent rounds of character surveys. In this way, a strong lesson would seem to lead to decreased perceived character growth, but may, over the longer term, contribute to students' actual character growth. This is a difference, i.e., perceived versus actual growth, that I will take up in the final chapter related to a consideration of how reference bias may have contributed to the negative association between these lessons and perceived character growth.

Conversely, I posited that the trend of positive association between rationales and students' character growth might suggest that teachers' rationales were particularly compelling, such that students were striving (and succeeding) to demonstrate character

growth. In viewing the lessons, it was difficult to determine how the rationales, in particular, might be positively associated with students' character growth. However, a clear pattern emerged, whereby the presence of grit was promoted through association with positive outcomes (e.g., higher grades and college matriculation) and the absence of self-control was deterred based on associated negative outcomes (e.g., calling out or hurting). While these findings do not explain the positive association with students' character growth, I will return to them in the final chapter where I will connect them to a broader vision of goal-setting and attainment that may explain students' differential grit and self-control growth outcomes.

VI—DISCUSSION AND CONCLUSION

This study put forth a hypothesis that a one-shot intervention in the form of a CE lesson could propel students' character growth, as other social-psychological interventions have propelled students' social and academic outcomes (Yeager & Walton, 2011). In this case, the hypothesis was not supported—the CE lessons were negatively (though weakly) associated with students' character growth. However, this and several other related findings—both expected and unexpected—suggest possibilities that might guide future practice and research. In particular, across research questions and methodologies, the findings suggest that students and teachers are positioning grit and self-control quite differently relative to each other and relative to academic endeavors. Additionally, and again across research questions and methodologies, the findings suggest that CE-related opportunities beyond the single CE lesson of interest may be positively associated with some of the character-related outcomes one sees in this study. Finally, both the quantitative and qualitative data suggest that the CE lessons re-position grit and self-control, such that students and teachers have a higher bar for these strengths after the lessons.

I detailed the individual quantitative and qualitative findings that led to these larger conclusions in the preceding two chapters. Here I will synthesize those individual findings across research questions and methodologies, organizing them by the three major categories of findings I outlined above: (1) Grit vs. Self-Control, (2) One-Shot vs. Ongoing CE, and (3) Re-Positioning Grit and Self-Control. I will connect these broad

conclusions to the salient extant literature. I will conclude with some suggested directions for future practice and research, as well as a final reflection on character as I conceive of it as an educator versus character as I investigated it herein.

Grit vs. Self-Control

A clear pattern emerged across both the quantitative and qualitative findings, such that grit and self-control were associated with different patterns of character and academic outcomes *and* were portrayed differently in the CE lessons. Quantitatively, the students in the classes that focused on grit demonstrated more character growth, and students deemed to have more grit demonstrated more reading growth. Conversely, students in classes focused on self-control demonstrated less character growth, and students who deemed themselves to have initially higher levels of self-control demonstrated less reading growth. In short, grit was associated with more desirable outcomes in this school-based context.

On the one hand, these differences may seem logical, given that the survey instrument presented grit in a somewhat more academic context and self-control in a somewhat more social context:

Grit
- Finishes whatever he or she begins
- Tries very hard even after experiencing failure
- Works independently with focus

Self-Control
- Allows others to speak without interruption
- Is polite to adults and peers
- Keeps his/her temper in check

And while one sees both grit and self-control associated with positive academic outcomes in the extant literature (e.g., Duckworth et al., 2007 and Duckworth and Seligman, 2005, respectively), one does not find positive associations between interpersonal self-control, in particular, and positive academic outcomes (e.g., Park et al., 2017). Thus, it might not be so surprising that grit was associated positively with reading growth, while self-control, as measured by the items above, was not.

On the other hand, there is nothing about these survey items that would explain—on the surface—why students focused on grit demonstrated notably more year-long character growth than students focused on self-control. Moreover, while the cognitive leap from the grit items to the academic and intellectual examples teachers and students were referencing in their grit-focused lessons is not vast, it is still a leap. There is nothing in these items, for example, that explicitly references the math problems, writing development, tests, or college matriculation that teachers and students alike were summoning when they offered examples of what grit looked like in action. And yet somehow, grit was often positioned as the "academic" strength and more positively associated with the desirable outcomes in these academic settings.

Prominent among—and perhaps partially explanatory of—the several grit- and self-control-related patterns was that associated with the rationales presented in the CE lessons. In these 14 lessons, the presence of grit was often promoted through association with positive outcomes (e.g., higher grades and college matriculation) and the absence of self-control was often discouraged based on associated negative outcomes (e.g., calling out or hurting). Related to this pattern, Lyubomirsky (2007) describes two broad categories of goal-setting and attainment: approach and avoidance. The former

category—approach—represents goals and / or goal-attainment defined or conceptualized as requiring one to do something new, or more, or better than before. The latter category—avoidance—represents goals and / or goal-attainment defined or conceptualized as requiring one to stop doing something, or do it less than before. For example, losing weight could be conceptualized in terms of approach (e.g., exercise more) or avoidance (e.g., eat less). Lyubomirsky (2007) and colleagues find that when goals and / or goal attainment are framed in terms of approach, people are more likely to succeed than when they are framed in terms of avoidance.

One sees these two conceptions of goals and goal-attainment in the rationales for grit and self-control in these 14 lessons. The rationales for having, using, and developing grit were framed exclusively in approach terms: have grit, do your best; use grit, go to college; get grittier, feel smarter. Whereas the rationales for having, using, and developing self-control were often framed in avoidance terms: have self-control, don't call out; use self-control, don't hurt your friend; build self-control, don't go to the dean's office. To be sure, some self-control rationales were framed in positive, approach-based terms, but many were framed in terms of avoidance, and in contrast with the grit rationales, which were *all* about approach, the self-control rationales are notably different.

This difference in the framing of the rationales for grit and self-control may partially explain why students focused on grit demonstrated more character growth than students focused on self-control. And beyond rationales, other elements of the lessons exhibited some of this same avoidance versus approach framing, for self-control versus grit, respectively. As but one example, the word "not" was used at least seven times in

the examples of self-control that students and teachers presented in the lessons. The word "not" was never used in the grit examples. If this differential positioning of grit and self-control—as the approach and avoidance strengths, respectively—was representative of how teachers and students were conceptualizing the reasons for building grit or self-control more broadly, students may have reaped some of the benefits of the approach-framed grit goal and fallen prey to some of the avoidance framing of self-control, as described by Lyubomirsky (2007) and colleagues.

One-Shot vs. Ongoing Character Education

A central hypothesis of the study was that one might harness the positive effects of robust, ongoing, school-wide CE (e.g., Benninga et al., 2003) into the potent, low dosage container of a one-shot intervention (e.g., Yeager & Walton, 2011) in the form of a single CE lesson. That hypothesis was not supported here, but the related findings raised larger questions about what might explain what *was* happening in and around the classrooms where one did see particularly strong character growth.

Both the strong—and strengthening as the school-year continued—relationships between students' and teachers' ratings of students' strength of character *and* the six interviewees' reports of their schools' approach to CE pointed towards ongoing CE as a factor in many of the classrooms in this study. That students' and teachers' conceptions of students' strength of character became ever more strongly associated as the school year went on may have been related to ongoing CE. (As mentioned previously, this convergence may also have been partially related to ceiling effects associated with the repeated use of the same survey items.) In talking with the four Group I interviews—all

of whom had classrooms with above average character growth—all reported robust, ongoing CE, both in their classrooms and in school-wide structures. And while one cannot draw generalizable conclusions from a sample so small, it bears noting that these four teachers worked across four different schools, four different CMOs, and three different grades. In short, from a teaching placement perspective, they represented almost as stratified a sample as possible based on the larger, quantitative sample (n = 88). In sum, the study presented modest evidence in support of school-wide CE.

Re-Positioning Grit and Self-Control

The Character Education Lesson, Definition, and Practice

The negative (though weak) relationship between these lessons and students' character growth is difficult to explain solely through the definitions and practice opportunities in the lessons alone. However, it is conceivable that the negative relationship between definitions and practice, in particular, and the CE lessons more broadly, and students' character growth may be partially explained by reference bias.

Reference bias, as first described in Chapter II is the relationship between how one rates oneself or others on self- or other-report surveys and the other people, groups, or contexts that serve as reference points for that evaluation (Heine, Lehman, Peng, and Greenholtz, 2002). It is one of several threats to the validity of self- and other-report data, and both Seider (2012) and West et al. (2014) cite it as a potential explanation for the decline in self-reported strength of character among the students in their respective studies.

As noted in Chapter IV, Seider (2012) found that students in schools with character-focused missions tended to remain stable or demonstrate a decline in strength of character, as measured by self-report surveys, from the beginning to the end of the school year. In other words, in the schools Seider studied, where certain strengths of character—integrity, risk, and grit—were clearly and consistently defined, students' estimation of their strength along those dimensions remained stable or decreased over the course of the school year. Seider explained this with the following logic:

> …surveys completed by students in the opening weeks of the academic year capture students at their most idealistic regarding their own levels of self-discipline and perseverance in carrying out their academic responsibilities. At the outset of the school year, students have not yet confronted decisions about completing their homework versus watching television, or studying for a test versus socializing with friends. (pp. 108-109)

It stands to reason that when students' beginning-of-year optimism is tested—and occasionally fails—against the school-based challenges that Seider describes, students' estimation of their own grit, for example, declines. West et al. (2014) presented a similar reference-bias hypothesis to explain why they, too, found that students' perceptions of their grit and self-control declined over the course of the school year.

Connecting Seider's (2012) findings from a study of schools with robust, school-wide CE to the findings from this study of classrooms with a single CE lesson—let alone one component, like the definition or practice, in that lesson—is a stretch, but one can see how the same logic *may* apply: if a CE lesson on grit, for example, is particularly clear and compelling, it may re-position grit for students, challenging their initial optimistic perceptions of their own grit. Equipped with a clearer and more ambitious vision of grit, students may rate themselves lower on grit in the next iteration of the survey (post-lesson) than they did at the beginning of the school year (pre-lesson). Quantitatively, this

would make it look like students had become "less gritty," when really, they had—as a result of the CE lesson—simply re-positioned "true grit" as higher relative to their own current expression and development of that strength. This logic would, in part, explain both why definition and practice quality, in particular, and CE lesson quality, more broadly, were all negatively (though weakly) related to students' character growth. In short, a better definition, practice, and overall lesson lend themselves to creating a clearer and more realistic vision of grit or self-control, thereby heightening students' reference bias, and thus decreasing their estimation of their strength of character relative to their own beginning-of-year ratings.

If one examines this possibility through the lens of practice in these CE lesson, for example, one might see more concretely how reference bias may be at play. In several of the classrooms with application-based practice, teachers were recreating situations like those studied by Mischel and colleagues (1989) where students were tempted and later rewarded if they could resist the temptation. In other classrooms, two teachers had students engage in versions of Mental Contrasting with Implementation Intentions (MCII), the goal-setting strategy developed by Oettingen (e.g., Duckworth, Kirby, Gollwitzer, & Oettingen, 2013). Teacher IIIa for example, had students choose a goal, write about what obstacles they may encounter in attempting to reach that goal, and then plan with a classmate what they would do when they encountered said obstacles. Also, several teachers had students write about a time they demonstrated grit or self-control, as in the intervention studied in Cohen et al. (2006). In all of these cases, the teachers may have been drawing on their GSE coursework as models for the lessons, as Mischel and

colleague's findings, Oettingen's MCII work, and Cohen et al.'s study were all part of course readings, class discussions, or activities.

