



Recruiting and Preparing Quality Teachers

Citation

Mancenido, Zid. 2022. Recruiting and Preparing Quality Teachers. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37371999>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

HARVARD UNIVERSITY
Graduate School of Arts and Sciences



DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Department of Education
have examined a dissertation entitled

Recruiting and Preparing Quality Teachers

presented by

Zid Mancenido

candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

A handwritten signature in blue ink that reads "Heather C. Hill".

Signature

Typed name: Prof. Heather C. Hill

A handwritten signature in blue ink that reads "Eric Taylor".

Signature

Typed name: Prof. Eric Taylor

A handwritten signature in blue ink that reads "Susan Moore Johnson".

Signature

Typed name: Prof. Susan Moore Johnson

Signature _____

Typed name: Prof.

Signature _____

Typed name: Prof.

Date: April 15, 2022

Recruiting and Preparing Quality Teachers

A Dissertation presented by

Zid Mancenido

to

The Committee on Higher Degrees in Education

in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy in

the subject of

Education

Harvard University

Cambridge, Massachusetts

15 April 2022

© 2022 Zid Mancenido

All Rights Reserved.

Recruiting and Preparing Quality Teachers

Abstract

In Paper #1, I investigate the career exploration processes of 96 recent graduates of elite higher education institutions with varying levels of interest and commitment to teaching. My aim is to understand how intrinsic motivations to teach (positive) and extrinsic concerns about teaching (negative) are developed and interact such that some academic high achievers choose to teach while others do not. I find that motivations to teach were often developed through engaging in teaching-related activities, participating in education-related academic experiences, and reflecting on the impact of teachers. I also find that while extrinsic concerns about teacher pay, prestige, and career advancement were ubiquitous, what distinguished academic high achievers who chose to teach was that they held beliefs that specifically addressed their extrinsic concerns and that these beliefs were often developed through conversations with other academic high achievers who chose to teach or supported teaching as a career. These findings suggest the possibility of attracting more academic high achievers into teaching through policies that develop their intrinsic/altruistic motivations and/or foster beliefs that directly address prevalent extrinsic concerns.

Paper #2 is a narrative inquiry drawing on the same dataset as Paper #1. My aim is to examine how academic high achievers learn through explicit and implicit signals that teaching is not appropriate for someone like them. I show that the process is social, with parents and peers playing a significant role in shaping beliefs. These findings suggest that policy efforts to recruit more high achievers into teaching may benefit from focusing earlier in the career exploration pipeline.

Paper #3 is a systematic literature review of the research designs and measurement approaches of 166 causal evaluations of teacher preparation practices published in peer-reviewed journals between 2002-2019. Although I find that very few studies address issues of internal validity, external validity, and measurement, I highlight some innovative approaches as guidance for future researchers. I also present a common template for designing and reporting on causal evaluations of teacher preparation practices.

Table of Contents

Title page	i
Copyright	ii
Abstract	iii
Table of Contents	v
Acknowledgements	vi
Introduction	1
Paper #1: How Do You Become Motivated To Be A Teacher? A Case Study of Academic High Achievers	4
Paper #2: How High-Achievers Learn That They Should Not Become Teachers	42
Paper #3: Impact Evaluations of Teacher Preparation Practices: Challenges and Opportunities for More Rigorous Research	74
Conclusion	131
Appendix A: Example codes and indicative quotes for Paper #1	133
Appendix B: Example summary memos of interviews transcripts for Paper #1	139
Appendix C: List of Included Studies for Paper #3	144
Appendix D: Full Coding Scheme for Paper #3	157

Acknowledgements

Heather C. Hill
Eric Taylor
Susan Moore Johnson
Boyd I. R. Hampton
Jeza Ann Mancenido
Felicitas Mancenido
Rex Mario Mancenido

Introduction

Quality teachers matter. Quality teachers have significant and persistent effects on student achievement (Aaronson, Barrow, & Sander, 2007; Gordon, Kane, & Staiger, 2006; Nye, Konstantopoulos, & Hedges, 2004; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004), on school attendance and graduation (Gershenson, 2016; Jackson, 2018), on social-behavioural outcomes (Jackson, 2018; Kraft, 2019), and on later-life income (Chetty, Friedman, & Rockoff, 2014; Hanushek, 2011). So how do we get more quality teachers in K-12 classrooms?

In this dissertation, I focus on two pathways for improving teacher quality at scale: attracting more academic high achievers into the teaching profession; and identifying effective teacher preparation practices. In the following paragraphs, I explain my motivations for investigating each of these two pathways, and how the papers in my dissertation contribute to the existing research literature.

My motivation for investigating how to attract more academic high-achievers into teaching is three-fold. First, high-achievers are better able to raise their students' test scores (D'Agostino & Powers, 2009; Hanushek, Piopiunik, & Wiederhold, 2019; Henry, Bastian, & Smith, 2012; Jacob, et al., 2018; Meroni, Vera-Toscano, & Costa, 2015; cf. Harris & Sass, 2011). Second, high-achievers are better able to implement the sorts of disciplinarily rich and cognitively demanding instructional practices that are the aim of recent education reforms (Baumert et al., 2017; Blazar, 2015; Hill, Blazar, & Lynch, 2015; Cohen, McLaughlin, & Talbert, 1993; Spillane & Thompson, 1997). And third, recruiting more high-achievers into the profession can in turn attract even more high-achievers into teaching, given how talent begets talent in the general labor market (Binder, Davis, & Bloom, 2016; Ho, 2009; Rivera, 2015). In

summary, finding ways to get more high-achievers into teaching is important because it could raise overall student achievement, increase the likelihood of success of future educational reform, and improve the occupational status of the teaching profession.

While numerous studies have investigated what motivates individuals to become teachers and how this has changed over the past fifty years (e.g., Cochran-Smith et al., 2012; Johnson, 1990; Lortie, 1975; Nias, 1981; Ornstein, 1982; Sikes, 1985; Richardson, Karabenick, & Watt, 2014), no study has yet analysed the distinct motivations and concerns of academic high achievers as they explore and decide on teaching as a career. To fill this gap in the literature, I investigate the career exploration and decision-making processes of 96 recent college graduates/graduating seniors, from across four elite college campuses and four highly selective graduate-level teacher preparation programs. I present my results in the first two papers of this dissertation. In Paper 1, I undertake a comparative case study where I compare the career exploration and decision-making processes of those who chose to teach ($n=33$) with those who were interested in teaching but chose other careers ($n=32$) and those who were never interested ($n=31$). Given that I find in Paper 1 that social influences are particularly important in academic high achievers' career exploration and decision-making, in Paper 2 I undertake a narrative inquiry where I examine how academic high achievers learn through explicit and implicit signals that teaching is not appropriate for someone like them.

In Paper 3, I shift my attention to investigating how to better identify effective teacher preparation practices. My motivation stems from the numerous calls by teacher education researchers for more rigorous impact evaluations within the field (Wilson, Floden, Ferrini-Mundy, 2002; Grossman, 2008). Researchers making these calls have expressed concerns regarding a range of issues related to internal validity, measurement, and external validity. These

concerns include: the ubiquity of single-group case studies that simply describe correlation but not causation (Cochran-Smith & Zeichner, 2005; Fallon, 2006; Cochran-Smith, et al., 2015; Sleeter, 2014); the lack of replicable measurement approaches that appropriately mitigate participant and researcher bias (Grossman, 2008; Grossman & McDonald, 2008); and, the lack of standardized reporting on characteristics of the study sample and site to allow others to assess the generalisability of findings to other contexts (Cochran-Smith et al., 2012; Zeichner, 2005; Zientek et al., 2008).

To identify ways to overcome these issues of research design and measurement, I review 166 impact evaluations of teacher preparation practices published in peer-reviewed journals between 2002-2019. I present my findings in Paper 3, which includes highlighting innovative and promising approaches that can be emulated by future studies, and presenting a checklist of considerations to assist future researchers in designing and reporting on more rigorous impact evaluations.

Paper #1

How Do You Become Motivated To Be A Teacher?

A Case Study of Academic High Achievers

Abstract

I investigate the career exploration and decision-making processes of 96 academic high achievers with varying levels of interest and commitment to teaching. I find that individuals interested in teaching generally reported intrinsic/altruistic motivations, and that these motivations were generally developed through: engaging in teaching-related activities, participating in education-related academic experiences, and reflecting on the impact of teachers. I also find that while extrinsic concerns about teacher pay, prestige, and career advancement were ubiquitous, what distinguished academic high achievers who chose to teach was that they held plans and beliefs that specifically addressed their extrinsic concerns and that these plans and beliefs were often developed through conversations with other academic high achievers who chose to teach or supported teaching as a career. These findings suggest the possibility of attracting more academic high achievers into teaching through policies that develop their intrinsic/altruistic motivations and/or foster plans and beliefs that directly address prevalent extrinsic concerns. This case study contributes to the broader literature on teacher motivation by demonstrating how prospective teachers' motivations to teach can be influenced by broader structural factors.

Introduction

Nearly fifty years ago, Dan Lortie (1975) found that individuals were primarily motivated to become teachers because they loved to teach, they loved children, they loved their subjects, and/or they wanted to make a positive impact on the lives of others. These *intrinsic* and *altruistic* motivations to teach, however, were often outweighed or frustrated by concerns such as low relative pay, poor professional status, and limited career advancement. Lortie (1975) argued that these *extrinsic* concerns often led intrinsically or altruistically motivated individuals to leave the profession and/or abandon their aspirations to teach. Since Lortie's seminal study, a number of researchers have replicated and extended these findings, seeking to more deeply understand what is enduring and what is distinct about how individuals have decided whether or not to teach over time (e.g., Cochran-Smith et al., 2012; Johnson, 1990; Johnson & Birkeland, 2002; Nias, 1981; Ornstein, 1982; Sikes, 1985; Richardson, Karabenick, & Watt, 2014).

In this paper, I contribute to this literature by investigating how one highly sought-after group of individuals -- *academic high achievers* -- explore and decide upon teaching as a career. I focus on academic high achievers because although there remains ongoing controversy regarding what constitutes effective teaching given the diverse goals of schooling (Fenstermacher & Richardson, 2005), recent research shows that incoming teachers' academic achievement predicts gains in their future students' academic achievement (Henry, et al., 2012; Meroni, Vera-Toscano, & Costa, 2015; Rockoff, et al., 2011). I also focus on academic high achievers because although the academic achievement of incoming teachers in the US significantly declined between the 1950s and 1990s (Bacolod, 2007; Corcoran, Evans, & Schwab, 2004), there have been recent signs of improvement (Goldhaber & Walch, 2014; Hill, Lovison, Kelley-Kemple, 2019), suggesting that the career choices of academic high achievers

are sensitive to broader labor market conditions and that there is potential in policies aimed at attracting more of them into the profession (Nagler, Piopiunik, & West, 2020; Podolsky & Kini, 2016).

To investigate what distinctly attracts and deters academic high achievers from the teaching profession, I recruited 96 recent college graduates/graduating seniors from residential mailing lists at four elite college campuses, and the interested/admitted students mailing lists of four highly selective graduate-level teacher preparation programs. I interviewed each participant individually and then compared the career exploration and decision-making processes of those who committed to teaching (n=33); those who were interested in teaching but decided not to pursue it (n=32); and those who were never interested (n=31).

Consistent with prior research, I find that individuals interested in teaching generally reported intrinsic and/or altruistic motivations. In contrast to prior research, however, I find that for academic high achievers in my sample, extrinsic concerns about teacher pay, prestige, and career advancement took on a heightened significance, with many attributing this significance to anxieties associated with elite institutional cultures. I also find that while all participants in the study registered extrinsic concerns about teaching, only those who were committed to teaching reported specific plans and beliefs that addressed these concerns, such as conceiving teaching as a first step in a career, or rationalizing how much pay should matter in career decision-making. Finally, I find that the above motivations, plans and beliefs were often developed through social interaction and encouragement from mentors and role models, as well as through feeling success in teaching-related experiences and through reflecting on the impact of teachers. Unfortunately, however, most academic high achievers in my sample report limited opportunity to engage in these sorts of experiences, particularly given strong and early social pressures towards more

well-paid and prestigious careers.