In each of these cases, however, the practice, rather than building grit or self-control, could have reminded students of how hard it is to demonstrate grit or self-control. To use the Mischel-esque self-control practice as an example, Mischel's initial self-control studies were not intended to develop self-control; they were intended to determine for how long students could exercise it. By analogy, running one timed mile does not make students more fit. Rather, it makes most of them far more cognizant of just how hard it is to run a timed mile. And if they thought themselves fit before, they have a far more realistic—and likely more negative—conception of where they stand relative to exemplary fitness. So, too, might one round of exercising self-control or participating in one well-facilitated MCII demonstrate for students what real and challenging obstacles lie between them and a desirable outcome or goal. And even when the CE lesson practice took the form of students identifying a time when they *had* shown grit, for example, students often named something they struggled to succeed at for a long time and at which they were still working—like developing strong writing skills, for example. Again reminding themselves—rightly—of just how hard it is to demonstrate "true grit." In all such examples, teachers' well-intended efforts to build students' strength of character may have simply proved just how difficult it is to demonstrate grit or self-control, thereby heightening the students' reference bias.

In short, perhaps the "best" (i.e., highest scoring on the CE lesson rubric) definitions and practice opportunities were best because they re-positioned grit or self-control for the students (and likely teachers, as well), offering a clearer, more realistic,

and likely higher standard for what exemplary grit or self-control might look like. And, as a measurement artifact of this re-positioning, students (and teachers) appeared to grow less with respect to that strength after those "better" lessons with "better" definitions and practice opportunities.

Rationales

Reference bias may explain the negative association between the lessons (broadly) and definitions / practice (particularly) and students' character growth. The same logic, though, does not hold for rationales. A particularly compelling rationale for why one should have, use, and/or build grit, for example, would not lend itself to a higher bar for grit. Rather it should compel students toward grittiness, wherever that standard was set. In other words, a better rationale would not re-position grit or self-control per se; it would re-position why one might want to build or use grit or self-control.

That the rationales in the CE lessons were positively (if again, only weakly) associated with students' character growth supports this view of re-positioning the goal. A strong rationale—one that presents "the importance of the character strength in a clear and inspiring fashion," to use the language of the CE lesson rubric (Appendix H)—would urge students to want to strive after further development of this strength and thus should, as in the study's findings, be associated with subsequent growth.

Implications for Practice and Research

Implications for Practice

The findings and connections above suggest several directions for future practice and research. From a practice standpoint, the study suggests three major implications for K-12 classrooms. First, helping students to see how they can bring key "non-cognitive" strengths to bear on academic endeavors may be yet another way to help students succeed in school in even the youngest grades. In this study, students' grit was positively associated with their reading growth, and all of the teachers focused on grit in the qualitative study associated grit and academic endeavors in their CE lessons. Building on this, however, teachers may need to be more attentive to how they are framing students' practice of non-cognitive strengths. A one-time exposure to the strength may serve more to discourage or dampen rather than to invigorate or build students' non-cognitive strengths.

Relatedly, a second implication for practice is the seemingly advantageous nature of ongoing classroom-based and school-wide CE. A one-shot lesson was not enough to promote character growth. In fact, in this study, a one-shot lesson was negatively (though weakly) associated with students' character growth as measured herein. For the four interviewees, however, whose students demonstrated above-average character and reading growth, robust ongoing classroom and school-wide CE were the norm. While these practices had flaws as described by the teachers, they were—across different organizations, schools, and grades—part of the context associated with students demonstrating notable growth, both academically and socially, during their year with those four teachers.

Finally, students need to know why they should strive toward a desired outcome in our classrooms and schools, as opposed to why they should avoid an undesired outcome. Grit was positioned in many of the classrooms in this study as a desirable outcome because of its association with other positive outcomes, such as success in class, school, and education more broadly. Self-control, on the other hand, was often described in photo-negative fashion: here is what you will miss or the trouble you will cause if you do *not* demonstrate self-control. This positioning of grit as a strength to be sought after and "approached," in Lyobmirsky's (2007) terms, perhaps explains why (a) students focused on grit demonstrated more character growth and (b) students with more grit demonstrated more reading growth. Moreover, even absent the approach / avoidance dichotomy and across the grit / self-control divide, lessons with stronger rationales for building the strength in question were associated with greater character growth. Telling students why they are learning something thus may be a crucial part of setting them up for engagement and success—be the goal demonstrating more grit or mastering long division.

The study also has practice implications for teacher education. From the broadest perspective, if the "goal" was to help teachers propel students' academic and character growth, perhaps preparing teachers to implement an established intervention of the sort detailed by Yeager and Walton (2011) might have been a more reliable means of helping them to meet the goal, as opposed to allowing them to chart their own course within a relatively flexible CE lesson structure. More particularly, the coursework supporting teachers' CE efforts should present sample strengths clearly and individually, helping teachers to see the academic and social implications of each (such that they exist). In this

study, teachers often positioned grit as the "academic" strength and self-control as the "social" strength, but teacher educators could help teachers to see both sides of both strengths and practice presenting them with nuance to their young students.

From a measurement perspective, teacher preparation coursework could help mitigate some of the limitations of self-report surveys, by providing teachers with opportunities to norm their administration in terms of tone, timing, and accommodations for the youngest, non-reading survey-takers. Additionally, coursework could provide teachers with shared, "touchstone" examples of the strengths that teachers could in turn use with their students to set a common standard for grit or self-control at the beginning of the school year.

Finally, given the apparent benefits of a school-wide approach to CE, teacher education could prepare teachers to advocate for whole-school CE structures. Many of the teachers in this study did so of their own accord, but they may have been better positioned to do so earlier in their careers had their coursework attended to such issues of advocacy.

Implications for Research

The study points toward several implications for future research, and in particular, to the challenges of "measuring" something as difficult to define, let alone measure, as strength of character. For example, there were many different character surveys in the literature. In some cases, the authors of one study would use several different iterations of their own instrument within the same study (e.g., Park et al., 2017). This makes comparing outcomes across studies difficult to impossible. It seems the field has come far

enough from a measurement perspective to settle on some agreed-upon and oft-used instruments, so that we can start to gather and compare reliable, valid, and similar data from sample to sample.

Additionally, much of the literature on character as measured by self-report surveys (including this study) is hindered by reference bias. Teacher-reports mitigate this somewhat, as do more objectively measured self-reports (e.g., how many times today did you call out in class?). Still, though, some "truer" measure of character would allow for more valid inferences, not just about character, but also about interventions designed to develop it. As a start, re-thinking standard school-based measurement intervals may contribute to the search for "truer" measures. Often, as in this study, Seider's (2012), West et al.'s (2014), and many other school-based studies, measures are taken at the beginning of the school year and then again at the end of the school year, in an effort to measure year-long growth. However, given reference bias might be a compelling potential factor in *any* (even "truer") self- and other-report measures of character, fall-to-fall and spring-to-spring comparisons may be better determinants of students' "true" year-long character growth than the more typical fall-to-spring comparisons. In other words, if students are at their most optimistic each fall, instead of comparing fall measures to spring measures, it may make more sense to compare, for example, students' character measures taken in the fall of second grade to those taken in the fall of third grade, so as to "naturally" control for beginning-of-year optimism. None of the literature referenced for this study took such an approach, but perhaps this would be a better standard measurement interval for determining growth of "non-cognitive" strengths.

Finally, it was surprising to me that no published character growth data existed for elementary school students. As evidenced by the strong correlations between students' and teachers' perceptions of students' character at the beginning of the school year, even the youngest school children understand concepts like grit and self-control. So…let us teach them…and study it! I would be eager to know how students in this study compare with same-age peers from other schools, experiences, and backgrounds.

Conclusion

I noted in the "Researcher Positionality" section in Chapter I, that I designed the teacher-level "inputs" that undergirded this study (e.g., the graduate level syllabi, curricula, classes, etc.), and therefore, I wanted them to "work." In some respects they did: students in the study demonstrated both reading and character growth beyond expectation. The teachers, too, performed beyond expectation in the two lessons of interest—both the CE lesson of primary concern and the academic lesson used as a comparison. In the main respect, though, the focus of the study did not "work": the CE lesson designed as the "intervention" had only a weak association—and a negative one at that—with students' character growth. In other words, teachers who earned better scores on the CE lesson rubric had students who demonstrated less character growth (than their peers in classrooms with lower performing teachers).

Given what I know now, I can re-shape the associated curriculum for future cohorts of teachers at the study site. In particular, I intend to prepare teachers to implement a validated set of interventions, designed to propel both students' academic and social growth. Additionally, I intend to design coursework that engages teachers in a

somewhat more philosophical examination of key character strengths. For example, asking teachers to describe the differences between grit and self-control; how they do and do not contribute to feelings of belonging, fulfillment, and success in school and life; and how children may or may not see them differently than their adult teachers. Finally, I want to help teachers prepare to advocate for and support CE efforts more broadly in their schools. In informal conversations, teachers often mention to me that they do not feel like they have adequate time or resources to incorporate CE as they would like into their classrooms. Sometimes this is a question of lack of teacher knowledge and skill, which I intend for our coursework to address; other times, however, this is positioned as a lack of investment in CE on the part of school leaders. Regardless of the obstacles, I would like for the teachers at the study site to feel prepared and empowered—in terms of knowledge, skills, mindsets, and grit—to help their students flourish.

Flourishing is what character, as defined by Peterson and Seligman (2004) in this study's theoretical foundations, both allows and exemplifies. This conception of character matches my conception of character in the world. It is a cross-cultural, pluralistic character comprising many strengths that people, in general, and children, in particular, can see in themselves and others. Cross-cultural character, though, is not to say culture-less or even context-less character. Quite the contrary, whether one is citing Graham and colleagues (e.g., Spivak, White, Juvonen & Graham, 2015) or Kidd et al. (2013), it is clear to me—and before I read either—that people see and value character strengths differently depending on their culture and their context. That the young students in this study saw grit and self-control shining prominently in themselves was inspiring to me.

That I had to limit the study—and therefore the study's investigation of character—to just grit and self-control was dissatisfying, at best, and deeply disappointing, at worst.

As I noted in Chapter I, to limit character, broadly, to just grit and self-control, in particular, seems thin. Do I see grit and self-control in children? Absolutely, as did the children and teachers in this study. Do I want children to build their grit and self-control in school? Absolutely, as did the teachers I interviewed. And yet, I want children to see and build so much more than simply grit and self-control. I want them to see themselves as curious, creative, kind, loving, and zesty, and build those strengths and many other strengths, as well.