These findings suggest that extrinsic concerns about teaching may not just be stopping academic high achievers already interested in teaching from deciding to pursue it as a career; rather, extrinsic concerns may also be inhibiting academic high achievers from developing intrinsic and altruistic motivations to teach in the first place. While focused specifically on the developing career motivations of academic high achievers, this case study contributes to the broader literature on teacher motivation by demonstrating how prospective teachers' motivations to teach can be influenced by broader structural factors (e.g., access to opportunities to explore teaching as a career, public discourses about teachers and the teaching profession). These findings also suggest the possibility of attracting more academic high achievers into teaching through policies that either develop their intrinsic and altruistic motivations and/or foster plans and beliefs that directly address prevalent extrinsic concerns.

Literature Review

As background to this study, I first review relevant research on prospective teacher motivation, identifying potential explanations for how motivations to teach may be developed. Then, I discuss relevant research on the distinct motivations of academic high achievers entering teaching, noting a gap in the literature regarding how academic high achievers committed to teaching learn to navigate extrinsic concerns about teaching. Finally, I briefly describe research on career exploration in elite educational contexts to identify the possibility that career funneling in elite educational contexts may be limiting the development of academic high achievers' intrinsic and altruistic motivations to teach.

What Attracts and Deters Individuals from Teaching?

There has been remarkable consistency in studies conducted across the US over the past

few decades demonstrating the importance of intrinsic and altruistic motivations for incoming teachers (Brookhart & Freeman, 1992; Guarino, Santibañez, & Daley, 2006; Rowsey & Ley, 1986). For example, Yarger, Howey, & Joyce (1977) surveyed over 2200 students across 175 teacher preparation programs in the US, finding that “desire to work with children” was the “overwhelming” reason why individuals chose teaching. Shipp (1999) surveyed 263 African American college students attending two different universities, finding that education majors ranked “contribution to society” and “intellectual stimulation” as the primary two reasons why teaching was attractive to them. Similarly, Chin & Young (2007) surveyed 1,862 teachers undertaking alternative certification programs in California, finding that the vast majority chose to teach for “service” and “personal and intellectual fulfillment”.

This is not to say that extrinsic motivations do not draw some individuals to teaching. Lortie (1975) described individuals -- particularly mid-career switchers -- who chose to teach for lifestyle reasons such as working hours that align with child care responsibilities (see also Castro & Bauml, 2009). However, extrinsic motivations are not common for the vast majority of new entrants to teaching: primarily recent college graduates who are more likely to register extrinsic concerns about teaching such as low relative pay, low prestige, and limited career opportunities (Johnson, 2004). For example, in their survey of 802 non-teaching college graduates, Farkas, Johnson & Foleno (2000) found that 78% of respondents believed teachers were seriously underpaid, 69% believed teachers did not have good opportunities for advancement, and 66% believed teachers do not get the sense they are respected and appreciated. Similarly, in their survey of high-achieving undergraduate students, Hiler & Hatalsky (2014) found education was ranked as the top profession that ‘average’ students choose, and only 35% of respondents described teachers as “smart.” These findings reflect the widely touted negative conceptions of

teaching in educational research and policy-making over the past few decades, conceptions exemplified by common adages like “those who can’t do, teach” and “teaching is a back-up option” (e.g., Holmes Group, 1986; Labaree, 2004).

Seen together, the research literature on prospective teacher motivation has established that intrinsic and altruistic are primarily what draw individuals into teaching; and extrinsic concerns about teaching are primarily what draw individuals away. As Johnson (2004) summarizes: “it would be a mistake to conclude that money is unimportant to teachers, even to those who are deeply committed to their work... [however,] Everyone knows that teachers do not make good money; they teach because it is good work that brings personal rewards, not financial ones.” (p. 50).

Recently, researchers have sought to understand how beginning teachers’ motivations change over time. Most of this work has focused on how motivations develop following the decision to become a teacher (e.g. Cochran-Smith et al., 2012; Geer, 1966; Johnson & Birkeland, 2003; Olsen, 2008). These studies have emphasized the critical importance of school contexts in retaining beginning teachers, noting that dashed hopes and feelings of failure can often negate new teachers’ motivations and lead them to move schools and/or leave teaching (Guarino, Santibañez, & Daley, 2006). Less investigated in the research literature, however, has been how motivations develop in the first place. That is, how do individuals develop various motivations to teach, and how does the way that they develop these motivations influence whether or not they eventually choose to become teachers?

Some researchers have described early experiences that have contributed to the development of beginning teachers’ motivations to teach. Lortie (1975) described the role of positive experiences with one’s own teachers as a way of stimulating some individuals’ early

interest in teaching as a career. Cole (1985) described how some teachers “drifted” into the profession after a number of positive experiences that are perceived as similar to teaching (e.g., babysitting, tutoring, or camp counselling). Johnson (1990) described how positive early experiences led some to choose teaching because “...it was not new or unfamiliar work, but a continuation of activities they had enjoyed since early adolescence” (p. 34). These findings in the teacher motivation literature align with what we know from broader research on adolescent career development. Career motivations are developed over time through a series of career learning experiences: career-related activities and social interactions that stimulate thinking about possible careers (Krumboltz, 2008; Perry, 1970). These experiences can be self-reinforcing, motivating individuals to seek further experiences; or they can be limiting, motivating individuals to stop exploring a particular career (Bandura et al., 2001; Harren, 1979).

In light of this existing literature, motivations to become a teacher may be developed in a similar way to how motivations to remain in teaching may be developed. That is, individuals’ motivations -- whether as a prospective teacher or a practicing teacher -- may increase when they are able to experience “a sense of success” in their work (Johnson & Birkeland, 2003) and/or feel part of a broader (pre-)professional community (Geer, 1966). If this is the case, then it suggests that access to experiences that help individuals consider a career in teaching may be critical to stimulating early motivations to teach. This may be particularly true for adolescents exploring future careers, given their educational and extracurricular opportunities are limited by contextual factors, such as school offerings, parental support, and local community context (Krumboltz, 2008). Furthermore, how individuals feel about early teaching-related career experiences may be critical to whether they may choose to foster their nascent interests in teaching, or pursue alternative careers. This may be particularly true for adolescents exploring various career paths

who often have very little information to base critical career and college decisions on, and therefore must rely more heavily on limited experiences provided by parents, mentors, and peers (Perry, 1970).

What Attracts and Deters *Academic High Achievers* from Teaching?

Studies investigating the distinct motivations of high achievers have primarily focused on the influence of extrinsic concerns like pay or entry costs. For example, Liu, Peske, & Johnson (2004) interviewed recipients of a large signing bonus in Massachusetts aimed at recruiting high achievers into teaching, finding that the money had little influence but the accompanying accelerated teacher certification pathway was highly valued. The researchers found that recipients were generally already intrinsically and altruistically motivated to become teachers, and that the Massachusetts policy simply accelerated their entry into the profession. Reback (2006) analyzed national survey data from the high school class of 1988, finding that individuals who attended undergraduate colleges without teacher certification programs were less likely to end up becoming teachers. Given highly selective colleges are less likely to have teacher certification programs, Reback concludes that “graduates from highly selective colleges are very sensitive to entry costs related to the number of years of schooling required for certification” (2006: p. 247). Finally, Coffman et al. (2019) studied the effects of randomly offering \$600 or \$1200 in financial assistance to those recently accepted into TFA. They find that the financial offer only increased the likelihood of joining TFA amongst applicants with the greatest financial need, and that without the offer they would likely have otherwise worked for private firms. Coffman et al. (2019) attributed these findings to lack of liquidity amongst high-need applicants: they were already motivated to teach, they just needed enough money to make the transition work financially.

More broadly, multiple studies have found that financial incentives (e.g., scholarships, loan forgiveness, changing labor market conditions) attract more academic high-achievers into teaching (Podolsky & Kini, 2016). For example, Henry, Bastian, & Smith (2012) evaluated a merit-based scholarship program (North Carolina Teaching Fellows) and found that in comparison to other teachers trained in-state, scholarship awardees had significantly higher academic qualifications, had higher value-added scores, and were more likely to stay in teaching (see also Feng & Sass, 2018; Steele, Murnane & Willett, 2010). Nagler, Piopiunik, & West (2020) show that more highly-skilled individuals choose to teach during recessions because of lower (expected) earnings in alternative occupations, suggesting that recent improvements in the academic skills (Goldhaber & Walch, 2014) and math knowledge of incoming teachers (Hill, Lovison, & Kelley-Kemple, 2019) may be “attributable to the 2008–9 financial crisis rather than an authentic reversal of long-term trend” (Nagler, Piopiunik, & West, 2020: p.458). Finally, Bacolod (2007) and Corcoran, Evans, & Schwab (2004) show how between the 1950s and 1990s high-achieving women increasingly gained access to (and therefore opted for) more highly paid and prestigious professions such as medicine and law over teaching.

Considered together, these studies affirm the findings discussed in the previous section: that extrinsic concerns regarding the low relative pay and prestige of teaching may be deterring some academic high achievers away from the profession, and that addressing these concerns may convince more intrinsically/altruistically motivated academic high achievers to choose it. Further investigation is required, however, to better understand: (a) how academic high achievers develop intrinsic/altruistic motivations towards and extrinsic concerns about teaching; and (b) how academic high achievers not presented with increased financial incentives have overcome extrinsic concerns regarding the low relative pay and prestige of teaching.

While not focused specifically on the case of teaching, research into career exploration in elite educational contexts offers some hypotheses for how academic high achievers' motivations may develop. Researchers have found that academic high achievers are generally funnelled early into high-prestige careers such as law, medicine, and engineering, often leaving little room for consideration of less prestigious career alternatives (Binder, Davis, & Bloom, 2016). This funnelling occurs due to strong explicit and implicit signals regarding what careers are socially appropriate, signals including parental expectations, availability of opportunities on elite college campuses, a sense of competitiveness and anxiety regarding "falling behind" one's peers (Deresiewicz, 2015). Career funnelling also occurs due to the numerous high-wealth high-status firms (usually from sectors such as finance, consulting, and technology) descending and actively recruiting on elite college campuses. These firms offer a range of material and symbolic benefits, such as large signing bonuses and pay, social status, and clear development pathways. Their outsized presence on elite campuses mean that for many academic high achievers, the choice to pursue a non-traditional career is distinctly one that involves numerous -- usually financial -- sacrifices (Ho, 2009; Rivera, 2015).

These findings from the career exploration literature suggest that academic high achievers interested in teaching may experience magnified extrinsic concerns (particularly those regarding pay, prestige, and career advancement) given the prevalence and allure of alternative career options and the social influences of parents, mentors, and peers. These findings also suggest that academic high achievers may have limited opportunities to explore teaching as a career given the career funnelling they experience in their elite educational contexts. While these findings do not offer insight into how academic high achievers may learn to address extrinsic concerns, they do suggest that structural factors may be inhibiting academic high achievers'

opportunity to develop intrinsic and altruistic motivations to teach in the first place. This is a challenge because it means that even if policies are developed to address academic high achievers' extrinsic concerns, they may have limited effects on overall teacher supply given the structural roadblocks that limit the development of intrinsic and altruistic motivations.

The Present Study

This study seeks *both* (1) to understand the distinct ways that one sought-after group of individuals -- academic high achievers -- currently explore and decide upon teaching as a career; and (2) to generate empirical evidence on how individuals develop intrinsic/altruistic motivations to teach, and how these motivations are weighed up against extrinsic concerns. My research questions are:

- What motivations distinguish academic high achievers who choose to become teachers?
- How do academic high achievers develop these motivations and/or ways to address extrinsic concerns?

Method

To answer these questions, I chose to conduct an in-depth qualitative comparative case study (Yin, 2018). I chose this approach because I wanted to understand how academic high achievers explore and decide on their career pathways from their own perspective, something only possible through in-depth qualitative study. I chose to pair this in-depth qualitative study with a comparative case approach because I wanted to identify what motivations and experiences distinguish academic high achievers who choose to teach, from those who were only interested in teaching and those who were never interested.