But context and culture do not just matter, they often reign. At the time of the study, Mischel and colleagues' marshmallow studies made self-control the stuff of popular thought. The Atlantic, The New Yorker, and National Public Radio were making his findings—which he had composed into a book for a lay audience—fodder for teachers' lounges everywhere in 2014. Similarly, Angela Duckworth and colleagues made grit the talk of the teacher town, as well. She published a widely-publicized book and came to speak to a cohort at the GSE during the time of the study. Consequently, it is not surprising that so many teachers at the study site—and likely across the country, if not world—were focused on grit and self-control during this time.

And why should they not be? Duckworth, Mischel, and colleagues proved—in study after study—that grit and self-control were huge predictors of the types of outcomes teachers want for their students: better grades, more education, better health, more wealth. What teacher would not want those things for her students? A teacher would be unlikely to wish her students to have worse grades, less education, worse

health, or less wealth. However, she might be likely to ask, "Are grit and self-control the only avenues to these ends?" She may also be likely to ask, "Are there not other outcomes—friends, happiness, self-fulfillment—that I also care about for my students?" And she may also wonder, "Might these outcomes be associated with more than just the students' character strengths? Might they also be associated with my own?" And armed with such curiosity, she may start to try to answer newer, perhaps better questions with her students' flourishing as her steady goal.

REFERENCES

Aristotle. (Trans. 2009) *The Nichomachean Ethics*. Translated by David Ross. Oxford: Oxford University Press.

Bennett, W. J. (1986). Moral literacy and the formation of character. (Address given to the Harvard Club, New York, NY. October 30. Washington, DC: Department of Education.

Bennett, W. J., & Delattre, E. J. (1978). Moral education in the schools. *Public Interest*, *50*, 81-98.

Benninga, J. S., Berkowitz, M. W., Kuehn, P., & Smith, K. (2003). The relationship of character education implementation and academic achievement in elementary schools. Journal of Research in Character Education, *1*(1), 19-32.

Berkowitz, M. W. (2011). Moral and character education. In K. R. Harris et al. (Eds), *APA educational psychology handbook, volume 2: Individual differences and cultural and contextual factors* (p. 3). Washington, DC: American Psychological Association.

Berkowitz, M. W., & Bier, M. C. (2004). Research-based character education. *The Annals of the American Academy of Political and Social Science*, *591*(1), 72-85.

Berkowitz, M. W., & Bier, M. C. (2005). *What works in character education: A report for policy makers and opinion leaders*. Washington, DC: Character Education Partnership.

The Bill and Melinda Gates Foundation. (2012). Asking students about teaching: Student perception surveys. Seattle, WA. Retrieved from http://k12education.gatesfoundation.org/wp-content/uploads/2015/12/Asking_Students_Practitioner_Brief.pdf

Blackwell, L.S., Trzniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246-263.

Blazar, D., & Kraft, M. A. (2015). Teacher and teaching effects on students' academic behaviors and mindsets. Cambridge, MA: Mathematica Policy Research. Retrieved from http://www.mathematica-mpr.com/our-publications-and-findings/publications/teacher-and-teaching-effects-on-students-academic-behaviors-and-mindsets

Borghans, L., Duckworth, A. L., Heckman, J. J., & Weel, B. (2008). The economics and psychology of personality traits. *The Journal of Human Resources*, *43*(4), 972-1059.

Calfee, R., & Sperling, M. (2010). *On mixed methods: Approaches to language and literacy research*. New York, NY: Teachers College Press.

Center for Public Education. (2015). Learning to read, reading to learn: Why third-grade is a pivotal year for mastering literacy. Alexandria, VA. Retrieved from http://www.centerforpubliceducation.org/Main-Menu/Policies/Learning-to-Read-Reading-to-Learn-At-a-Glance/Learning-to-Read-Reading-to-Learn-Full-Report-PDF.pdf

Character Education Partnership. (2010). *Eleven principles of effective character education*. Washington, DC: Character Education Partnership.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307-1310.

Cohen, J. (2006). Social, Emotional, Ethical, and Academic Education: Creating a climate for learning, participation in democracy, and well-being. *Harvard Educational Review*, *76*(2), 201-237.

Coleman, J. S. (1966). Equality of educational opportunity. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education/National Center for Education Statistics.

Collaborative for Academic, Social, and Emotional Learning. (2005). *Safe and sound: An educational leader's guide to evidence-based social and emotional learning programs—Illinois edition*. Chicago, IL: CASEL.

Cremin, L.A. (1961). *The transformation of the school*. New York: Alfred A. Knopf.

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

Dewey, J. (2009) [1916]. Democracy and education: An introduction to the philosophy of education. New York, NY: WLC Books.

Duckworth, A. L. (2011). The significance and self-control. *Proceedings of the National Academy of Science*, *108*(7), 2639-2640.

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2014). Self-control in school-age children. *Educational Psychologist*, *49*(3), 199-217.

Duckworth, A. L., Grant, H., Loew, B., Oettingen, G., & Gollwitzer, P. M. (2011). Self-regulation strategies improve self-discipline in adolescents: benefits of mental-contrasting and implementation intentions. *Educational Psychology*, *31*, 17-26.

Duckworth, A. L., Kirby, T. A., Gollwitzer, A., & Oettingen, G. (2013). From fantasy to action: Mental contrasting with implementation intentions (MCII) improves academic performance in children. *Social Psychological and Personality Science*, *4*(6), 745-753.

Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate practice spells success: Why grittier competitors triumph at the national spelling bee. *Social Psychological and Personality Science*, *2*(2), 174-181.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92(*6), 1087-1101.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (Grit-S). *Journal of Personality Assessment*, *91*(2), 166-174.

Duckworth, A. L., Quinn, P. D., & Seligman, M. E. P., (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology*, *4*(6), 540-547.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance in adolescents. *Psychological Science*, *16*(12), 939-944.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*(4), 237-251.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405-432.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Taylor and Francis.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.

Dweck, C. S., Walton, G. M., & Cohen, G. L. (2011). Academic tenacity: Mindsets and skills that promote long-term learning. White paper prepared for the Gates Foundation. Seattle, WA. Retrieved from https://ed.stanford.edu/sites/default/files/manual/dweck-walton-cohen-2014.pdf

The Education Trust—New York. (2017). *See our truth*. New York, NY. Education Trust. Retrieved from https://1uxzg93r9wfm26lrln198mbd-wpengine.netdna-ssl.com/wp-content/uploads/sites/5/2017/10/See-Our-Truth.pdf

Elementary and Secondary Education Act, Pub. L. 89-10 (1965).

Emmons, R.A., & McCullough, M. E. (2003). Counting blessings versus burdens. *Journal of Personality and Social Psychology*, *84*(2), 377-389.

Eskreis-Winkler, L., Shulman, E. P., Beal, S.A., & Duckworth, A.L. (2014). The grit effect: Predicting retention in the military, the workplace, school, and marriage. *Frontiers in Personality Science and Individual Differences*, *5*, 36.

Every Student Succeeds Act, Pub. L. 114-95 (2015).

Farrington, C.A., Roderick, M. Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become better learners: The role of noncognitive factors in shaping school performance: A critical literature review.* Chicago, IL: The University of Chicago Consortium on Chicago School Research. Retrieved from http://ccsr.uchicago.edu/sites/default/files /publications/Noncognitive%20Report.pdf

Froh, J. J., Sefick, W. J., & Emmons, R. A. (2008). Counting blessings in early adolescents: An experimental study of gratitude and subjective well-being. *Journal of School Psychology*, *46*(2), 213-233.

Giltay, E.J., Geleijnse, J.M., Zitman, F.G., Hoekstra, T., & Schouten, E.G. (2004). Dispositional optimism and all-cause and cardiovascular mortality in a prospective cohort of elderly Dutch men and women, *Archives of General Psychiatry,* 61, 1126–1135

Gollwitzer, A., Oettingen, G., Kirby, T. A., Duckworth, A. L., & Mayer, D. (2011). Mental contrasting facilitates academic performance in school children. *Motivation and Emotion*, *35*, 403-412.

Hartshorne, H. (1930). Sociological implications of the character education inquiry. *American Journal of Sociology*, *36*(2), 251-62.

Hartshorne, H., & May, M. A. (1928-1930). *Studies in the nature of character:* Vol. 1, *Studies in deceit;* Vol. 2, *Studies in service and self-control*; Vol. 3, *Studies in the organization of character*. New York, NY: Macmillan.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, *82*, 903-918.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172-177.

Hunter, J. D. (2000). *The death of character*. New York, NY: Basic Books.

James, W. (2008). *Talks to Teachers on psychology: And to students on some of life's ideals*. Rockville, MD: Arc Manor. (Original work published in 1899)

Kerbow, D., & Bryk, A. (2005). STEP literacy assessment: Technical report of validity and reliability. Chicago, IL: STEP.

Kidd, C., Palmieri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, *126*(1), 109-114.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed), *Handbook of Socialization Theory and Research*. Chicago, IL: Rand McNally.

Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Cambridge, MA: Harvard University Press.

Kukla-Acevedo, S. (2009). Do teacher characteristics matter?: New results on the effects of teacher preparation on student achievement. *Economics of Education Review*, *28*(1), 49-57.

Leming, J. S. (1993). In search of effective character education. *Educational Leadership*, *51*(3), 63-71.

Leming, J. S. (2008). Theory, research, and practice in the early twentieth century character education movement. *Journal of Research of Character Education*, *6*(2), 17-36.

Levine, A. (2006). Educating school teachers. The Education Schools Project. Retrieved from http://edschools.org/pdf/Educating_Teachers_Report.pdf

Lickona, T. (1991). *Educating for character: How our schools can teach respect and responsibility*. New York, NY: Bantam.

Lickona, T. (2004). *Character matters: How to help our children develop good judgment, integrity, and other essential virtues*. New York, NY: Touchstone.

Lickona, T., & Davidson, M. (2005). *Smart and good high schools: Integrating excellence and ethics for success in school, work, and beyond*. Courtland, NY: Center for the 4th and 5th Rs.

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

Lyubomirsky, S. (2007). *The how of happiness: A new approach to getting the life you want*. New York, NY: Penguin.

Martin, J. E. (2016). We should measure students' noncognitive skills. *Education Week*, July 27, 2016. Retrieved from http://www.edweek.org/ew/articles/2016/07/27/we-should-measure-students-noncognitive-skills.html

McClellan, B. E. (1999). *Moral education in America: Schools and the shaping of character from colonial times to the present*. New York, NY: Teachers College Press.

McGuffey, W. H. (1836). The eclectic second reader. Cincinnati, OH: Truman and Smith.

McGuffey, W. H. (1837). The eclectic third reader. Cincinnati, OH: Truman and Smith.

Mischel, W., & Baker, N. (1975). Cognitive appraisals and transformations in delay behavior. *Journal of Personality and Social Psychology*, 31, 254-261.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*, 933-938.

Milson, A. J. (2003). Teachers' sense of efficacy for the formation of students' character. *Journal of research in character education*, *1*(2), 89-106.

Milson, A. J., & Mehlig, L. M. (2001). *The efficacy beliefs of elementary teachers regarding character education*. Paper presented at the annual meeting of the American Education Research Association, Seattle, WA.

Milson, A. J., & Mehlig, L. M. (2002). Elementary school teachers' sense of efficacy for character education. *The Journal of Educational Research*, *96*(1), 47-53.

Moffitt, T. E. et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*, 2693-2698.

Moore, B., Mischel, W., & Zeiss, A. (1976). Comparative effects of the reward stimulus and its cognitive representation in voluntary delay. *Journal of Personality and Social Psychology, 34*(3), 419-424

Nash, R. J. (1997). *Answering the "Virtuecrats": A moral conversation on character education*. New York, NY: Teachers College Press.

National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: U.S. Department of Education.

*National Defense Education Ac*t, Pub. L. No. 85-864 (1958).

No Child Left Behind Act of 2001, Pub. L. No. 107-110 (2002).

Noddings, N. (1988). An ethic of caring and its implications for instructional arrangements. *American Journal of Education*, *96*(2), 215-230.

Noddings, N. (1995). Care and moral education. In W. Kohli (Ed), *Critical conversations in philosophy of education* (pp. 137-148). New York, NY: Routledge.

Noddings, N. (2002). *Educating moral people: A caring alternative to character education*. New York, NY: Teachers College Press.

Nucci, L. (2001). *Education in the moral domain*. Cambridge, UK: Cambridge UP.

Nucci, L., Drill, K., Larson, C., & Browne, C. (2005). Preparing preservice teachers for character education in urban elementary schools. *Journal of Research in Character Education*, *3*(2), 81-96.

Oxford English Dictionary. (1971). Oxford, UK: Oxford University Press.

Park, A., Tsukayama, E., Goodwin, G. P, Patrick, S., & Duckworth, A. (2017). A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual competencies in children. *Contemporary Educational Psychology*, *48*, 16-27.

Peterson, C. (2006). *A primer in positive psychology*. Oxford, UK: Oxford UP.

Peterson, C., Park, N., & Seligman, M. E. P. (2005). Assessment of character strengths. In G. P. Koocher, J. C. Norcross, & S. S. Hill, III (Eds), *Psychologists' desk reference* (2nd ed., pp. 93-98). New York, NY: Oxford University Press.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. New York, NY: Oxford University Press.

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew Effects in reading. *Review of Educational Research*, *84*(2), 203-244.

Piaget, J. (1965). The moral judgment of the child. New York, NY: Free Press. (Original work published in 1932)

Power, F.C., Higgins, A., & Kohlberg, L. (1989). *Lawrence Kohlberg's approach to moral education*. New York, NY: Columbia University Press.

Ready, D. (2013). Associations between student achievement and student learning: Implications for value-added school accountability models. *Educational Policy*, *27*(1), 92-120.

Robertson-Kraft, C., & Austin, K. (2014). The character of achievement: An analysis of teachers' instructional practices for character education. *Journal of Character Education*, *11*(2), 109-128.

Robertson-Kraft, C., & Duckworth, A. L. (2014). True grit: Trait-level perseverance and passion for long-term goals predicts effectiveness and retention among novice teachers. *Teachers College Record*, 116(3), No. 17352. Retrieved from http://www.tcrecord.org.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, *94*(2), 247-252.

Ryan, K. (1988). Teacher education and moral education. *Journal of Teacher Education*, *39*(5), 18-23.

Salinger, T. (2006). Policy decisions in early literacy assessment. In D. K. Dickinson & S. B. Neuman (Eds), *Handbook of early literacy research, volume 2* (pp. 427-444). New York, NY: The Guilford Press.

Seider, S. (2012). *Character compass: How powerful school culture can point students toward success*. Cambridge, MA: Harvard Education Press.

Seider, S., Gilbert, J. K., Novick, S., & Gomez, J. (2013). The role or moral and performance character strengths in predicting achievement and conduct among urban middle school students. *Teachers College Record*, *115*(8), 1-18.

Seligman, M. E. P., Ernst, R. M., Gillham, J., Reivich, K., & Linkins, M. (2009). Positive education: Positive psychology and classroom interventions. *Oxford Review of Education*, *35*(3), 293-311.

Seligman, M.E.P., Steen, T.A., Park, N., & Peterson, C. (2005). Positive psychology progress: Empirical validation of interventions. *American Psychologist*, *60*(5), 410- 421.

Spivak, A. L., White, S. S., Juvonen, J., & Graham, S. (2015). Correlates of prosocial behavior of students in ethnically and racially diverse middle schools. Merrill-Palmer Quarterly, 61(2), 236-263.

Stephens, J. M., Young, M. F., & Calabrese, T. Does moral judgment go offline when students are online? A comparative analysis of undergraduates' beliefs and behaviors related to conventional and digital cheating. *Ethics and Behavior*, *17*(3), 233-254.

Study Site Graduate School of Education. (2014). Reading growth assessment evaluation checklist. Internal document.

Taylor, A. Z., & Graham, S. (2007). An examination of the relationship between achievement values and perceptions of barriers among low-SES African American and Latino students. *Journal of Educational Psychology, 99*(1), 52-64.

Tough, P. (2012). *How children succeed*. New York, NY: Houghton Mifflin.

Tsukayama, E., Duckworth, A. L., & Kim, B. (2013). Domain-specific impulsivity in school-age children. *Developmental Science*, *16*(6), 879-893.

Tyack, D.B. (1974). *The one best system: A history of American urban education*. Cambridge, MA: Harvard University Press.

Tyack, D.B., & Hansot, E. (1982). *Managers of virtue: Public school leadership in America, 1820-1980*. New York, NY: Basic Books.

United States. National Commission on Excellence in Education. (1983). A nation at risk: the imperative for educational reform: a report to the Nation and the Secretary of Education, United States Department of Education. Washington, D.C.: The Commission: [Supt. of Docs., U.S. G.P.O. distributor], 1983.

U.S. Department of Education, Office of Educational Technology. (2013). *Promoting grit, tenacity, and perseverance: Critical factors for success in the 21st century*. Retrieved from http://www.ed.gov/technology.

U.S. Department of Education, Social and Character Development Research Consortium. (2010). *Efficacy of school-wide programs to promote social and character development and reduce problem behavior in elementary school children*. Retrieved from http://ies.ed.gov/ncer/pubs/20112001/pdf/20112001.pdf

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard UP.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic outcomes and health outcomes of minority students. *Science*, *331*, 1447-1451.

West, M. R., Kraft, M. A., Finn, A. S., Martin, R., Duckworth, A. L., Gabrieli, C. F., & Gabrieli, J. D. (2014, June). Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. In *CESifo Area Conference on Economics of Education Munich: CESifo (September)*.

White, J. T. (1909). *Character lessons in American biography for public schools and home instruction*. New York, NY: The Character Development League.

Wilson, T. D. (2006). The power of social-psychological interventions. *Science*, *313*, 1251-1252.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*(2), 267-301.

Appendix A

The Language of Character

The *Oxford English Dictionary* (*OED*) (1971) contains no less than 19 distinct definitions of the noun form of "character," divided between seven rather concrete denotations (e.g., "a distinctive mark impressed, engraved, or otherwise formed") and twelve more humanistic qualities (e.g., "the sum of the moral and mental qualities that distinguish an individual…") (p. 280). Of these 19 distinct definitions, eight comprise two or more sub-definitions, and thus the actual number of uses of the noun number over 40. And not a single one describes precisely the vision of character espoused in this study.

When one adds the variety of verbs, adjectives, and nouns used to represent very similar concepts across the fields of philosophy, psychology, economics, policy, and education, the possibilities are myriad. Table A.1 represents some of the most frequently used terminology attempting to describe these similar concepts. For example, the middle and right-hand columns of Table A.1 (on the next page) can be combined to create many of the common phrases used to describe the concept of interest in this study. For example: "moral habits," "social strengths," "personality traits," and "21st century skills" are just a few of the myriad phrases one reads when investigating "character."

Table A.1

Character Terminology across Philosophy, Psychology, Economics, Policy, and Education

Verbs	Nouns / Adjectives	Nouns
Instill	Ethics / ethical	Factors
Impart	Morals / morality / moral	Traits
Teach	Virtue	Qualities
	Caring	Mindsets
Learn	Social	Dispositions
Understand	Emotional	Strengths
	Socio-emotional	Behaviors
Promote	Inter-personal	Skills
Develop	Intra-personal	Habits
Build	Personality / personal	Competencies
Strengthen	Non-cognitive	Strategies
	Soft	Tactics
	21st century	Process(es)
		Approach(es)

Appendix B

Peterson and Seligman's (2004) 24 Character Strengths

1. **Creativity:** coming up with new and productive ways to think about and do things
2. **Curiosity:** taking an interest in experience for its own sake; finding things fascinating
3. **Open-mindedness:** examining things from all sides and not jumping to conclusions
4. **Love of learning:** mastering new skills and topics on one's own or in school
5. **Wisdom:** being able to provide good advice to others
6. **Bravery:** not running from threat, challenge, or pain; speaking up for what's right
7. **Grit:** finishing what one starts; completing something despite obstacles
8. **Integrity:** speaking the truth and presenting oneself sincerely and genuinely
9. **Zest:** approaching life with excitement and energy; feeling alive and activated
10. **Love:** valuing close relationships with others; being close to people
11. **Kindness:** doing favors and good deeds for others; helping them; taking care of them
12. **Social intelligence:** being aware of motives and feelings of other people and oneself
13. **Citizenship:** working well as a member of a group or team; being loyal to the group
14. **Fairness:** treating all people the same; giving everyone a fair chance
15. **Leadership:** encouraging a group of which one is a valued member to accomplish
16. **Forgiveness:** forgiving those who've done wrong; accepting people's shortcomings
17. **Modesty:** letting one's victories speak for themselves; *not* seeking the spotlights
18. **Prudence:** being careful about one's choices; not taking undue risks
19. **Self-control:** regulating what one feels and does; being disciplined
20. **Appreciation of beauty:** noticing and appreciating all kinds of beauty and excellence
21. **Gratitude:** being aware of and thankful for the good things that happen
22. **Hope:** expecting the best in the future and working to achieve it
23. **Humor:** liking to laugh and tease; bringing smiles to other people; seeing a light side
24. **Spirituality:** having beliefs about the higher purpose and meaning of the universe

Appendix C

The Language of Character Education

As noted earlier, some scholars and practitioners whose work might be classified as CE resist the label. For example, Cohen (2006) includes "the caring approach….[which] stresses the importance of recognizing and developing caring relationships and infusing caring, relational, and social-emotional themes into school curricula" (p. 205) under the umbrella of CE. However as mentioned above, Noddings (2002), who identifies as a "care theorist" and promotes "care ethics," seeks to distance herself and her practice from CE, which from her perspective, "relies on indoctrination, and indoctrination is not an acceptable method of education" (pp. 3-4).

Noddings's point is represented more subtly by the verbs in the left-most column of Table A.1. Again, each of these verbs is the predicate of choice in one or more of the papers cited in this study and each verb represents a slightly different conception of CE. The first three verbs—instill, impart, and teach—all imply the teacher as the subject and the students as the indirect objects in a sentence describing CE. For example, "The teacher **instills moral habits** [in her students]." The vision of CE that unfolds when one uses the first three verbs is—if not necessarily "indoctrination"—certainly one where the teacher has the knowledge and wisdom and her students are the recipients thereof.