I recruited 96 recent college graduates/graduating seniors across three subsamples (cases): (1) those who self-report never having been interested in teaching as a career (n=31); (2) those who self-report having been interested in teaching as a career but chose not to pursue it following graduation, as defined by enrolling in a US teacher certification pathway or teaching in a private school (n=32); or (3) those who self-report choosing to pursue teaching following graduation (n=33). I recruited participants through the academic and residential mailing lists of four elite college campuses spread across the US, and through the interested/admitted students mailing lists of four highly selective graduate-level teacher preparation programs: two traditional one-year Masters programs; one two-year charter-based teacher residency program; and one national two-year alternative certification program¹. I chose to recruit through elite colleges and highly selective graduate-level teacher preparation programs because high academic achievement is a criterion for admission. I used interested/admitted students mailing lists of teacher preparation programs that target elite college graduates because I wanted participants with demonstrated interest in teaching as a career. I recruited participants using a maximum variation approach (Patton, 1990), purposefully seeking out and selecting participants from different demographics (gender, race/ethnicity, undergraduate major) rather than just interviewing anyone who initially expressed interest. I also recruited across two time periods (May-July, 2018 (n=41) and May-July, 2020 (n=55)) to guard against time-specific influences.

¹ Elite colleges and graduate-level teacher preparation programs were selected based on convenience through my own informal networks as a teacher educator and undergraduate college advisor.

Table 1.1 Demographics of study participants

<u>Age</u>		<u>Achievement Scores¹</u>		Mean
Mean	22.53	Undergrad GPA (n=93)		3.74
Range	21-37	SAT score (n=53)		
		-	Mean	2216.8
		-	Range	1860 - 2400
<u>Gender</u>		% (n)		
Male	36.5% (35)	ACT score (n=36)		
Female	59.4% (57)	-	Mean	33.2
Trans/Non-Binary/Other	4.2% (4)	-	Range	29 - 36
<u>Race</u>		% (n)	<u>Undergraduate Institution Rank</u>	% (n)
White	40.6% (39)			
Asian	20.8% (20)	Top 20 National		86.5% (83)
Black	16.7% (16)	Top 20 Liberal Arts		2.1% (2)
Latinx	10.4% (10)	Top 200 National		5.2% (5)
Mixed	9.4% (9)	Top 100 Regional		6.3% (6)
<u>Interest in teaching</u>		% (n)	<u>Undergraduate major</u>	% (n)
Never interested	32.3% (31)	Social Studies		32.3% (31)
Interested not committed	33.3% (32)	Humanities		20.8% (20)
Committed	34.4% (33)	STEM		45.8% (44)

¹ Not all participants chose to report achievement scores. Most who did not report could not remember their scores, some chose not to disclose for personal reasons.

Given this recruitment procedure, “academically high-achieving” in this paper refers

primarily to a distinct population of highly-selective elite college graduates (Table 1.1). Participants' average SAT score was 2211 (above the 98th percentile of all SAT test-takers); and 89.5% (n=86) graduated from a college ranked in the Top 20 National Universities or Top 20 Liberal Arts Colleges in the US News & World Report 2019 Best Colleges Rankings (2018). I acknowledge that this sampling procedure means my findings may not generalize beyond this distinct population (e.g., to academic high achievers who attend less selective institutions like large state schools). I look to future research to test the generalizability of my findings to these different contexts. That said, given my maximum variation sampling approach, I recruited quite a diverse sample from within this distinct population: 35.4% identified as male; 58.3% identified as non-white; and 50% majored in STEM. This is not representative of the broader US teacher labor force, where approximately 23% of teachers are male, 20% are non-white (Goldring, Taie, & Riddles, 2014), and there is a well-acknowledged shortage of qualified STEM teachers (Cowan, et al., 2016).

To collect data from participants, I conducted semi-structured interviews. These interviews lasted between 50-75 minutes and were conducted in person, via online video conferencing, or over the phone. As the aim of the study was to identify the key learning experiences that comprised individuals' career exploration and decision-making process, interviews were guided by a narrative inquiry approach (Carter, 1993; Tripp, 1994). In this approach, participants are asked to recall significant events, activities and social interactions in their lives and to make explicit how these experiences influenced their choices and actions. For this study, individuals were asked a single question: "Tell us the story of how you came to be doing what you are doing now career-wise. Where you start is completely up to you; as long as you can help us understand how your career interests started and how they have developed over time". Follow-up questions were then

asked chronologically, diving into various themes including: college experiences, including academics, extra-curricular activities, and internships; influences from parents, mentors, and peers; campus recruitment events and culture, etc. If teaching was not mentioned during the interview, participants were asked a final set of questions regarding any teaching-related experiences or career conversations. While the focus of the study was about teaching, I told participants that the study was about general career exploration and decision-making to ensure that I did not bias their recollections by indicating teaching-related hypotheses through specific questions. Interviews were audio-recorded and then transcribed verbatim.

I analysed data using a grounded theory approach. After the first round of participant recruitment, I selected a random sample of six transcripts and read each multiple times to identify key themes that were relevant to the research questions. These themes formed the basis of an initial coding scheme. Example codes and indicative quotes are presented in Appendix A. I then tested this scheme on four more randomly selected transcripts and amended it to ensure that codes were present for all relevant themes. I then coded all remaining transcripts using NVivo 12 to identify key trends across all participants and within subsamples. Following the coding of each transcript, I wrote short memos to summarize each participant's distinct decision-making process and the key experiences that led them becoming interested in or committed to teaching (or not). Example short memos are presented in Appendix B. I then aggregated all this data to identify key themes that differentiated each of the three subsamples (i.e. those *Not Interested*, those *Interested but not Committed*, and those *Committed*).

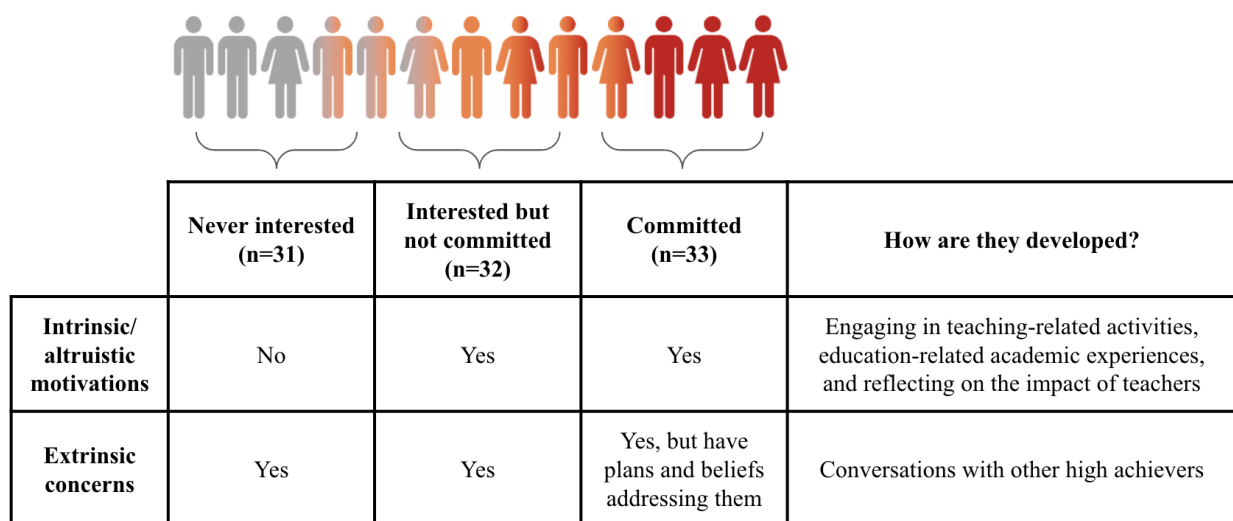
Results

In the following sections, I summarize participants' self-reported motivations to teach and

the experiences that led to the development of these motivations. I use the study’s two research questions to structure my findings. In section one, I first affirm that high achievers primarily report intrinsic and altruistic motivations when asked what motivated them to teach. I then show how, for the specific sample of high achievers in this study, extrinsic concerns about teaching take on a heightened significance given their elite educational contexts. Finally, I describe two ways in which high achievers committed to teaching address these extrinsic concerns.

In section two, I first describe three types of experiences through which high achievers report becoming motivated to teach, whether through developing intrinsic or altruistic motivations or learning ways to address their extrinsic concerns. Then I discuss how most high achievers report limited opportunities to avail of these sorts of experiences; instead, they are more likely to experience opportunities and social pressures that funnel them towards other more prestigious careers. Throughout these two sections, I compare the experiences of participants across the three subsamples to draw out the motivations and experiences that distinguish high achievers who choose to teach. A high-level summary of these findings is presented in Figure 1.1.

Figure 1.1. Summary of findings



What Motivates Academic High Achievers to Teach?

All participants in the study interested in teaching as a career attributed their interest to intrinsic or altruistic motivations. These motivations ranged from loving working with children or loving the craft of teaching through to feeling that teaching is meaningful work or wanting to have a positive impact on others. For example:

I really like working with people and I really like-- I love kids. I love seeing growth and I think the connections you make with kids are really, really special. (Alexandra)

... teaching is just fulfilling. I don't know, it's a little cliché, but like that feeling of sharing knowledge with student, like helping them enter the field, and watching students become better than you, stuff like that. (Jessica)

I realized that teaching is something that I'm really passionate about and that giving people information allows them to just spread the impact. You can never make more impact than you can as a teacher... it's the ultimate form of public service. (Naia)

While these intrinsic and altruistic motivations mirror those reported in previous research, participants generally framed their motives in ways specific to their elite educational contexts. For example, altruistic motivations were often cast within the frame of privilege, whether this was a feeling of wanting to pass on to others what has been given to them, or a drive to make the most of their own privilege by working for social justice:

The fact that I went to [elite college] has not been a reason for me to say that I don't want to teach, rather I want to use my education as an opportunity to bring more good and to help people think and achieve the things that they may not feel are possible. (Edwin)

I'm a privileged guy without really a care in the world in comparison to others. I think anyone in my position has to think very hard about your duties and obligations as a person in society given all the things that you've been given. (Nicholas)

Intrinsic motivations were often cast within the frame of doing what you loved rather than what was expected of a high achiever attending an elite college. Some participants contrasted their motivations with other students who were “selling out”; other participants

described turning towards teaching as a way to finally choose a life based on one's own interests.

For example:

My freshman year roommate double concentrated in business and computer science. She's like, "Yes, I'm 100% going to sell out." She already was pretty rich to start with... I didn't see the point in that. Like I definitely felt the pressure to sell out and go into something like consulting. But I thought about it for a while and then was like, "No, I want to feel good about what I do." (Paola)

I spent a lot of time at [elite college] making myself really miserable doing things that I thought that I should be doing or that others wanted me to do... I was like, "We're over that. I've been doing this for four years. It's time for me to do something that I want to do and do something that I think is actually meaningful rather than something that will look good on paper." (Deanna)

Regardless of participants' level of intrinsic or altruistic motivation towards teaching, extrinsic concerns loomed large. All high achievers in the study variously expressed concerns related to low pay, limited career prospects, and low prestige of teaching:

I come from Michigan which has just completely defunded public education-- teachers just don't get paid what they deserve to be paid. It's hard to make ends meet as a teacher if you have a family, particularly as a single breadwinner. I feel like a fancy degree from [elite college] is not going to protect you from that fate. (Nicholas)

In terms of the stuff I could be doing, the income levels I could be at in the future, I just know there are different careers or different jobs that I could take on now that would give me more opportunities in the future than teaching ever could. (Harrison)

I have a teammate who's going to be teaching next year in a high school, and I've definitely heard people say, "Oh, it's a waste of her [elite college] education." (Johanna)

As these quotes imply, concerns about pay, career, and prestige, were particularly heightened given others' and their own expectations of success as a high achiever. As Ryan explained:

There are both internal and external pressures on students to get high prestige jobs out of [elite college]. It's like getting a prestigious job is how you continue to prove that you're a high achiever, that you've worked hard in college, and are better than your peers.

Some participants in the study even noted concerns about whether teaching was ‘worth’ it given all the money, effort, and sacrifice it took in order to even get into an elite college. As Blair described:

You sacrifice a lot by attending a school like this-- it's like \$250,000. You have to consider if you're getting your worth for that education. I think that partially explains why people value prestige so much. Because like it's almost embarrassing if you went to [elite college], and you're doing the same job as someone that just went to a less prestigious college and didn't work as hard. It makes you wonder whether all of your work, even going back to high school was worth it.