Similarly, with the two verbs in the middle of the left-most column of Table A.1—learn and understand—the students, while now the implied subjects of a sentence describing CE, are still construed as the recipients of knowledge or understanding. For example, "The students **learn caring approaches** [from their teacher]." As we will see in later sections, we do not yet know enough about CE to know if the effects—such that they exist—of such instruction are so cleanly hierarchical. In other words, a great middle school math teacher is a master of the content and pedagogy required to teach middle school math (among myriad other skills required to excel at this difficult and complex work). She has a more sophisticated understanding and far higher level of mastery of math content and skills than any of her students (rare prodigies excepted). We do not know, though, what it takes to be a great teacher of gratitude. Thus, while students may build strength of character, in part, via what they learn from teachers, to construct one's vision of CE exclusively around such a one-way model seems absent both empirical and practical support.

The final three verbs in Table A.1—develop, build, or strengthen—broaden the vision of possibilities for CE. They could imply the teacher, the students, or both as the subjects:

- The teacher **builds character strengths** [in her students].
- The students **build** [their] **character strengths**.
- The teacher and students **build** [their] **character strengths**.

The more ambiguous, two-way, and fluid conception of CE implied by these constructions seems appropriate to the empirical and practical evidence at hand. Consequently, when describing the CE lesson that is the focus of this study, I will use verbs like "develop," "build," and "strengthen" to imply that children are not blank slates with respect to character and neither are teachers the "masters" of character by virtue of their position in the classroom.

Appendix D

Reading Growth Using Standardized Measures

While standardized achievement tests proliferate, the ways to compare students' performance across them (or even across grades on the same type of test) are not straightforward. To try to better understand students' progress from grade to grade across tests, Hill, Bloom, Black, and Lipsey (2008) examined the effect sizes associated with both the reading and math portions of a subset of commonly used K-12 standardized achievement tests. Using the difference between students' mean scale scores in adjacent grades on the reading portions of the tests, for example, Hill and colleagues calculated average annual reading growth (measured in effect size) for each set of adjacent grades.[1] They found that students' growth in reading test scores from grade to grade is largest from kindergarten to first grade (1.52 SD) and declines steadily from there (e.g., 0.97 SD from first to second grade, 0.32 SD from fifth to sixth grade, and 0.06 SD from 11th to 12th grade). To quote Hill et al.: "The natural growth in test scores declines as students age" (p. 173). These findings make intuitive sense: in the early grades, one is learning to read, making notable strides each year. As one's education progresses, one's reading skills become more nuanced and refined, but to a lesser degree each year. At some point, most of us become the readers we are for life and do not continue to add ever-increasingly sophisticated reading skills to our repertoire.

Hill and colleagues' (2008) findings also support an oft-cited teacher adage: from kindergarten through third grade, students are learning to read, and from fourth grade on, they are reading to learn (Center for Public Education, 2015). This teachers' dictum essentially describes the rapid growth that young children make in reading ability during their early years of formal schooling, which is then followed by the use of those reading skills to acquire knowledge in other disciplines—history, literature, and science, for example. The divide, though, between "learning to read" and "reading to learn" is not as clear cut in Hill et al.'s results as it is in the adage. Hill et al. find that students' reading gains decline steadily over time, as opposed to sharply at some particular grade; though it is during the third-to-fourth grade transition that students' average reading growth drops below 0.50 SD (i.e., less than one third of the growth from the first to second year of schooling) (p. 173). Regardless of the findings' support of teachers' lounge truisms, the trend is clear: students' year-to-year reading growth declines through the course of their formal schooling.

---

[1] Because the tests Hill and colleagues (2008) choose to use in their analyses are administered to students in the spring of each academic year, the reading growth calculated in these analyses incorporates both learning during the school year and learning (or lack thereof) during the summer. This differs from many classroom-based reading assessments in which "baseline" measures are taken at the start of the school year and "final" measures are taken at the end of the school year, leaving out a formal measure of what students learned (or did not) over the course of the summers previous or subsequent to that school year.

Appendix E

Pre-21st Century Character Education in America

CE in Colonial through Late 19th-Century America

The earliest examples of formal schooling in America—the schools serving the children of the colonial era—were far closer to what one might see in a contemporary Sunday School classroom than a contemporary public school classroom. The emphasis in colonial schooling was overwhelmingly on religion and morality. While students were expected to acquire literacy and some rudimentary mathematical skills, the materials were largely religious (McClellan, 1999). Thus, Colonial schools focused on character through an almost exclusively religious lens.

The formation of the United States of America brought about a new emphasis on education. Thomas Jefferson and others advocated strongly for expanding the nation's public schools, with an emphasis on "republican values" and the teaching of "virtue" (McLellan, 1999, pp. 12-13). This expansion of the public school system began in the early 19th century, and accelerated particularly with Horace Mann's Common School movement. From the mid-19th century through to the early 20th, Horace Mann's Common School was the model for schooling in the United States. As noted previously, scholars generally agree that Mann's vision of the nonsectarian, universal Common School provided the foundation for America's public schooling (Cremin, 1961; Tyack, 1974). For Mann, the ultimate purpose of the Common School was creating an American citizenry united by shared ethics and common skills, regardless of background (Hunter, 2000). Though Protestant morality was certainly part of the education in the Common Schools, Mann admonished his supporters when they were "teaching creeds offensive to immigrants" (McClellan, 1999, p. 36). Mann's vision was the Common Schools and the education therein was for everyone, regardless of their particular religious beliefs.

The McGuffey (1836 & 1837) readers, however, the most common text seen in mid-19th century American schools, emphasize Protestant religion and moral values. The books are full of biblical passages like "The Ten Commandments" (McGuffey, 1836, pp. 229-233) and moral tales like "How to Guard against Temptation" (McGuffey, 1837, p. 60). The "Ten Commandments" passage exemplifies the McGuffey structures. First, it is written in language geared to elementary school students and begins with a short pre-amble explaining the origins of the Ten Commandments (i.e., "written by God himself, on two tables of stone" [McGuffey, 1836, p. 230]). The passage is followed by a set of ten questions, from simple recall—"By whom were the Ten Commandments written?"—to more rhetorical—"Wasn't God very merciful to give His son to die for our sins?" (p. 232). The questions are followed by a list of focus vocabulary words (e.g., wherefore, commandments, adultery, etc.) (p. 233). This cycle repeats itself throughout the scores of passages across the six volumes: introduction and passage, questions, and vocabulary. The questions—particularly about the moral and biblical texts—are straight forward; they ask the reader to recall information or relay key points from the same perspective as the authors. They typify how teachers and schools approached CE for much of the 19th century: highly-structured, morally- or religiously-focused, and with little practice beyond the rote.

CE from the Progressive Era through the 1930s

The early twentieth century saw the rise of the Progressive movement and the decline of religiosity in the nation's schools (Tyack and Hansot, 1982). Perhaps the most famous of the progressive educators, Dewey (1916) saw school as a site for character development. He and other progressive educators conceived of CE as organized around democratic ideals, in which students developed both an understanding of democratic society and the values and virtues necessary to participate in a democracy.

In the popular parlance of the times, the approach to CE endorsed by Dewey and other progressives was deemed "indirect," meaning that their approach was more "child-centered" and activity-based. These techniques—hailed as more modern by their champions—were contrasted with the "direct" approach. A "direct" approach to CE (branded as old-fashioned by many of the researchers and education administrators of the era) had a specific place in the formal curriculum and teachers used structured materials and standard pedagogies like lectures and basal readings (e.g., McGuffey readers) to lead character-focused instruction (Leming, 2008).

One such "direct" approach to CE is exemplified by James White, author of *Character Lessons…* (1909) and founder of the Character Development League.[2] White identified and described 31 traits—from obedience to patriotism—that together compose a person's character. In addition, White designed "character lessons" to accompany each trait, whereby "*direct*, *specific*, and *systematic*" instruction led a child learn how to "adjust morality to the various conditions and situations of life" (p. ii). For example, White's passages on the "elucidation and training" of self-control suggest that the teacher "instill the idea that it is generally the unwary whose passions are so often allowed the mastery" (p. 44). More actively, White suggested "how to bear annoyances and vexations without exhibiting any outburst of temper" (p. 45) as another fruitful topic of instruction. Finally, White gives several examples of Americans who exemplified self-control (e.g., George Washington, Edgar Allan Poe[3], and Henry Ward Beecher) and offers some suggestions for further reading on the topic, such as various fables, Emerson's "Essays on Character," and Psalm XCI. The latter is one of relatively few biblical references in White's work. In many ways, White's work is the "direct," character-only corollary to the McGuffey readers without the religious emphases.

The "direct" versus "indirect" debate was fanned, in part, by the CE research of the late 1920's. Hartshorne and May's (1928-1930) "Character Education Inquiry" served as a rallying point for many of the researchers and education administrators who were advocating more progressive, "indirect" approaches to CE. This oft-cited line from the conclusion of the first volume of the study decried the traditional "direct" practices of CE: "the mere urging of honest behavior by teachers or the discussion of standards and ideals…has no necessary relation to conduct…the prevailing ways of inculcating ideals probably do little good and may do some harm" (Hartshorne & May, 1930, p. 413, cited in Leming, 2008, p. 21). Progressives and other "modern" educators used Hartshorne and

---

[2] At the time of publication of *Character Lessons…* (1909), the Character Development League was located at 70 Fifth Avenue in New York. I do not know if White would be encouraged or dismayed to know that building is now within a block of both a Taco Bell and a Curves gym—opportunities to test one's self-control in the forms of both restraint and exertion.

[3] Apparently Poe's reputed gambling, drinking, and drug use were not common knowledge in 1909 or perhaps White had a more liberal definition of self-control than his oeuvre may otherwise suggest.

May's (1928-1930) findings to support their vision of "indirect" approaches to CE, while other groups (e.g., the Character Education Association) continued to advocate for more "direct," structured approaches to CE.

Early CE and This Study

In sum, from Colonial times through the 1930s in America, many philosophers, theologians, publishers, principals, and teachers focused on promoting character in educational settings. Though their emphases—religious, moral, civic virtue—were varied and their approaches perhaps all the more so, there was a surprising continuity of focus on the social aspects of schooling across these several centuries.

When one compares these early enactments of CE in American classrooms with Seider's (2012) conception and the enactment of CE in this study, both connections and discontinuities appear. First, both the schools in Seider's study and the enactment of CE investigated in this study deviate from the religiosity of early CE in this country. The values of interest in Seider's schools (e.g., integrity, compassion, perseverance, civic commitment, and risk-taking) and the strengths of interest at the GSE that served as the study site (i.e., curiosity, gratitude, grit, love, optimism, self-control, social intelligence, and zest) are far from the religious doctrines of early CE. That said very few (if any) of the values and strengths of interest to Seider and this study lie in direct conflict with the religious messages of early CE. Moreover, one will find several of the contemporary strengths of interest in the "character codes" of the late 19th and early 20th centuries' CE curricula and materials (McClellan, 1999, p. 50). Additionally, some of the grander desires of proponents of CE of that era live on in both Seider's schools and the form of CE in this study. For example, both the schools and the study see knowledge of and shared commitment to at least a core set of strengths among school-aged children as a desirable outcome of CE, in much the same way that Horace Mann and early 20th educators advocated for CE as a means to ensuring a citizenry united around shared values. Finally, as it concerns the pedagogical, Seider's vision and the CE lesson of interest in this study allow for and encourage both "direct" and "indirect" approaches to CE. While clearly defining a character strength for students would be an example of a "direct" approach—something both the teachers in Seider's schools and the teachers in this study do consistently—allowing children to practice this strength via athletics, clubs, or self-directed activities (again, as done in Seider's schools and many of the classrooms of interest in this study) exemplifies the "indirect" approach that Dewey and others advocated. In short, while the religiosity and prescriptive moralism of early CE is not an element of CE as conceived, enacted, and investigated in this study, one can see many traces of earlier approaches to CE—from content to pedagogy—in these contemporary examples.

Mid-to-Late 20th Century CE

CE in schools has taken many forms over the more than three hundred years of established formal education in America. From the dawn of World War II through the early phases of the Cold War, CE became far less prominent in America's public schools. The CE of earlier eras—of both the "direct" and "indirect" styles—was replaced by citizenship, civics and social studies coursework that sought to bolster national identity and ideals in the face of international threat (Leming, 2008). That said, there were a few

researchers whose work propelled CE (though not by that term) through this era. Lawrence Kohlberg and Carol Gilligan were among the most prominent and prolific, though their work took different orientations and approaches, relative to each other and to the traditional value-centered CE of earlier eras.

Kohlberg (1969) approached the social aspects of development and, later, schooling through the lens of cognitive moral development. In the tradition of Piaget (1932/1965) and Vygotsky (1978), Kohlberg proposed a stage-based theory of moral development comprising six stages—from heteronomous morality (Stage 1) to universal ethical principles (Stage 6). He proposed these as universal stages, with individuals making different rates of progress through them and reaching differing maximal stages. In Kohlberg's research, individuals' current level of moral development was assessed via their responses to case studies or scenarios featuring moral dilemmas. Those individuals who had more sophisticated responses, by Kohlberg's measures, were deemed to have reached higher stages of moral development.

After defining and testing his theories of moral development, Kohlberg turned his efforts to moral education, founding a school based on the promotion of moral development as manifested in a democratic approach to schooling (Power, Higgins, and Kohlberg, 1989). Kohlberg believed that "at least in American public schools teachers have the responsibility to teach values, though they do not have the right to impose their own, or any, set of values on their students" (p. 15). He advanced these educational theories through "just community schools," in which not only the moral dilemmas, but also the rules, consequences, and structures of daily school life were debated and decided democratically by students and teachers alike (McClellan, 1999). Kohlberg believed dialog and decision-making about the factors that defined justice in students' lives would serve as an ideal moral education.

Once Kohlberg's research assistant and later his philosophical sparring partner, Carol Gilligan thought Kohlberg's emphasis on justice, as well as his findings that men fared better on his moral dilemma assessments than women, a male-centric view of moral development. She proposed a related, but alternative stage-based theory of development focused on caring (McClellan, 1999). Nel Noddings was Gilligan's protégé and was more focused than her mentor on the educational implications of caring. In her definition, caring is a purely relational act, and it can be natural or ethical. In the former case, we care because we want to (e.g., because the object of our care is a loved one), while ethical caring is "...aimed at establishing, restoring, or enhancing the kind of relation in which we respond freely because we want to do so" (Noddings, 1995, p. 138). Noddings (2002) saw school as one of many places where children can and should develop their ability to care, buttressed both by adult modeling and ample opportunities for practice.

Kohlberg, Gilligan, and Noddings all positioned themselves in opposition to traditional, values-based CE. Kohlberg referred to traditional CE as the "bag of virtues" approach (cited in McClellan, 1999, p. 83). Noddings (2002), as noted previously, viewed traditional CE as indoctrination. That said, both camps—the cognitive developmentalists and the feminists—picked what some would call a particular character strength (i.e., justice and caring, respectively) upon which to focus their work. Moreover, both advocated for both "direct" and "indirect" exposure to and practice with this strength in schools. Finally, while Kohlberg's camp was the only one of the two to attempt to measure the concept of interest in quantitative terms, both groups were interested and

engaged in the work of helping children to build and develop related knowledge and skills in school. In short, though these researchers pitted themselves against CE (and the CE advocated by the next era of intellectuals, described in the section below, would in many ways in turn pit themselves against Kohlberg and Gilligan), there was likely more overlap of foci and approach than either camp would admit.

Neo-Classical CE

Many in the field of CE (e.g., Bennett and Delattre, 1978; Ryan, 1988; and Lickona, 1991) saw the latter half of the twentieth century as an era of sharp moral decline in America. For these scholars, the rising rates of single-parent households, drug use, and incarceration of the late 1980's and early 1990's signaled a tolling of the bell for America's collective morality. Their vision of our moral decline was so pervasive and consistent that Nash (1997), describing the field of character educators of that era, referred to the group collectively as "declinists" (p. 17). The declinists claimed that the America they knew and loved was going down; the Kohlbergs and the Gilligans had done little to save it (and likely some to accelerate its sinking); and the only way to save the country was to right the moral ship through traditional CE.[4]

The neo-classical (or if one prefers, declinist) approach to CE could be dubbed—in earlier terms—a "direct" style. Before he became Secretary of Education, Bennett—arguably *the* spokesman for neo-classical CE—exhorted educators to reemphasize moral education, but not the "cognitive moral development" of Kohlberg and colleagues (1969). Bennett saw this and related approaches as over-intellectualized and failing to give students a clear sense of right and wrong (Bennett and Delattre, 1978). As Secretary, Bennett described his vision of CE more clearly: specific traits (e.g., honesty, courage, kindness, etc.); clarity regarding right and wrong; strong examples in the form of teachers, schools, and stories; and a "begin with the basics" approach, as opposed to starting with thorny moral issues like "nuclear war, abortion, creationism, or euthanasia" (Bennett, 1986, pp. 13-14). Bennett and his contemporaries were reacting directly to (and against) Kohlberg, Gilligan, and the other heirs to Dewey's progressive approach to CE. They represented one more swing of the CE pendulum—like McGuffey to Dewey, like Dewey to civics, like civics to Kohlberg and company, so too from Kohlberg and company to the neo-classicists.

---

[4] The "declinist" vision of American morality is not the orientation of this study, but regardless of whether one thinks America's children are in the midst of a moral decline, incline, or plateau, this study operates from the perspective that increased strength of character and academic achievement are positive goals regardless of the current state of either. Thus the question of *where* one student or one teacher (let alone an entire country) starts is not the issue, but rather how, from *wherever* one starts, one develops strength of character and academic achievement.

Appendix F

Character Growth Survey

Each teacher at GSE picked one of eight sets of survey items (shown below), each associated with a different character strength. The survey items used a 5-point Likert scale ranging from 1 = *Very much unlike me* to 5 = *very much like me*.

Zest
- Actively participates
- Shows enthusiasm
- Invigorates others

Grit
- Finishes whatever he or she begins
- Tries very hard even after experiencing failure
- Works independently with focus

Curiosity
- Is eager to explore new things
- Asks and answers questions to deepen understanding
- Actively listens to others

Love
- Demonstrates care for others
- Demonstrates care for self
- Recognizes and appreciates others' strengths

Social Intelligence
- Is able to find solutions to conflicts with others
- Demonstrates respect for feelings of others
- Knows when and how to include others

Gratitude
- Recognizes and shows appreciation for others
- Recognizes and shows appreciation for his/her opportunities

Self-Control
- Allows others to speak without interruption
- Is polite to adults and peers
- Keeps his/her temper in check

Optimism
- Gets over frustrations and setbacks
- Believes that effort will improve his/her future

Appendix G

The Academic Lesson Rubric

Rubric Row (Mod#: Row#)	(4) Exemplary	(3) Proficient	(2) Foundational	(1) Attempting	(0) Lacking
OBS-201: 1 Classroom Culture	a. Based on the indicators assessed, the teacher's application of Classroom Culture indicators is exemplary	a. Based on the indicators assessed, the teacher's application of Classroom Culture indicators is proficient	a. Based on the indicators assessed, the teacher's application of Classroom Culture indicators is foundational	a. Based on the indicators assessed, the teacher is attempting to apply Classroom Culture indicators	a. Based on the indicators assessed, the teacher's application of Classroom Culture indicators is lacking
OBS-201: 2 Teaching Cycle	a. Based on the indicators assessed, the teacher's application of Teaching Cycle indicators is exemplary	a. Based on the indicators assessed, the teacher's application of Teaching Cycle indicators is proficient	a. Based on the indicators assessed, the teacher's application of Teaching Cycle indicators is foundational	a. Based on the indicators assessed, the teacher is attempting to apply Teaching Cycle indicators	a. Based on the indicators assessed, the teacher's application of Teaching Cycle indicators is lacking
OBS-201: 3 Content	a. Based on the indicators assessed, the teacher's application of Content indicators is exemplary	a. Based on the indicators assessed, the teacher's application of Content indicators is proficient	a. Based on the indicators assessed, the teacher's application of Content indicators is foundational	a. Based on the indicators assessed, the teacher is attempting to apply Content indicators	a. Based on the indicators assessed, the teacher's application of Content indicators is lacking
OBS-201: 4 Self and Other People	a. Based on the indicators assessed, the teacher's application of Self and Other People indicators is exemplary	a. Based on the indicators assessed, the teacher's application of Self and Other People indicators is proficient	a. Based on the indicators assessed, the teacher's application of Self and Other People indicators is foundational	a. Based on the indicators assessed, the teacher is attempting to apply Self and Other People indicators	a. Based on the indicators assessed, the teacher's application of Self and Other People indicators is lacking
SWITCH ROW OBS-201: *The teacher will demonstrate application of the Relay GSE Elements of Effective Instruction in her classroom*	a. Overall, the teacher's application of the Relay GSE Elements of Effective Instruction is exemplary	a. Overall, the teacher's application of the Relay GSE Elements of Effective Instruction is proficient	a. Overall, the teacher's application of the Relay GSE Elements of Effective Instruction is foundational	a. Overall, the teacher is attempting to demonstrate application of the Relay GSE Elements of Effective Instruction	a. Overall, the teacher's application of the Relay GSE Elements of Effective Instruction is lacking