Concerns over whether teaching was ‘worth it’ were most often voiced by participants who self-reported as being first-generation and/or low-income. Unlike some teachers from Lortie’s (1975) study almost 50 years ago, these high achievers did not conceive teaching as a comfortable passport into the middle-class nor a respectable outcome of an elite education. For example:

Feeling like you'll graduate [Elite University] and you're not, I guess, fulfilling it... is still difficult for me, to be completely honest. [After learning more about teaching] I was so excited, and I am so excited, but it also actually scared me a bit. It's like, "What if this is where I'm at? What if I'm a teacher forever?" That shouldn't be a bad thing, I'm aware of that, but low-income Rory is thinking... that's going to be difficult. (Rory)

... teachers are just greatly underappreciated by American society. I didn't want to be feeling like I came from a low-income background and went to an Ivy University and all I could become was a teacher. Like I could see through my peers how good it could get, how much I could do with my degree, how much I could accomplish. Being a teacher kind of felt like settling. (Faith)

Given the ubiquity of extrinsic concerns about teaching, it may appear surprising that there are high achievers who choose to teach. By comparing the motivations of the participants in my sample who did choose to teach with those from the other two subsamples, I found that only high achievers committed to teaching reported plans and beliefs that directly addressed extrinsic concerns about teaching. Two types were commonly reported. First, committed high achievers reported conceiving teaching as only the first step in their career, as something that

could lead them to other more prestigious opportunities. As Tabitha recounted:

I think talking to a lot of [people who started as teachers] influenced my decision in that I realized a lot of different pathways that you could take. There were some people who succeed in teaching and do Fulbright; some people stayed in the classroom that they were in, or went to another city and taught, or went to grad school, or went in another industry or stayed in education but did policy. It just seems something—that could set me up to be in the position, where I can do the work that I want to do in multiple capacities.

While never made explicit by respondents, conceiving teaching in this way suggests that teaching is only perceived as low prestige if one is planning to do it for the long term; having other more prestigious career plans post-teaching enables high achievers to feel more comfortable with their choices. This findings aligns with prior research acknowledging a generational shift in how most incoming teachers conceive of the prospect of classroom teaching—a shift away from teaching as “a calling and a lifelong commitment” and towards teaching as a short-term launching pad for other education and non-education related careers (Johnson, 2004). This generational shift aligns with broader research on new entrants to the labor market. As Hiler & Hidalsky (2014) find from their survey of high-achieving undergraduate students, “only one-third of undergraduates believe they will stay in their first job for ten years or more, and... 43% intend to leave their first job in less than five years”.

Many participants acknowledged that having post-teaching plans was particularly helpful in allaying the concerns of family and friends, something that made choosing teaching easier given others’ expectations of them as a high achiever. As Rafael admitted:

My father knows that I'm going to, at least I plan to do policy afterwards. Whenever he tells his friends about me, right now, what the conversation always looks like, he's so proud of me, he tells all his friends, “Yes, my son, he's going to [Elite University].” They say, “What's he doing?” Then his voice lowers, “He's getting his masters.” They're like, “Masters in what?” He's like, “Education. He's going to be a teacher, but he's got a plan. He's going to teach for 10 years and then go back to law school and be a lawyer.” It always comes back to that.

When participants were asked how they learned that teaching could be the first step of a broader career, the most common response was: by talking to other high achievers who had previously chosen teaching. While some conversations occurred informally through personal networks (e.g., friends of friends, friends of the family), others happened when teacher education programs provided structured opportunities for applicants to get to know their alumni and/or program staff. These conversations enabled high achievers to expand their understanding of teachers' various career prospects, as well as build a sense of connectedness with the teaching community as a prospective teacher. For example:

It was helpful meeting alumni of [TEP1] and definitely swayed me toward it. I don't know. I met really intelligent and outgoing and driven people that were alumni and I was like, "Wow, that's great. It seems like a program for the type of person that I want to be. What I'm trying to be." That was helpful, for sure. (Valerie)

There was one event, this was before I applied to [TEP2], and they actually brought in the regional director. It was an intimate setting. It was only seven undergraduate students and him, just having a conversation about educational equity and education reform. He talked about his experience in [teaching] and his career path... We were all just really impressed and motivated by him... it was the thing that motivated and pushed me to actually apply to [TEP2]. (Ryan)

The second most commonly reported way that high achievers addressed extrinsic concerns regarding teaching was by justifying that other career goals like meaning and enjoyment were more important to them than money and prestige. In this way, the intrinsic and altruistic aspects of teaching were conceived as 'worth' more. For example:

I'm obviously not doing teaching for money. I'm doing it so that a kid will understand something that they didn't know before doing it, so that a kid can see themselves growing intellectually doing it, so that a kid can go to college. (Kyle)

I'm not really that interested in money. I'm interested in having enough to be comfortable and have a few things, but I don't really care about making \$100,000 a year in my first year or anything like that... (Jessica)

When high achievers committed to teaching were asked how they learned to conceive their extrinsic concerns about teaching in this way, most reported that it was generally through conversations with other high achievers about the drawbacks of more conventional career paths. These conversations highlighted that the high pay and prestige of other careers often came with significant costs: whether in terms of long and gruelling hours of tedious work, or meaningless work that did not enable them to do something ‘that matters’. As Francis described:

At a career event I went to there was someone who had gone to Chicago and become a teacher. He told the story about how he was rooming with someone who went into investment banking or something like that. He said that one day his roommate came home, just looked at him, and asked, "What is it like to enjoy what you do?" I think about that a lot. I know that I probably won't be making as much money as a lot of the people around me, and that won't precisely be fun, but it's more important that I'd be doing something that I enjoy and that I can feel good about.

Finally, it is important to note that how participants rationalized teachers’ low relative pay depended on their family background. Some economically privileged participants reported feeling like money did not need to be a primary driver in their decision-making given their background; in contrast, some low-income participants reported needing to budget to evaluate whether the low teacher pay was enough to live on. As Isaac and Rory described respectively:

My family is-- we're right in the middle. We're healthily middle class and I think that gives me the ability to feel like I don't need to make a ton of money, and that money is not the primary thing I'm looking for in a job. (Isaac)

Honestly, I did a lot of budgeting. I looked at the salary for teaching and was like, "Can I live on this? Can I still save money? Can I still be a productive person? Will I have to cut corners constantly?" That was probably the hardest thing for me to wrestle with. (Rory)

In summary, by comparing common themes across participants in my three subsamples, I find that high achievers who chose to become teachers reported *both* intrinsic/altruistic motivations to teach *and* plans and beliefs that addressed extrinsic concerns about teaching.

These plans and beliefs were often developed through conversations with other high achievers who helped them see that teaching has multiple potential career paths, and/or that higher paid and more prestigious careers have significant drawbacks relative to teaching. In contrast, high achievers who were interested in teaching but chose other careers reported intrinsic/altruistic motivations to teach but did not report any plans and beliefs that addressed extrinsic concerns about teaching. Further, high achievers who never became interested in teaching did not report either intrinsic/altruistic motivations to teach or plans and beliefs that addressed extrinsic concerns about teaching.

How do Academic High Achievers Develop Motivations to Teach?

High achievers who reported being interested in teaching generally described three types of motivational experiences: (1) engaging in teaching-related activities; (2) perceiving the impact of teachers as a student; and, (3) undertaking education-related academic experiences. In this section, I explain each of these experiences, identifying potential critical features by contrasting the experiences of participants who reported interest in teaching vs participants who reported never becoming interested. I then explain the finding that most high achievers report difficulty accessing such motivational experiences. This is because of strong and early social pressures that funnel high achievers away from teaching and towards more prestigious careers.

The most common motivational experience reported was engaging in teaching-related activities (e.g. tutoring; volunteering for after-school programs; teaching in summer schools; coaching sports or teaching music). Through these activities, high achievers were able to reflect on their interests and realise their motivations to engage in teaching. For example, Brooke recounts:

I'm a camp counselor at a summer camp in Maine. It's a sleep-away camp, seven weeks, same girls. I've been there for three summers now, and each summer I'm with the same girls, so I get to watch them grow up, which is really cool, and I love it. In my second year, I had this realization more or less, that if I love doing that why not do it all year round, not just for seven weeks out the year. I like hanging out with kids and I like history, so why not do that for a living?

Similarly, Angela describes realizing her interests in teaching after volunteering at an after-school program:

I went in there with no expectations, having no idea. All my students were first and second generation Somali immigrants... I knew I had to make science personal. So I just tried to do that for them, to center their lives and their culture in it. I fell in love with it.

Critical to these personal experiences of teaching was observing a sense of success, which in turn fostered intrinsic motivations to teach. This sense of success also enabled some high achievers to overcome various challenges or difficulties associated with teaching, and to perceive teaching as “worth it”. These findings resonate with research that finds perceived self-efficacy predicts both occupational choice and persistence (Bandura, et al., 2001). The theory is that when someone feels that they are good at something (or that they could be) they are motivated to investigate it further and do more of it (Harren, 1979). As Francis demonstrates:

[Teaching summer school] was intense, it was tiring, but it was still so fun, just learning all that I did. I just wanted to learn more about teaching, I wanted to learn more strategies. And just at the end of the summer, realizing that although teaching was really challenging, although it was hard... I don't know, I was still positive and wanted to continue through.

In contrast, participants in the study who did not experience a sense of success in their teaching-related activities often reported magnified extrinsic concerns about teaching. Descriptions of these negative experiences were often followed by respondents lamenting that while they enjoyed some aspects of teaching, they felt that it was overall not worth the stress or that there were other better alternatives. For example, Kathryn was interested in becoming a

teacher until a negative experience volunteering at a school. She recounts:

So the students were all really great. But I just really did not enjoy a lot of the other aspects... Part of it was that it was the first time [the program] was happening but a lot of things could have gone better. I was just doing so much work, grading so late, getting up at 6 AM. I was like, "This is not the best thing in the world. No one gets paid enough to do this". I was just so disenchanted from that experience.

A few participants reported enjoying teaching-related activities but that this did not lead them to becoming interested in teaching. One common theme was that these participants received encouragement to channel their extra-curricular enjoyment towards non-teaching careers. For example, Alexandra describes receiving encouragement from others that made her consider tutoring as a pathway to a particular medical specialty, rather than teaching:

People would often be like, "You love kids. Do you want to go into pediatrics?" No one ever said like, "You love kids. Do you want to be a teacher?"

The second type of experience that high achievers reported as motivating them towards teaching was perceiving the impact of teachers on students' lives. These perceptions were often developed through reflecting on their own and their peers' experiences as students. As Nicholas describes:

I've had these teachers and peers believe in me, value my opinions, invest in my education, invest in my behavior, all these things... After a while I just realised that I wanted to teach because I had such a good experience at my high school.

A common theme amongst participants who reported this feeling was finding a teacher who they conceived of as a role model and who spoke positively about teaching as a potential career:

One of my favorite teachers had a PhD in mathematics from Caltech, and he was teaching at our public high school. Seeing how many students he was able to care for, mentor, and guide was pretty amazing. He was just so positive whenever he would talk about teaching, it's why I thought I might give teaching a shot right out of college. (Harrison)

... One year my brother had this teacher who just went out of her way to make sure that

he got all his [special education] accommodations. We talked a lot about her teaching, and it was very touching to me, and something that made me think, like, "Oh, I want to be able to do stuff like that for all students." I think that's a lot of where my initial motivation to teach came from, like the feeling of being able to support students who otherwise fall through the cracks. (Amiyah)

In contrast, many high achievers who reported never becoming interested in teaching described teachers who discouraged thinking about teaching as a career. These experiences often routed nascent interests, or consolidated perceptions that teaching was not an appropriate career for high achievers like them:

Growing up, a lot of people including teachers told me you shouldn't teach for lots of reasons. They were like, "No, don't do that. You're too smart for that," or teaching was like, I guess, getting more political. It was like, "Don't do teaching. It's swallowed in bureaucracy, now, you have to follow these rules. There's no freedom," and all this stuff. (Luisa)

I did definitely have teachers who even told me, you're too smart for that, go to grad school, get a Ph.D., become a professor if teaching is really what you want to do, go become a professor. (Lizete)

Finally, the third type of motivational experience reported by high achievers was education-related academic experiences (e.g. taking education-related courses in college; reading and discussing books about educational inequality). Critical to these experiences was intellectually engaging with issues in education, in ways that helped them perceive the importance and possibility of teaching as a career. For example, Jacqueline describes taking an international education course during a study abroad:

I'm like, "Oh my God, so much is wrong with the system." Initially, I didn't think about teaching, I thought, "I need to go into education policy. I need to just change the whole system-- That's how it needs to be." But my teacher in that class said something that really stuck with me, which was that the people making these policies, they've never been in a classroom-- they're not student-centered because they don't know the experiences of these students on a day-to-day basis. That's when I started thinking about teaching.