Appendix H

The Character Education Lesson Rubric

Rubric Row (Mod#: Row#)	(4) Exemplary	(3) Proficient	(2) Foundational	(1) Attempting	(0) Lacking
SOP-222: 1 The teacher will (re)introduce the character strength to students	a. OBJECTIVE: The teacher presents a clear and inspiring character-focused lesson objective b. DEFINITION: Student(s) (re)introduce(s) the character strength with a clear definition c. RATIONALE: Student(s) explain(s) to fellow students the importance of the character strength in a clear and inspiring fashion (with minimal teacher guidance)	a. OBJECTIVE: The teacher presents a clear character-focused lesson objective b. DEFINITION: The teacher (re)introduces the character strength with a clear definition c. RATIONALE: The teacher explains to students the importance of the character strength in a clear and inspiring fashion	a. OBJECTIVE: The teacher presents a character-focused lesson objective that lacks clarity b. DEFINITION: The teacher (re)introduces the character strength with a definition that lacks clarity c. RATIONALE: The teacher explains to students the importance of the character strength, but his explanation lacks clarity or inspiration	a. OBJECTIVE: The teacher presents a lesson objective, but not one that relates to character b. DEFINITION: The teacher (re)introduces the character strength but without a definition c. RATIONALE: The teacher explains to students the importance of the character strength, but his explanation lacks clarity and inspiration	a. OBJECTIVE: The teacher does not present a lesson objective b. DEFINITION: The teacher does not (re)introduce the character strength c. RATIONALE: The teacher does not explain to students the importance of strong character
SOP-222: 2 The teacher will model the character strength for students	a. MODEL: Student(s) describe(s) or enact(s) a clear and inspiring vision of what the character strength looks like in action	a. MODEL: The teacher gives students a clear vision of what the character strength looks like in action	a. MODEL: The teacher gives students a vision of what the character strength looks like in action, but the vision lacks clarity	a. MODEL: The teacher gives students a vision of the character strength, but the vision lacks clarity	a. MODEL: The teacher does not give students a vision of the character strength
SOP-222: 3 The teacher will guide students in developing the character strength	a. PRACTICE: The teacher gives all students the chance to better understand AND build/exercise the character strength through effective practice	a. PRACTICE: The teacher gives all students the chance to better understand OR build/exercise the character strength through effective practice	a. PRACTICE: The teacher gives all or most students the chance to better understand OR build/exercise the character strength through practice, but the practice lacks efficacy	a. PRACTICE: The teacher gives some students the chance to better understand OR build/exercise the character strength through practice, but the practice lacks efficacy	a. PRACTICE: The teacher does not give students the chance to better understand OR build/practice exercising the character strength
SOP-222: 4 The teacher will reflect on the strengths and growth areas of his character-focused lesson(s)	a. REFLECTION: The teacher accurately, clearly, and compellingly describes strengths and growth areas in his character-focused instruction in the clip, making insightful connections between ideas from SOP character modules and his classroom	a. REFLECTION: The teacher accurately and clearly describes strengths and growth areas in his character-focused instruction in the clip, making connections between ideas from SOP character modules and his classroom	a. REFLECTION: The teacher describes strengths and growth areas in his character-focused instruction in the clip, but less accurately and/or clearly	a. REFLECTION: The teacher describes strengths and growth areas in his character-focused instruction in the clip, but inaccurately and/or with little clarity	a. REFLECTION: The teacher does not describe either strengths or growth areas in his character-focused instruction in the clip

Appendix I

Study Site GSE Instructional Sequence and Related Data

Over the course of their two-year program, each cohort of teachers in the study took several dozen "modules" (the GSE's term for the short courses that compose the MAT curriculum), designed to teach them about various instructional concepts and techniques. The teachers' work from several of these modules generated much of the data used in the study. For example, the teachers obtained, analyzed, and submitted their students' quarterly reading levels as part of their standard coursework. To better understand the modules and related data relevant to this study, the teachers' instructional experience is described below.

In the summer preceding their second year of the program, the teachers took "Introduction to Character Strengths," in which they learned about Peterson and Seligman's (2004) taxonomy of character strengths, as well as some of the many positive life outcomes associated with these strengths. As part of this module, teachers took the VIA survey and reported their top five character strengths, as well as positing a strength that may be relevant for their students' focus in the upcoming school year.

In the early fall of their second year in the program, the teachers took "SGA-201: Year 2 Pathway." This module taught teachers how to set year-long academic and character goals and measure students' beginning-of-year academic and character status. As part of this module, teachers assessed their students' beginning-of-year reading levels, using the reading assessments described in Chapter II, and beginning-of-year character, using the survey items designed by Park et al. (2017), also described in Chapter II and replicated in Appendix F. Teachers submitted these beginning-of-year data for their students to the GSE in November. As described above, these beginning-of-year data sets serve as predictor measures in this study.

In the late fall of their second year in the program, the teachers took "Teaching Character Strengths." This module introduced teachers to research-based approaches to CE and social-psychological interventions. As part of this module, each teacher planned, taught, and filmed a character-focused lesson in his classroom. Professors then reviewed these tapes and recorded each graduate student's score on a common, five-row, five-point rubric (see Appendix H). As described above, teachers' scores on these rubrics serve as predictors in this study.

Over the course of their two years in the program, all teachers had four formal classroom observations conducted by a GSE faculty member. Each of these observation modules was evaluative and focused on four different elements of instruction: classroom culture, planning, content, and relations. This study will focus on just one such observation module: OBS-201. Of the four observation modules teachers participate in over the two years of the program, OBS-201 is the observation module of focus for the study for two reasons: (1) it took place in the same term as the CE lesson and (2) like the CE lesson, it was assessed by observation with professors scoring those observations using a common rubric. The rubric scores teachers earned in OBS-201 serve as predictor measures, as described above.

Finally, in the late spring of their second year at the Study Site GSE each cohort took "Year 2 Outcomes," in which they recorded and submitted end-of-year student

reading GLEs and character ratings. As described above, these data—when converted to growth measures—serve as the outcomes for this study.

In sum, this study was designed to investigate the data associated with several aspects of the second year coursework in a two-year MAT program: from students' beginning-of-year reading and character levels, to the quality of teachers' efforts to promote both, and finally to the students' end-of-year reading and character outcomes.

Appendix J

Study Site Letter of Consent

███████████
Chief Research Officer
███████████ School of Education
███████████

Teachers College, Columbia University
Department of Curriculum and Teaching

July 21, 2017

To Whom It May Concern,

I write to grant my consent to Mayme Hostetter, a doctoral student in the Department of Curriculum and Teaching at Teachers College, to use de-identified academic, demographic, and interview data from ███████ alumni and their K-12 students for her Teachers College dissertation. This consent is, of course, contingent on Ms. Hostetter's successful completion of the Teachers College IRB process.

Sincerely,

[signature]

███████████
Chief Research Officer

Appendix K

Proposed Interview Participant Letter of Consent

Protocol Title: Cultivating Character in the Classroom

Interview Consent

Principal Investigator: Mayme Hostetter, maymehostetter@gmail.com, 917-655-1311

INTRODUCTION

You are being invited to participate in this research study called "Promoting the Social: Cultivating Character in Urban Public Charter Elementary School Classrooms." You may qualify to take part in this research study because you were an elementary school teacher working at a public charter school when you were enrolled at ▮▮▮ Graduate School of Education (▮▮▮ GSE). Approximately ten people will take part in the interview portion of this study, and it will take one hour of your time to complete.

WHY IS THIS STUDY BEING DONE?

The goals of this study are two-fold: (1) to better understand the character growth that students make in a school year and (2) to better understand how teachers think about and work towards cultivating this growth in the classrooms. If you choose to participate, your responses to a brief survey, your Teaching Character Strengths (SOP-222) video from your ▮▮▮ GSE coursework, reflection, and interview will inform (2).

WHAT WILL I BE ASKED TO DO IF I AGREE TO TAKE PART IN THIS

STUDY?

If you decide to participate, you will be asked to take a brief survey, review your SOP-222 lesson video and reflection (i.e., the submissions associated with the character lesson you taught and filmed during your time as a graduate student at ▮▮▮ GSE) and participate in a 30-minute interview with the researcher about character education, broadly, and your SOP-222 lesson, in particular. This interview will be recorded. The survey, video review, and interview can be completed anywhere that you prefer, so long as you have access to the internet. The exact date and time of the interview will be scheduled at your request.

WHAT POSSIBLE RISKS OR DISCOMFORTS CAN I EXPECT FROM TAKING PART IN THIS STUDY?

This is a minimal risk study, which means the harms or discomforts that you may experience are not greater than you would ordinarily encounter in daily life. **You do not**

have to answer any questions or divulge anything you don't want to talk about. You can stop participating in the study at any time without penalty. The principal investigator is taking precautions to keep your information confidential and prevent anyone from discovering or guessing your identity, such as using a pseudonym instead of your name and keeping all information on a password protected computer and locked in a file drawer.

WHAT POSSIBLE BENEFITS CAN I EXPECT FROM TAKING PART IN THIS STUDY?

There is no direct benefit to you for participating in this study. Participation may benefit the teaching profession more broadly, given the possibility that the research will catalyze the development of better preparation for new teachers at ▮▮▮▮ GSE and elsewhere, particularly with respect to cultivating social growth in K-12 classrooms.

WILL I BE PAID FOR BEING IN THIS STUDY?

You will not be paid to participate; however, the principal investigator will send participants a $25 Amazon, Barnes and Noble, or Starbucks (their choice) gift card as a token of appreciation for the hour they spend completing the pre-interview survey, (re)viewing their CE lesson video, and participating in the interview. There are no costs to you for taking part in this study.

WHEN IS THE STUDY OVER? CAN I LEAVE THE STUDY BEFORE IT ENDS?

The study is over when you have completed the survey, the video review, and the interview. Your participation will take approximately 60 minutes (i.e., 10 minutes completing a pre-interview survey, 20 minutes of video and reflection review, and 30 minutes of interview). However, you can leave the study at any time even if you haven't finished.

PROTECTION OF YOUR CONFIDENTIALITY

The investigator will keep all written materials locked in a desk drawer in a locked office. Any electronic or digital information (including audio recordings) will be stored on a computer that is password protected. What is on the audio-recording will be written down and the audio-recording will then be destroyed. There will be no record matching your real name with your pseudonym. The master list of student IDs that allows the principal investigator to contact you is kept separately from the principal investigator's research records. Regulations require that research data be kept for at least three years.

HOW WILL THE RESULTS BE USED?

The results of this study will be published in journals and presented at academic conferences. Your name or any identifying information about you will not be published. This study is being conducted as part of the dissertation of the principal investigator.

CONSENT FOR AUDIO AND OR VIDEO RECORDING

Video recording is part of this research study in that both your initial SOP-222 lesson and your interview have a video record. You can choose whether to give permission to be recorded in your interview. If you decide that you don't wish to be recorded, you will not be able to participate in this research study.

_____I give my consent to be recorded _____

                  Signature

_____I **do not** consent to be recorded _____

                  Signature

WHO MAY VIEW MY PARTICIPATION IN THIS STUDY

___I consent to allow written, video and/or audio taped materials viewed at an

educational setting or at a conference outside of Teachers College

_____

       Signature

___I **do not** consent to allow written, video and/or audio taped materials viewed outside

of Teachers College Columbia University _____

                  Signature

WHO CAN ANSWER MY QUESTIONS ABOUT THIS STUDY?

If you have any questions about taking part in this research study, you should contact the principal investigator, Mayme Hostetter at maymehostetter@gmail.com or 917-655-1311.
If you have questions or concerns about your rights as a research subject, you should contact the Institutional Review Board (IRB) (the human research ethics committee) at 212-678-4105 or email IRB@tc.edu**. Or you can write to the IRB at Teachers College, Columbia University, 525 W. 120th Street, New York, NY 1002. The IRB is the committee that oversees human research protection for Teachers College, Columbia University.**

PARTICIPANT'S RIGHTS

- I have read and discussed the informed consent with the researcher. I have had ample opportunity to ask questions about the purposes, procedures, risks and benefits regarding this research study.
- I understand that my participation is voluntary. I may refuse to participate or withdraw participation at any time without penalty to future employment; student status or grades; alumni status or opportunities; services that I would otherwise receive.
- The researcher may withdraw me from the research at his or her professional discretion.
- If, during the course of the study, significant new information that has been developed becomes available which may relate to my willingness to continue my participation, the investigator will provide this information to me.
- Any information derived from the research study that personally identifies me will not be voluntarily released or disclosed without my separate consent, except as specifically required by law.
- I should receive a copy of the Informed Consent document.

My signature means that I agree to participate in this study

Print name: _____ **Date**: _____

Signature: _____

Appendix L

Video-Coding Protocol

I used a video-coding protocol to determine the extent to which the teachers who were recruited and selected for interviews were employing the structures and strategies recommended by Seider (2012) and contemporary CE scholars, as well as aspects of recent social-psychological intervention work in their CE lessons.

I reviewed each video three times. The first time, I simply watched. The second time, I recorded—running-record-style—what was happening with both the teacher and the students for every three minutes of instruction. In the third viewing, I used a coding protocol to identify and categorize the teacher and student actions that were related to CE. In this third viewing, I paused the footage after each minute of instruction and used the set of codes below to summarize the CE-related features of the previous minute of instruction. For example, after viewing and coding a 20-minute video, I had 20+ codes describing the CE-related instructional structures and strategies the teacher employed during those 20 minutes. When appropriate, I supplemented a code with additional notes describing in more detail what was happening during that minute of instruction.

Code	Meaning
D/E-T	Teacher presenting definition, explanation, and / or example of character strength
D/E-S	Student(s) presenting definition, explanation, and / or example of character strength
Mod-T	Teacher modeling of character strength
Mod-S	Students(s) modeling of character strength
R/W	Students reading and / or writing character-related text
Int	Students interacting about character-related material (e.g., discussion)
Prac	Students practicing character-related habits or skills
Self-Soc	Students having opportunities to see the character strength in themselves, as in several of the social-psychological interventions
Other-Soc	Some other aspect of published social-psychological intervention
Other	Some other character-related instructional strategy or structure
N/A	Non-character-related (e.g., students passing out papers, lining up for class, etc.)

Appendix M

Pre-Interview Survey Protocol

Introduction:

Thank you for participating in this study! Before you re-watch the character lesson you planned and filmed while still a student at Study Site GSE, please take ten minutes to answer a few questions on the following survey. These answers will help to contextualize your thinking about character and character education within the larger vision of the study *and* provide important interview-shaping information for Mayme Hostetter, the primary investigator. Additionally, Mayme hopes that thinking "big picture" about character and character education before reviewing your own footage will help you see this footage again with the big picture in mind.

Survey:

Page 1

1. Name: [open response]
2. School at which you're currently teaching: [open response]
3. If you're not currently teaching, what do you do now? [open response]
   a. How many years ago did you leave teaching? [drop-down menu with a range from "less than 1" to "3"]
4. How many years had you been teaching when you filmed this lesson (during the fall of your second year at Study Site GSE): [drop-down menu with a range from "less than 1" to "more than 10"]
5. How many years have you taught in total? [drop-down menu with a range from "less than 1" to "more than 10"]

Page 2

Please state the extent to which you agree with the following statements: [five-point Likert scale from "Strongly Disagree" to "Strongly Agree"]

1. I am confident in my ability to be a good role model.
2. Teachers are usually not responsible when a child becomes more persistent.
3. When a student shows greater self-control, it is usually because teachers have effectively modeled that trait.
4. I am usually at a loss as to how to help a student be more persistent.
5. I know how to use strategies that might lead to positive changes in students' character.
6. I am not sure I can teach my students to have more self-control.

7. When students demonstrate grit, it is often because teachers have encouraged students to persist with tasks.
8. Teachers who spend time encouraging students to demonstrate more self-control will see little change in students' actions.
9. Some students will not become more persistent even if they have had teachers who promote this quality.
10. I often find it difficult to persuade a student that self-control is important.
11. When a student develops more persistence, it is usually because teachers have created classrooms where this strength is the norm.
12. I will be able to influence the character of students because I am a good role model.
13. Teaching students what it means to have self-control is unlikely to result in students who demonstrate more self-control.
14. I am continually finding better ways to develop the character of my students.

Page 3

Please state the extent to which you agree with the following statements: [five-point Likert scale from "Strongly Disagree" to "Strongly Agree"]

1. Character education works.
2. Character education helps students develop CHARACTER strengths and skills.
3. Character education helps students develop ACADEMIC strengths and skills.
4. It is a teacher's job to lead character education, in addition to academic endeavors, in his/her classroom.
5. All students possess character strengths.
6. Finally, is there anything else you'd like me to know before we meet for your interview? [open response]

Thank you! Please now take 20 minutes or so to review the footage of your lesson and to read your reflection on that lesson. I will look forward to speaking with you during your interview!

Appendix N

Interview Protocol

Introduction:

This interview will have two major foci. First, your big-picture vision of "character" and "character education" as a teacher: how you define those terms, how you see those concepts playing out in a classroom, and what you see as some of the goals or outcomes for that part of a teacher's work. And second, the particular, character-focused lesson you planned, taught, and filmed for SOP-222 while you were a graduate student. We'll start with the broad, big-picture questions and then move to the particulars of your SOP-222 lesson.

1. What's your name? Who and where were you teaching at the time you submitted this video and reflection? Are you still teaching? If so, who and where?[5]

2. People mean lots of different things when they use the word "character" (in the context of someone's personal qualities). We talked and read about this concept a lot in your second year at Study Site GSE, and I'm interested in your definition and vision.
   a. How would you define "character"? How would you describe "character" in a classroom context?
   b. How would you define "character education"?

3. Teachers are often interested in whether or not what they're doing in their classrooms with their students is "working"—whether reading instruction, for example, is helping students to become better readers.
   a. What does it mean for CE to "work"?
   b. How would you determine if it's "working"? How did / do you determine this in your classroom?

4. What—if any—is the relationship between CE and academic growth? How did / do you see this in your classroom?

5. Moving to your SOP-222 video, why did you choose the character strength you chose?

6. What did you notice about how you taught the lesson? Were there certain elements that seemed more important than others? Worked better than others? Why?
   a. How did you choose the definition of the strength you chose? Why did you choose to introduce it that way?
   b. Why did you choose to have kids model / model yourself?
   c. Why did you choose the type and number of practice opportunities you did?

---

[5] All identifying information will, of course, be removed from all published transcripts and analysis.

      d. At the end of the lesson, what do you think kids knew / were able to do that they may not have / been able to do at the beginning?

      e. How does the lesson as enacted compare with the lesson as you envisioned / planned? To what extent do you think this contributed to the efficacy of the lesson?

7. While SOP-222 focused on what you and your students were doing to develop a particular character strength in just one lesson, perhaps you and / or your school offered students other CE-related opportunities.

      a. When you think about your classroom that year, to what extent was this lesson the only CE versus one of several CE-related opportunities / structures?

      b. If SOP-222 was one of several CE-related opportunities / structures in your classroom, please describe what other CE-related work looked like. To what extent was SOP-222 representative of or different from that other CE?

      c. Similarly, school programming beyond your classroom may or may not have provided students further CE. For example, CE curriculum or coursework in particular grades / classes? CE-focused assemblies? When you think about your school that year, to what extent was SOP-222 the only CE versus one of several CE-related opportunities / structures?

      d. If SOP-222 was one of several CE-related opportunities / structures in your school, please describe what other school-wide CE-related work looked like. To what extent was SOP-222 representative of or different from that other school-wide CE?

      e. And finally, broadly speaking, how would you describe your school environment with respect to character and CE?

8. What else would you like to say about character and CE in your professional practice?

Appendix O

Sample Character Education Lesson Objectives

Many of the lessons began with the teacher and / or students reading aloud an objective. The lesson objectives were not an in-depth focus of the study because they were not associated with character growth in the regression analyses. For reference, though, the table below shows a sampling of lesson objectives:

Sample CE lesson objectives

Teacher	Grade	Strength	Objective
Ia	1	Self-Control	Describe examples of self-control.
Ic	3	Self-Control	Demonstrate self-control by using our check-list.
IIa	2	Grit	How does Wilma Rudolph show grit?
IIc	K	Self-Control	Show self-control by coloring neatly and waiting for what we want.
IIIb	1	Grit	Define and explain what it means to show grit.
IIIc	3	Self-Control	Show self-control by not dancing and build self-control by completing an "if…then…" chart.

Appendix P

Pre-Interview Survey Items and Response Data

Pre-interview survey items, source language, and sample and source response data

#	Domain (Coding)	Pre-interview survey item (with study edits *italicized*)	Milson & Mehlig (2001) language	Study M (SD) n = 6	Milson M (SD) n = 920+
1	PTE(+)	I am confident in my ability to be a good role model.	Same	4.67 (0.52)	4.68 (0.56)
2	GTE(-)	Teachers are usually not responsible when a child becomes more *persistent*.	*courteous*	4.50 (0.55)	4.03 (0.69)
3	GTE(+)	When a student shows greater *self-control*, it is usually because teachers have effectively modeled that trait.	*respect*	3.83 (0.98)	3.68 (0.79)
4	PTE(-)	I am usually at a loss as to how to help a student be more *persistent*.	*responsible*	3.83 (0.98)	4.23 (0.66)
5	PTE(+)	I know how to use strategies that might lead to positive changes in students' character.	Same	3.83 (1.17)	4.09 (0.66)
6	PTE(-)	I am not sure I can teach my students to *have more self-control*.	*be more honest*	4.50 (0.55)	3.76 (0.97)
7	GTE(+)	When students demonstrate *grit*, it is often because teachers have encouraged students to persist with tasks.	*diligence*	4.17 (0.41)	3.85 (0.76)
8	GTE(-)	Teachers who spend time encouraging students to *demonstrate more self-control* will see little change in students' actions.	*be respectful*	4.17 (0.75)	4.12 (0.75)
9	PTE(-)	I often find it difficult to persuade a student that *self-control* is important.	*respect for others*	4.50 (0.55)	3.78 (0.88)
10	GTE(+)	When a student *develops more persistence*, it is usually because teachers have created *classrooms where this strength is the norm*.	*becomes more compassionate / caring classroom environments*	4.17 (0.41)	3.73 (0.72)
11	PTE(+)	I will be able to influence the character of students because I am a good role model.	Same	4.17 (0.41)	4.08 (0.64)
12	GTE(-)	Teaching students what it means to *have self-control* is unlikely to result in students who *demonstrate more self-control*.	*be honest / are more honest*	4.00 (0.00)	3.77 (0.79)