Jacqueline's experience is different to participants who took education courses but did not end up becoming interested in or committed to teaching. These participants often described coming away from the courses with the belief that working at a system-level may lead to greater impact, particularly given the systemic issues such as poor pay and prestige. As Theo explains:

I don't need to make a lot of money, but just learning about how poorly paid and treated teachers are made me definitely reassess whether teaching was the right place for me to make an impact. Like being a teacher is obviously very important and a big role to play in people's lives but taking that class made me wonder what's the replacement value of being a teacher versus, I don't know, if I started an organization that hired all these people to create more jobs, or worked in politics to change education policies. That is doing good for the world also, maybe that's a much bigger impact on our world.

In reflecting broadly on their career exploration trajectories, high achievers who committed to teaching often acknowledged that the three types of motivational experiences were self-reinforcing, that they compounded in ways that deepened their intrinsic and altruistic motivations to teach. In other words, high achievers who were committed to teaching often described themselves as self-selecting into teaching-related activities, or perceiving the impact of teachers, or education-related academic experiences, because they had previously enjoyed these experiences and wanted to do more. Committed high achievers also often acknowledged that engaging in motivational experiences helped them learn about future opportunities, making it easier for them to continue to develop their interests. This aligns with broader research on how positive career exploration experiences can be self-reinforcing, motivating individuals to become more interested in a career (Bandura et al., 2001).

In contrast, high achievers who did not choose teaching reported limited opportunities to engage in the above motivational experiences. This was often attributed to their elite educational backgrounds and contexts, which participants claimed funnelled them towards more prestigious careers in ways that never even allowed them to consider teaching as an option. For example,

many high achievers recalled early experiences of doing well in certain subjects in school, which led parents and peers to encourage them towards the relevant prestigious careers or towards “default” pathways for high achievers. Lori, who received consistently high grades in social science subjects in high school, recounts:

I think [my parents’] only understanding of what social sciences look like in real life was a lawyer and so they encouraged me to think about it as a possibility... and so definitely at that point in my life I thought law school would be the step for me.

Similarly, Yeo-Jin describes:

I think for me, pre-med was a bit of a default choice. [laughs] Like any other Asian parents, my parents were like, "It's a safe choice." [laughs] I didn't have any particular inklings in any direction whatsoever so I was like, "Yes. You know, I might as well go down this path." All of my friends were going to be pre-med. All of their parents were doctors. It seemed like it was the thing that you just did.

This funnelling continued through to college, where participants generally reported a lack of dedicated teaching-related career opportunities on elite campuses:

[Elite University] has very few undergraduate opportunities for understanding education. There are a million different groups and career fairs that you learn about BCG and McKinsey and [management consulting firms] like that, which I guess given how many students going in makes sense. But also, it's a self-fulfilling prophesy. (Kathryn)

Some of this funnelling away from teaching-related experiences was implicit rather than explicit. Participants often stated that they felt that being interested in teaching wasn’t appropriate given no one talked about it or that people who were lower-achieving were pursuing it; and/or, because other more prestigious careers like medicine and law were discussed with much more regularity and enthusiasm.

I feel as I got older, people in my life and in my community... if I pitched the idea of teaching they would respond one way. Not in a negative way but they would respond kind of maybe in a neutral way but then I pitched the idea of maybe wanting to go into the Peace Corps or be a lawyer, then their response was really enthusiastic. (Ashley)

I generally got the impression when I would go to internship or career fairs that—when

there were education-related opportunities—the people applying for these jobs, either they were few and no one was even like at the booths, or that the people who are applying were generally the people who are C students or otherwise not the kind of students that I necessarily thought would be the kinds of people I wanted to be working with. (Graham)

That there are explicit and implicit structures that funnel high achievers away from teaching is not to say that high achievers have no agency in their career choices. Rather, these findings suggest that there are structures and discourses embedded within elite educational contexts that exclude teaching from many high achievers' consideration set of potential careers. As Sean aptly summarises:

... at school I thought that the only career paths were being a doctor, being a lawyer, going into consulting, going into finance, going into tech, academia and maybe one or two other things. Every single other path just seemed out of the question. Not even out of the question, they just never presented themselves. I just didn't know they existed.

Discussion and Conclusion

Over fifty years ago, Dan Lortie (1975) described teaching as a “safety net” option for high achievers, where “people who cannot make it as doctors, business executives, or professors can avoid slipping into work of considerably less social standing” (p. 50). The findings of the present study suggest this is no longer the case. First, all academic high achievers in this study reported numerous explicit and implicit signals highlighting that teaching was of lower social standing and not appropriate for elite college graduates. Participants reported receiving these signals from an early age, often in ways that meant that teaching was never even considered as a possible career. Second, interested academic high achievers who chose to teach were not distinguished from other academic high achievers in this study because they had no other alternative career options. Rather, participants committed to teaching reported *both* intrinsically/altruistically motivations to teach *and* beliefs that directly addressed prevalent

extrinsic concerns about teaching, such as conceiving teaching as providing prestigious further career opportunities, and rationalising pay and prestige as less important in career decision-making than meaning and enjoyment.

These findings highlight the increased challenge of attracting academic high achievers to become teachers in today's context. However, it is not all bad news. This study has also shed light on the sorts of experiences and relevant critical features that may generate intrinsic and altruistic motivations to teach (e.g., engaging in teaching-related activities; reflecting on the impact of teachers; participating in education-related academic experiences). This study has also shed light on the types of interactions that may foster beliefs that directly address extrinsic concerns about teaching (e.g., conversations with other academic high achievers that reveal possible career paths that stem from teaching, and/or highlight the downsides of well-paid and prestigious careers).

One implication from these findings is that it may be possible to improve the academic achievement of incoming teachers through policies that focus on fostering academic high achievers' intrinsic and altruistic motivations to teach. For example, high-schools and elite colleges could offer more mentoring or tutoring programs that get students engaging and feeling a sense of success in teaching-related activities. These programs could be run by slightly older peers who can share their own interests in teaching as a career, which would in turn encourage younger students to reflect on their own career interests and the possibility of teaching. Elite colleges could offer more education-related classes that get students intellectually engaged with educational issues and reflecting on the fundamental role of teachers, experiences commonly reported by those in this study who became interested in teaching. Further, given the critical roles of mentors in the career exploration process of participants, teachers and professors could

do more to share their positive experiences of teaching and explicitly encourage academic high achievers to consider teaching as a career.

These efforts will require coordination from school districts and teacher preparation programs who must identify and recruit academic high achievers much earlier in their career exploration process. For example, school districts could develop “Grow Your Own” teacher preparation programs (Gist, Bianco, & Lynn, 2018) to identify young people with latent interests in teaching in high-achieving magnet schools and/or promote the possibility of teaching as a career as part of existing initiatives like Teacher Appreciation Day. Graduate-level teacher preparation programs could partner with elite college professors teaching introductory STEM courses to promote teaching as a possible career and/or facilitate volunteering in local high school tutoring or mentoring initiatives.

It may also be possible to convince more academic high achievers already interested to commit to teaching through policies that help them develop beliefs that directly address prevalent extrinsic concerns. For example, teacher preparation programs and school systems could facilitate more social connections between academic high achievers with emerging interests in teaching and academic high achievers already committed to the profession. School districts could offer summer internships or professional mentoring relationships that pair experienced teachers, principals, and/or senior district leaders with students from their alma maters with demonstrated interest in teaching. Teacher preparation programs could organize “back to school” days, where their alumni return to their undergraduate colleges and high schools to share personal experiences and reasons for choosing to teach.

As noted above, given the specific elite contexts of participants in this study, I cannot say whether my findings may generalize to other highly-sought after groups of prospective teachers,

such as individuals from racially minoritized backgrounds. I also cannot say whether there may be uninvestigated factors that lead some sorts of experiences and interactions to be more or less effective at generating motivations to teach if implemented in other contexts such as large public universities. Regardless of these limitations, the study findings suggest a general insight regarding policies for increasing teacher supply: that it may be possible to generate interest in teaching as a career by offering experiences focused on developing intrinsic or altruistic motivations.

References

- Bacolod, M. P. (2007). Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Quality. *The Review of Economics and Statistics*, 89(4), 737–751. <https://doi.org/10.1162/rest.89.4.737>
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-Efficacy Beliefs as Shapers of Children's Aspirations and Career Trajectories. *Child Development*, 72(1), 187–206. <https://doi.org/10.1111/1467-8624.00273>
- Binder, A. J., Davis, D. B., & Bloom, N. (2016). Career Funneling: How Elite Students Learn to Define and Desire “Prestigious” Jobs. *Sociology of Education*, 89(1), 20–39. <https://doi.org/10.1177/0038040715610883>
- Brookhart, S. M., & Freeman, D. J. (1992). Characteristics of Entering Teacher Candidates. *Review of Educational Research; Washington*, 62(1), 37.
- Carter, K. (1993). The Place of Story in the Study of Teaching and Teacher Education. *Educational Researcher*, 22(1), 5–18. <https://doi.org/10.3102/0013189X022001005>
- Castro, A. J., & Bauml, M. (2009). Why Now? Factors Associated with Choosing Teaching as a Second Career and Their Implications for Teacher Education Programs. *Teacher Education Quarterly*, 36(3), 113–126.
- Chin, E., & Young, J. W. (2007). A Person-Oriented Approach to Characterizing Beginning Teachers in Alternative Certification Programs. *Educational Researcher*, 36(2), 74–83. <https://doi.org/10.3102/0013189X07299192>
- Cochran-Smith, M., McQuillan, P., Mitchell, K., Terrell, D. G., Barnatt, J., D'Souza, L., Jong, C., Shakman, K., Lam, K., & Gleeson, A. M. (2012). A Longitudinal Study of Teaching Practice and Early Career Decisions: A Cautionary Tale. *American*

Educational Research Journal, 49(5), 844–880.

<https://doi.org/10.3102/0002831211431006>

Coffman, L. C., Conlon, J. J., Featherstone, C. R., & Kessler, J. B. (2019). Liquidity Affects Job Choice: Evidence from Teach for America. *The Quarterly Journal of Economics*, 134(4), 2203–2236. <https://doi.org/10.1093/qje/qjz018>

Cole, M. (1985). “The Tender Trap?” Commitment and Consciousness in Entrants to Teaching. In S. J. Ball & I. F. Goodson (Eds.), *Teachers’ Lives And Careers* (pp. 89–104). Falmer.

Corcoran, S. P., Evans, W. N., & Schwab, R. M. (2004). Changing Labor-Market Opportunities for Women and the Quality of Teachers, 1957-2000. *American Economic Review*, 94(2), 230–235. <https://doi.org/10.1257/0002828041301920>

Cowan, J., Goldhaber, D., Hayes, K., & Theobald, R. (2016). Missing Elements in the Discussion of Teacher Shortages. *Educational Researcher*, 45(8), 460–462. <https://doi.org/10.3102/0013189X16679145>

Deresiewicz, W. (2015). *Excellent Sheep: The Miseducation of the American Elite and the Way to a Meaningful Life*. Simon & Schuster.

Farkas, S., Johnson, J., & Foleno, T. (2000). *A Sense of Calling: Who Teaches and Why*. Public Agenda. <https://eric.ed.gov/?id=ED443815>

Feng, L., & Sass, T. R. (2018). The Impact of Incentives to Recruit and Retain Teachers in “Hard-to-Staff” Subjects. *Journal of Policy Analysis and Management*, 37(1), 112–135. <https://doi.org/10.1002/pam.22037>

Fenstermacher, G. D., & Richardson, V. (2005). On Making Determinations of Quality in Teaching. *Teachers College Record*, 107(1), 186–213.

- Geer, B. (1966). Occupational Commitment and the Teaching Profession. *The School Review*, 74(1), 31–47.
- Gist, C. D., Bianco, M., & Lynn, M. (2019). Examining Grow Your Own Programs Across the Teacher Development Continuum: Mining Research on Teachers of Color and Nontraditional Educator Pipelines. *Journal of Teacher Education*, 70(1), 13–25.
<https://doi.org/10.1177/0022487118787504>
- Goldhaber, D., & Walch, J. (2014). Gains in Teacher Quality. *Education Next*, 14(1), 38–45.
- Goldring, R., Taie, S., & Riddles, M. (2014). Teacher Attrition and Mobility: Results from the 2012-13 Teacher Follow-Up Survey. First Look. NCES 2014-077. In *National Center for Education Statistics*. National Center for Education Statistics.
<https://eric.ed.gov/?id=ED546773>
- Guarino, C. M., Santibañez, L., & Daley, G. A. (2006). Teacher Recruitment and Retention: A Review of the Recent Empirical Literature. *Review of Educational Research*, 76(2), 173–208. <https://doi.org/10.3102/00346543076002173>
- Harren, V. A. (1979). A model of career decision making for college students. *Journal of Vocational Behavior*, 14(2), 119–133. [https://doi.org/10.1016/0001-8791\(79\)90065-4](https://doi.org/10.1016/0001-8791(79)90065-4)
- Henry, G. T., Bastian, K. C., & Smith, A. A. (2012). Scholarships to Recruit the “Best and Brightest” Into Teaching: Who Is Recruited, Where Do They Teach, How Effective Are They, and How Long Do They Stay? *Educational Researcher*, 41(3), 83–92.
<https://doi.org/10.3102/0013189X12437202>
- Hiler, T., & Hatalsky, L. (2014). *Teaching: The Next Generation*. The Third Way.
- Hill, H. C., Lovison, V., & Kelley-Kemple, T. (2019). Mathematics Teacher and Curriculum Quality, 2005 and 2016. *AERA Open*, 5(4), 2332858419880521.

<https://doi.org/10.1177/2332858419880521>

Ho, K. (2009). *Liquidated: An Ethnography of Wall Street*. Duke University Press.

Johnson, S. M. (1990). *Teachers at work: Achieving excellence in our schools*. Basic Books.

<http://hdl.handle.net/2027/mdp.39015026896608>

Johnson, S. M. (2004). *Finders and keepers: Helping new teachers survive and thrive in our schools* (1st ed.). Jossey-Bass.

Johnson, S. M., & Birkeland, S. E. (2003). Pursuing a “Sense of Success”: New Teachers Explain Their Career Decisions. *American Educational Research Journal*, 40(3), 581–617. <https://doi.org/10.3102/00028312040003581>

Krumboltz, J. D. (2009). The Happenstance Learning Theory. *Journal of Career Assessment*, 17(2), 135–154. <https://doi.org/10.1177/1069072708328861>

Labaree, D. F. (2004). *The Trouble with Ed Schools*. Yale University Press; JSTOR. <https://www.jstor.org/stable/j.ctt1njmnb>

Liu, E., Johnson, S. M., & Peske, H. G. (2004). New Teachers and the Massachusetts Signing Bonus: The Limits of Inducements. *Educational Evaluation and Policy Analysis*, 26(3), 217–236. <https://doi.org/10.3102/01623737026003217>

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

Meroni, E. C., Vera-Toscano, E., & Costa, P. (2015). Can low skill teachers make good students? Empirical evidence from PIAAC and PISA. *Journal of Policy Modeling*, 37(2), 308–323. <https://doi.org/10.1016/j.jpolmod.2015.02.006>

Nagler, M., Piopiunik, M., & West, M. R. (2020). Weak Markets, Strong Teachers: Recession at Career Start and Teacher Effectiveness. *Journal of Labor Economics*, 38(2), 453–500. <https://doi.org/10.1086/705883>

- Nias, J. (1981). 'Commitment' and Motivation in Primary School Teachers. *Educational Review*, 33(3), 182–190. <https://doi.org/10.1080/0013191810330302>
- Olsen, B. (2008). How “Reasons for Entry into the Profession” Illuminate Teacher Identity Development. *Teacher Education Quarterly*, 35(3), 23–40.
- Ornstein, A. C. (1982). Motivations for Teaching. *The High School Journal*, 66(2), 110–116. JSTOR.
- Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods* (2nd ed., p. 532). Sage Publications, Inc.
- Perry, W. G. (1970). *Forms of Ethical and Intellectual Development in the College Years: A Scheme*. Holt, Rinehart and Winston.
- Podolsky, A., & Kini, T. (2016). How Effective Are Loan Forgiveness and Service Scholarships for Recruiting Teachers? Policy Brief. In *Learning Policy Institute*. Learning Policy Institute. <https://eric.ed.gov/?id=ED606351>
- Reback, R. (2006). Entry Costs and the Supply of Public School Teachers. *Education Finance and Policy*, 1(2), 247–265. <https://doi.org/10.1162/edfp.2006.1.2.247>
- Richardson, P. W., Karabenick, S. A., & Watt, H. M. G. (2014). *Teacher Motivation: Theory and Practice*. Routledge.
- Rivera, L. (2015). *Pedigree: Elite Reproduction in Hiring*. Princeton University Press.
- Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can You Recognize an Effective Teacher When You Recruit One? *Education Finance and Policy*, 6(1), 43–74. https://doi.org/10.1162/EDFP_a_00022
- Rowsey, R. E., & Ley, T. C. (1986). Perceptions of Teachers' Salaries and Non-salary Benefits. *Journal of Teacher Education*, 37(2), 42–45.

<https://doi.org/10.1177/002248718603700208>

- Shipp, V. H. (1999). Factors Influencing the Career Choices of African American Collegians: Implications for Minority Teacher Recruitment. *The Journal of Negro Education*, 68(3), 343–351. <https://doi.org/10.2307/2668106>
- Sikes, P. J. (2005). The Life Cycle of the Teacher. In S. J. Ball & I. F. Goodson (Eds.), *Teachers' Lives and Careers*. The Falmer Press.
- Steele, J. L., Murnane, R. J., & Willett, J. B. (2010). Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from California. *Journal of Policy Analysis and Management*, 29(3), 451–478. <https://doi.org/10.1002/pam.20505>
- The Holmes Group. (1986). *Tomorrow's Teachers: A Report of The Holmes Group*. The Holmes Group, Inc. <https://eric.ed.gov/?id=ED270454>
- Tripp, D. (1994). Teachers' lives, critical incidents, and professional practice. *International Journal of Qualitative Studies in Education*, 7(1), 65–76. <https://doi.org/10.1080/0951839940070105>
- Yarger, S. J., Howey, K., & Joyce, B. (1977). Reflections on Preservice Preparation: Impressions from the National Survey. *Journal of Teacher Education*, 28(6), 34–37. <https://doi.org/10.1177/002248717702800607>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (Sixth edition.). SAGE Publications, Inc.

Paper #2

How High Achievers Learn That They Should Not Become Teachers

Abstract

In this article, Mancenido uses portraiture to examine how high-achieving students are socialized to believe that they should not become K-12 classroom teachers. Research has well-established that high-achievers are generally disinterested in teaching as a career, however there has been little attention to how this disinterest is developed through the process of career exploration. To address this gap in the literature, Mancenido draws on interviews with 96 high-achieving recent college graduates/graduating seniors. He presents six representative portraits to demonstrate how high-achievers learn – through manifold explicit and implicit signals – that teaching is not appropriate for someone like them. This process is heavily social, with parents and peers playing a significant role in shaping high-achievers' perspectives.

How High Achievers Learn That They Should Not Become Teachers

Feeling like you'll graduate [Elite University] and you're not, I guess, fulfilling it... is still difficult for me, to be completely honest. [After learning more about teaching] I was so excited, and I am so excited, but it also actually scared me a bit. It's like, what if this is where I'm at? What if I'm a teacher forever? That shouldn't be a bad thing, I'm aware of that, but low-income Rory is thinking... that's going to be difficult.

When Rory tells me this over the phone, I'm walking through a quadrangle enclosed by centuries-old buildings made of ivy-covered red-brick walls². I am walking slowly, ambling as we talk, thinking more about Rory's words than my next steps. Those who pass me by are students like Rory -- undergraduates shuffling from their dormitories to various classes and libraries. There is a general excitement mixed with anxiety in the air, a common feeling that comes with the fast-approaching end of Spring semester. I know Rory is in a similar quadrangle to me a few hundred miles away; a quadrangle also lined by ivy-covered red-brick student dormitories, a quadrangle also filled with tourists and tour groups who visit in the hopes of learning some of the secrets behind one of the world's most selective undergraduate institutions.

"Where do you think it comes from, this feeling that becoming a teacher isn't enough?" I ask. I hear Rory sigh. "A lot of places. Last week was the graduating seniors' dinner for [an extracurricular club]. We were all getting drinks and another senior was like, 'you're doing [teacher preparation program]? You're just going to be a teacher?' She just ended it there and I felt like I needed to say that no, I'll be a principal, I'll be a superintendent, I'll rise above blah blah blah. But it's like, no, you're right. I'm going to just be a teacher. It shouldn't be like that, but it is what it is. It's something I'm still wrestling with."

² Names have been changed, institutions have been anonymized, and any identifying features, such as club names, place names, have all been anonymized to protect the confidentiality of the participants.

This paper is about how high-achieving undergraduate college students like Rory are socialized to think and feel like they should not choose to become classroom teachers. It is about how high achievers, specifically those who attend elite highly-selective universities, learn to believe that teaching is not an appropriate career choice; that choosing to be a teacher would be ‘settling’; that it would be ‘beneath’ them; that it is a ‘waste of potential’. Negative conceptions of teaching amongst high-achievers in the US have been widely touted in educational research and policy-making over the past few decades, something exemplified by common adages like “those who can’t do, teach” and “teaching is a back-up option” (e.g. Darling-Hammond, 1984; Holmes Group, 1986; Labaree, 2004). For example, in his seminal sociological study of teachers in the 1970s, Dan Lortie (1975) describes teaching as a “safety net” option for high achievers, where “people who cannot make it as doctors, business executives, or professors can avoid slipping into work of considerably less social standing” (p. 50). Multiple nationally representative surveys attest to these negative conceptions: the occupational prestige of teaching has consistently been reported as relatively low in comparison to other professions requiring a college degree (Smith et al., 2019).

Despite poor conceptions of teaching being well-known and long-standing, there has been surprisingly little attention to how they are *socially constructed*. That is, where does the idea that it is not right for high achievers to become teachers come from? When do high achievers first learn about it? How does it manifest? How is it perpetrated? In what contexts? Research has generally side-stepped these micro-level questions, focusing more on the macro-level, i.e. how certain social and material conditions of teaching lead to poor perceptions. For example, many researchers have discussed why teaching is not perceived as a highly-skilled profession. Reasons include: that teachers are not experts in their disciplinary fields; that

teaching involves giving away one's expertise such that students do not need them; that the concept of good teaching is contested given the diverse ways that individuals prefer to learn; that teachers' impact on their students is hard to isolate, particularly because it is contingent on their students' cooperation; that teachers are caught 'in-between' the world of children and adults, leading many to perceive their work as lower status; and, that one's 'apprenticeship of observation' through 13 years of schooling generally leads individuals to underestimate the work of teaching (e.g. Cohen, 2011; Etzioni, 1969; Hoyle, 2001; Kennedy, 2006; Labaree, 2004; Lortie 1975). The effects of these social conditions on perceptions of the teaching profession are mutually reinforced by its poor material conditions – i.e. its poor relative pay, poor relative working conditions, poor preparation, etc. (e.g. Auguste, Kihn, Miller, 2010; Barber & Mourshed, 2007; Farkas, Johnson & Foleno, 2000; Rowsey & Ley, 1986; Sutchter, Darling-Hammond, & Carver-Thomas, 2016). There is no doubt that these social and material conditions are at the root of negative conceptions of teaching amongst high-achievers. However, acknowledging this at the macro-level sheds little insight into how conceptions are developed at the micro-level. For example, how do high-achievers learn about these poor social and material conditions of teaching? How do they develop negative conceptions about teaching? Where and/or whom do they learn all of this from? Which comes first?

There is some emerging sociological research into how students are socialized in elite institutions that offer some suggestive answers. This research identifies high achievers as committed yet highly impressionable (Binder, Davis, & Bloom, 2016; Schneider & Stevenson, 1999), using cues from their learning environments to construct an understanding of what are "socially-appropriate" (read: "prestigious") career paths. While some of these cues are explicit (e.g. family and friends encouraging high achievers to pursue certain careers) mostly are implicit

(e.g. who gets reified on elite campuses through alumni spotlights, public events, and major awards; the availability of information about certain career on campus, including which firms are present at campus career events and/or which industries are more prominent on campus). Together these cues lead to a sort of “career funneling” where high-achievers are streamed towards a narrow set of conventional pathways like law, medicine, and financial services (Binder, Davis, & Bloom, 2016; Ho, 2009; Rivera, 2015).

In this paper, I take a look at these social processes within the specific case of teaching. I do this not just out of conceptual interest but also because of the policy implications. We know that poor conceptions of teaching contribute to why high achievers choose not to become teachers (Auguste, Kihn, Miller, 2010; Farkas, Johnson, & Foleno, 2000), but until we can identify the sites, experiences, and interactions that create and maintain these conceptions, we will continue to struggle to identify concrete actions that may disrupt them (Lawrence, Suddaby, & Leca, 2011; Battilana, Leca, & Boxenbaum, 2009). This is born out in contemporary policy discussions. When discussing the negative perceptions of the teaching profession, many US policymakers often tout simple policy platitudes (e.g. “we need to raise the status of the teaching profession”; “increase public respect for teachers”), and/or simply recommend that we first need to improve teachers’ material conditions like pay and working conditions (e.g. Duncan, 2011; Sutchter, Darling-Hammond, & Carver-Thomas, 2016; Weingarten, 2019). To step beyond and develop more specific and targeted interventions that can subvert negative conceptions of teaching, we first need to understand how these conceptions are actually developed and fostered; we need to first understand how high achievers learn that they should not become teachers

To do this, I draw on a subset of interview data from a broader research project on the career exploration and decision-making processes of 96 high-achieving recent college

graduates/graduating seniors with varying levels of interest in and commitment to teaching. I use the method of *portraiture* (Lawrence-Lightfoot & Davis, 1997) to identify and present five portraits representative of the range of the broader sample's experiences. Read together, these portraits reveal how high-achievers learn and have reinforced negative conceptions of teaching all throughout their career exploration and decision-making process. High-achievers learn it explicitly from family and friends, who question and reroute any interests in teaching; they learn it implicitly from peers and mentors, who deprioritize and eschew teaching as a long-term career; and they learn it consistently throughout their lives, as they navigate others' expectations that they pursue other more prestigious career paths appropriate for high-achievers.

These portraits are not meant to exhaustively sketch how high-achievers learn about teaching as a career. Nor is the aim of these portraits to causally link negative conceptions of teaching to high-achievers' unwillingness to become teachers. Rather, in this paper, I simply highlight and elevate various ways that high-achievers learn about and grapple with negative conceptions of teaching. My hope is that educational researchers may use this work to extend existing conceptual models for why high-achievers may choose not to become teachers; and that policymakers may use this work to learn more deeply about the struggles, aspirations, and preoccupations of the target demographic that they seek to recruit into the teaching profession. I also hope that educators and prospective teachers who read these portraits feel seen in the experiences of my participants – that through these portraits they are given a voice and a vocabulary to express feelings of uncertainty and confusion that permeated throughout their own career exploration and decision-making processes. Overall, my theory of action is that by nuancing our conceptual and empirical understanding of high-achievers' career exploration and decision-making can help us better target and improve the effectiveness of policies aimed at

convincing more of our best and brightest into teaching.

Method

The goal of this study is to provide an empirical grounding for the development of theory (Glaser and Strauss, 1967) regarding when, where, and how high achievers are socialized to become disinterested in teaching. Given the challenge of needing to *both* capture the idiosyncrasy of an individual's process of socialization *and* to draw generalizations from across diverse experiences, I use portraiture, a qualitative social science methodology that “seeks to document and illuminate the complexity and detail of a unique experience or place” in order to “discover resonant universal themes” (Lawrence-Lightfoot & Davis, 1997: p. 14). Key to theory building through portraiture is *identification*, the use of aesthetic features like voice and storytelling to evoke “the energizing shock of sympathy and of human community” (Featherstone, 1989: 376). In this way, the aim of portraiture is to “make companionable for you a person who is identified or unknown” (Schjeldahl, 2019), such that “the familiar scene outside, the road and the bustle, often look transformed and transfigured” (Gombrich, 1969: p. 305 cited in Lawrence-Lightfoot & Davis, 1997: p. 35). I also choose to use portraiture because of its commitment towards *goodness*, a commitment that eschews conceptual negativism in the hope of drawing novel insights about its focal phenomenon. As Lawrence-Lightfoot and Davis (1997) write: “The researcher who asks first ‘what is good here?’ is likely to absorb a very different reality than the one who is on a mission to discover the sources of failure” (p. 9). I make this choice because, as flagged above, existing research into why individuals choose against teaching has focused on systematically “discovering the sources of failure” (Lawrence-Lightfoot & Davis, 1997: p. 9), i.e. what teaching *lacks*, such as pay and prestige. In holding this deficit focus,

research has too quickly rendered high achievers as overly deterministic and narrowly focused on the negative aspects of teaching. I choose portraiture to step back from and problematize these deficit-based assumptions.

Data for this study is drawn from a broader project on career exploration and decision-making. 96 recent college graduates/graduating seniors were recruited through the academic and residential mailing lists of four elite college campuses and the interested/admitted students mailing lists of four highly selective teacher preparation programs. I recruited participants using a maximum variation approach (Patton, 1990), purposefully seeking out and selecting participants from different demographics (gender, race/ethnicity, undergraduate major) rather than just interviewing anyone who initially expressed interest. I also recruited across two time periods (May-July, 2018 (n=41) and May-July, 2020 (n=55)) to guard against time-specific influences. My hope was that through this diverse sampling methodology, I would be better able to collect and represent the many different ways that high-achievers learn about teaching. This led to a final sample of which 35.4% of participants identified as male; 58.3% identified as non-white; and 50% majored in STEM. This is not representative of the broader US teacher labor force, where approximately 23% of teachers are male, 20% are non-white (Goldring, Taie, & Riddles, 2014), and there is a well-acknowledged shortage of qualified STEM teachers (Cowan, et al., 2016).

Given this recruitment procedure, “high-achieving” in this study refers to a distinct population of highly-educated elite college graduates. 89.5% (n=86) graduated from a college ranked in the Top 20 National Universities or Top 20 Liberal Arts Colleges in the US News & World Report 2019 Best Colleges Rankings (2018)). Participants’ average SAT score was 2211, which is above the 98th percentile of all SAT test-takers. For comparison, only around 10% of

first-year teachers starting in the 2011-12 school year came from selective higher education institutions (Ingersoll, Merrill, Stuckey, 2014); and the average first-year teacher starting in the 2008-9 school year was in the 50th percentile of SAT test-takers (Goldhaber & Walch, 2014). I discuss in my conclusion the limits to generalizability and possibilities for further research given this recruitment procedure.

Semi-structured interviews lasted between 50-75 minutes and were conducted in person, via online video conferencing, or phone. Interviews were guided by a narrative inquiry approach (Carter, 1993; Kelterchmans & Vandenberghe, 1994; Tripp, 1994). In this approach, participants are asked to recall significant events and individuals in their lives to make explicit the influences on their present choices and actions. For this study, individuals were asked a single question: “Tell us the story of how you came to be doing what you are doing now career-wise. Where you start is completely up to you; as long as you can help us understand how your career interests started and how they have developed over time”. Follow-up questions were then asked chronologically, diving into various themes including: college experiences, including academics, extra-curricular activities, and internships; influences from parents, mentors, and peers; campus recruitment events and culture, etc. If teaching was not mentioned throughout the interview, participants were asked a final set of questions probing into any experiences and conversations they may have had where teaching as a career was brought up. Interviews were purposefully structured to in this open-ended way to ensure that the interviewer did not bias participants’ recollections by indicating hypotheses through specific questions. Interviews were audio-recorded and then transcribed verbatim.

For the present study, I analyzed interview data through two steps: *first*, I listened and re-listened to my participants’ life histories, coding the influences on their career exploration and

decision-making as they spoke about them. Codes were developed through listening--as I listened to one then another, I would reflect on whether any particular influences conformed or diverged from any previously identified code. After carefully listing and analyzing of these codes, I selected six participant experiences that seemed representative of the broader sample's experiences. These six experiences are what I present as portraits in this article (Rory's story is above, the five remaining are below). In selecting these six, I balance the range and frequency of codes across the sample. For example, family and peer influence were often reported across the sample, and therefore are presented across multiple portraits. Conversely, influences by university administrators was reported only a handful of times, and therefore is presented in only one portrait. Below I first present each of the individual portraits, and then in the final section, I draw out common themes and further questions. I purposefully reserve analysis and synthesis of the portraits until this final discussion and conclusion section to preserve the independence of each of the portraits. This is important given that interviews were conducted individually and across different times and contexts.

In analyzing and presenting my data in this way, the development of this assortment of portraits was akin to what others have described as "puzzle building" (Goetz & LeCompte, 1984; Lawrence-Lightfoot & Davis, 1994): first, a steady and systematic analysis of the data to draw out the 'pieces' from my participants' life histories; and then, a deliberate and iterative process of putting together various pieces to reveal the broader 'puzzle'. I quote my participants extensively, hoping to preserve as much as possible their distinct voices and the sentiments that undergird them. In doing this, I see my role as a curator, finding ways to put together the various pieces in ways that amplify their messages for the reader. As Lawrence-Lightfoot & Davis (1997: 85) describe it: "The actors sing the solo lines, the portraitist supporting their efforts at

articulation, insight, and expressiveness.”

Admittedly, in collecting, analyzing, and presenting data in this way, there may be -- as is always the case -- concerns about holistic fallacy (“interpreting events as more patterned and congruent than they really are”), elite bias (“overweighting data from articulate, well-informed, usually high-status informants”), and going native (“being coopted into the perceptions and explanations of local informants”) (Miles & Huberman, 1994: 263). To guard against these concerns, I did two things. First, as is a key tenet of portraiture, I sought out the “deviant voices” amongst my data, with the aim of constructing “a view that resists caricature” (Lawrence-Lightfoot, 2000). Below, I present some of these voices that seem not to fit the dominant conception with the recognition “that the contrasting image does not detract from the conception, but actually underscores its centrality and power” (Lawrence-Lightfoot & Davis, 1994: 248). And second, to ensure that I remained grounded in the broader motivation and context of the study, I collaborated with other researchers who had an understanding of the research literature and the context of elite colleges, but who did not participate in the data collection and analysis process. I discussed emerging themes and asked them to challenge me for evidence on themes that seemed conceptually muddy or empirically unsupported.

Finally, a note on subjectivity. Critical to both the portraiture and narrative inquiry approach is the role of subjectivity in data collection and analysis. In particular, interviewer positionality can lead interviewees to reveal or obscure certain information in conversation; and researcher positionality can lead to the over-privileging or the over-looking of certain findings. As such, it is important to note that I am both an insider and outsider to the focal phenomenon and context in this study. I am an insider in that I work closely as an undergraduate advisor and teacher preparation instructor at an elite institution. I am also an insider in that I have personally

experienced what my participants report: explicit and implicit signals that teaching is not an appropriate career for high achievers. However, I am an outsider to this work in that I was neither born in the US nor grew up in the context of elite higher education institutions. During data collection and analysis, I leveraged this insider/outsider status to advance the validity of my findings. I used my insider status to facilitate relationship-building with my participants, encouraging honesty and vulnerability in their responses; I used my outsider status to question participants' implicit assumptions, drawing out what was strange from what they consider familiar. During data analysis, I used my insider status to appropriately contextualize interview data, making efforts to ensure participants were accurately represented; I used my outsider status to step back and draw out generalizable themes that cut across contexts and speak to broader issues of teacher recruitment, career exploration, and high achievers' socialization.

Results

#1. "It's subtle"

It's only been two weeks since Ashley finished her Peace Corps placement teaching English to elementary school students in Nicaragua; and it's only been thirty minutes since she finished teaching a history lesson to grade 10 summer school students at the local high school down the road. And yet when we sit down to talk in the lobby of her highly-regarded graduate school of education, Ashley looks at me straight in the eyes with a presentness and warmth that belies any sense that her life over the past month has been a whirlwind. She has just started her one-year Masters in Teaching; the summer school class is her first practicum placement. "How are you adjusting to being back from Nicaragua?" I ask. "I love it," she beams. "It's so good to be here. I might crash eventually but right now, I'm running on adrenaline... I love working with

kids, I've always loved it. I've always felt so in my element in a classroom."

While I ordinarily might have relegated such platitudes to things one says in a job interview, Ashley's tone is kind and rings with an authenticity that is disarming, almost intimidating. I ask her whether she's always been so sure of this, whether she's always felt this positively about teaching. "No, definitely not. That's why I did the Peace Corps. That's why I didn't major in education." Ashley shrugs. I lean back and smile, signaling her to elaborate.

"In my undergrad, I was an International Studies major," Ashley replies. "I was interested in teaching but there was just a certain type of person that did the education program and I didn't really necessarily feel like I was that type of person... This was reinforced by my academic advisor, I just brought up in one of our meetings junior year that I had been volunteering in schools, and thinking about doing education, and he was kind of like 'ditch that, you have really high academic scores. You could be a lawyer, you can be a diplomat. Why would you want to go into teaching?'"

On the wall above Ashley is a framed blackboard; the phrase "every child, every classroom, every day" written in cursive in chalk. It's just one of the many positive messages about education and the value of teaching that adorn the walls of this campus. Ashley looks up at it, noticing my gaze. The irony isn't lost on her: "Don't get me wrong, I really enjoyed my college experience, but there was this moment where I feel like if people had encouraged me to go into education instead, I could have been tipped in that direction..." "What encouragement did you get instead?" I ask.

Ashley pauses for a moment. She starts slowly: "I feel... as I got older... people in my life and in my community... if I pitched the idea of teaching they would respond one way--not in

a negative way but they would respond maybe in a neutral way. But then when I pitched the idea of maybe wanting to go to the State Department or be a lawyer, their response was really enthusiastic.” I ask whether this was a feeling she got across the board, or if it was only specific people, like her advisor. “Not everyone but some people. I don’t know. It was really subtle but I just remember a couple conversations with professors and extended family members where I got that distinct-- If I said one thing, the reaction would be like, ‘oh, okay’ or like, ‘teachers work so hard but it’s such an important job’ ... Instead of like, ‘wow, that’s awesome, you’d be so great at that’ and ‘oh, that’s such a glamorous job, that would be such an exciting job to have’ which is the response I got when I mentioned I might be interested in the State Department or Peace Corps and stuff like that.”

I wonder how persistent this discourse has been throughout Ashley’s life: “Do people still say this stuff to you now?” Ashley smiles; she responds quickly: “Some people are a little bit surprised now when I tell them I’ve shifted away from International Studies and back towards doing education. No one is negative about it. Some people are surprised though.” I follow-up: “Can I ask how that surprise manifests?” “Yeah... Just like in conversations mostly with old friends and mentors from college, people who I didn’t talk to much while in Nicaragua.... Like when I said, ‘I’m applying to these Masters in Teaching programs’, and then they’re like, ‘Oh but I thought you were thinking about going into the foreign service or like doing law school or something like that?’ It’s kind of just a question, but it’s subtle, you know, the surprise that I ended up choosing to do teaching.”

#2. It’s Difficult

It’s Tuesday at 9:30pm halfway through Spring semester. Micah sits opposite on a large dark brown leather sofa, his backpack and jacket strewn on the floor beside him. We’re meeting

late because it's the only time he can make given the hours he's spending this semester in the engineering lab. We are sitting in a student common room that looks like a stereotype of an elite college's student common room. Dark wood bookshelves line the walls. A shield with house symbols and insignia hangs above a fireplace. Purple velvet curtains, pulled aside by cords with tassels, frame the floor-to-ceiling windows. A grand piano sits in the corner.

At 6'2", well-built, with a perfectly shaped fade, Micah looks exactly like he was a star Texas high school footballer. We've already discussed how he was recruited to play football at college; how he played one year and then decided it wasn't for him, preferring other extracurriculars like mentoring in schools and acapella. We've also talked about how he ended up doing engineering: "It was a fluke." In his first year at college, a friend told him he'd like the class. He went, he liked it, and he's just kept taking the classes ever since. I asked him if he had other career plans before arriving at Elite University, whether people around him had suggested he should study something in particular: "A lot of people associate [elite college] with the law school. That's the impression that comes to mind, so everybody used to think that I was going there to get a degree in law.... People always just thought I was going to be a lawyer. People would say things like you could be a future president, stuff like that. Nobody ever really said anything else or made any other assumptions."

I ask him about his family of teachers, wondering whether they held the same expectations. "I think it was beneficial for me, having a mother, grandmother, and aunt in teaching," Micah reflects. He sits contemplatively, leaned forward with his elbows on his knees, his hands pressed lightly together near his face. His posture matches his thought-process: his words are even-keeled. "It helped me establish and respect my teachers for one, and it also helped me with what I was learning. But I also saw the kinds of struggle they had working with

kids every day, some of the stresses they had from having to work for different administrators, or from working to get students ready for tests because that's like the big push nowadays. I guess I saw both the rewarding aspects of it and also some of the things that were less rewarding, the things that made the job a little bit harder for them."

"Did you talk about teaching with your mother, grandmother, or aunt while you were trying to figure out what to do for a career?" I ask. "Yeah," he responds quickly. "They really didn't have any direct advice. It was just do what you think is best because you're the one who has to live with that decision." I follow up: "But did hearing from them about how difficult teaching was make you stray away from considering it?" He demurs. "Maybe. My aunt and I, we talk everyday and share what we're both going through. And you know, sometimes the bad outweighs the good and when we talk, she mentions how bad the kids were that day, how disrespectful they were, the things that they have done, so on, and so forth."

I venture a hypothesis that these conversations with his family stimulated a disinterest in teaching as a career. Micah counters: "I think it was mostly experiencing it in schools myself." He starts telling me stories about volunteering in schools as a teaching assistant. "I was only there an hour a week so you don't get a great picture of what it'd be like to be a teacher full time... but I feel like it was very disrespectful for the teacher, even with someone else like me present, the kids didn't want to act appropriately. Just the way students were acting, the things they would say... The teachers were always struggling to get them to listen and to actually be interested in learning. And it wasn't all the students... you just have a couple in the class who act that way, and then it kind of takes away from the entire class." I press him to explain his observations of teaching further. I openly wonder whether he was just extrapolating from a couple of bad experiences, whether if he was just in a different classroom with different students

and a different teacher, everything might have turned out a little differently. “I wouldn’t say that I have a whole understanding of what being a teacher is like, but I’ve heard at home what my mother and aunt go through; and I’ve volunteered in schools a bit... and together it just gave me a snapshot of what it would be like to be a teacher and how difficult it would be.”

#3. Something to do “in between”

“I was very young, very very young... there’s a picture of me in my dad’s doctor’s office probably in junior kindergarten.” Ife laughs. She is responding to my question about when she knew that her family expected her to become a doctor. Ife and I both know that it isn’t an expectation that she’s resisted. As we’ve talked, she’s regaled me with stories of childhood: a doctor father and a nurse practitioner mother; a high school summer spent at a neuropsychology program that cultivated an intellectual curiosity; a relay of inspirational science and math teachers who challenged her and made her feel at home in the sciences; a community of pre-med college students with whom she has infinitesimally dissected and discussed a career in the life sciences. Ife punctuates these stories with sidebars and laughs, she is self-deprecating and self-aware.

I ask Ife about whether there has been anything else other than medicine in her pathway. Throughout our conversation she’s flagged an interest in education; she said she even once looked into Teach for America and her college’s fifth year teaching program as something “in between” college and medical school. “When I say I’m obsessed with school, it’s half a joke but I’m actually serious... I think I’ve just always really been into school, loved it, loved my teachers, I’ve just always really cared about how things are taught, when I got to college, I would just roast classes when I thought they weren’t taught to standard. It would just be in so many of my conversations with my family and friends.”

I follow-up: “So have you ever discussed with your parents about being a teacher?” Ife laughs: “I don’t know if we explicitly talked about being like a teacher teacher, but... we definitely talk about me being in academia. So if I told [my father] I was thinking about being a teacher, he would think I’m talking about being a college professor... He would be very comfortable about that because LOL you’re still going to medical school to teach something.” As Ife talks, she traces the deeply ingrained expectation of attending medical school and becoming a doctor. “But I think my mom... it’s very much in my family’s persona that I love school, so doing something like teaching wouldn’t be too out there. There would definitely be side-eye but my mom would probably eventually be very accepting of it, especially because in her mind it would be: ‘this will still lead to medical school!’ ... in her mind, teaching would just be something that I’d be doing ‘in between’ college and medical school.”

I wonder and ask where this idea of teaching as “in between” comes from. Ife wavers a little. “I think it’s just what’s more common for people here”. She mentions a friend’s partner, who completed the elite college’s fifth-year teacher certification program, taught for three years, and then went to law school. She notes a college advisor connecting her with a medical resident, also an alumnus, who taught for three years before medical school. Both spoke glowingly of their time in teaching, emphasizing how well it prepared them for their future careers. As Ife tells me about these alumni, I note how despite these glowing reviews, despite the social encouragement, Ife is still does not sound convinced about teaching. I wonder what it would have taken.

I ask her directly. Ife laughs. “My dad nearly cried when I told him I thought about becoming a lawyer. I took a constitutional law class in freshman year and was like ‘this is interesting’ ... yeah, that was a tough conversation.” Ife nods and smiles; she knows it’s funny but

also that it's not a joke. "The most side-eye I would get if I chose to become a teacher would be from my parents. But generally from [students at elite university] overall, they're either just neutral -- 'ok cool, you want to do that's great' -- but no one is wildly excited by it. I think no one is steering me away from it but no one is like 'Oh my gosh! You should definitely do this!' It's just something that you do in between."

#4. "I need to be more strategic"

It's been a month since graduation and Graham is halfway through his beach holiday before starting at a well-known management consulting firm in New York City. He tells me he's back home in California, renting a beach house with friends to pass the summer. The house is only 15 or so miles from the high-performing urban charter high school he attended. That said, early in our conversations Graham flags that given the friends he is on holiday with, the cultural references they make, and how much they are spending on recreational activities, the beach house is much more a reflection of the elite college he attended nearly 3000 miles away in the Northeast.

Graham is articulate and matter-of-fact. The sort of college student you talk to and get the impression they have it all figured out. He tells me he started college thinking he was interested in education: "I didn't know exactly where I wanted to go, or exactly what type of teaching or education that I wanted to do. I'd heard things like going to get a PhD and getting a professorship might be one way to approach it." He is purposeful in flagging that he was never told while in high school that doing a teacher preparation program was a possible option: "Everyone was just focused on getting me into a good college."

When asked, Graham flags his interest in education came from seeing the disparity between his urban charter school and the others around him. His mother was a teacher and ever

since early on, she encouraged him to see and investigate all the educational inequities that surrounded him. He wrote his college application essay on it: “I went through an incredible charter school in Los Angeles next to other charter schools that either lost their accreditation or otherwise sent less than 20% of people to college.” That said, when he got to college, “I took a few courses in education, did not get a minor or major or anything, and the teacher practicum at [elite college] was not considered particularly prestigious or robust and was generally thought of as a way to increase your GPA, or otherwise just not really respected as a program.”

“Where did you get the sense that the education courses weren’t that well respected?” I ask. Graham looks at me and smiles. He looks at me like the question is too obvious. Like any simple question about how we ‘know’ about prevalent socio-cultural norms, the answer is everywhere and nowhere at the same time. “Well for one, in freshman year a lot of other people wanted to do education but by the time I made it to senior year, no one talked about it at all... And then, if I told someone that I was considering doing teaching, I might get a sympathetic look. I might get something like, ‘That’s so nice of you. I don’t think I could ever do that. I’m not good with children. I can’t be an implementer or a low-level worker, I need to be more strategic.’ That I think was the most explicit that it got, that people just didn’t really respect the field of teaching.” Graham shrugs. These are codes I have heard from others as well; teaching described as ‘technical’ or ‘frontline’, simple codes used by some high-achievers to distinguish teaching from their personal interests focusing on ‘broader’, ‘high-level’, ‘policy’ and/or ‘strategic’ issues. I hear these codes as language games: delineations of occupational prestige and status masked otherwise as personal preferences.

I probe for more: “How else was it revealed that teaching was not respected?” Graham is unfazed. “Most of it was non-explicit; it wasn’t what people said but the ways they spent their

time. Even though people said at the very beginning that they were very interested in education, they would take courses in economics or they would take courses in political science but not in education and not in teaching. Or when it came to career fairs, we would walk down the hallway, and the private sector side would have lines of 10 to 20 people next each booth and the nonprofit side would have 5 people walking across 20 different companies. It was just visual. People were voting with their feet. It was just eminently clear that the two were not treated the same.”

Graham is quick to admit that these early and consistent signals from peers were persuasive for him. Instead of pursuing education, he majored in mechanical engineering. In his words, “just being surrounded by peers who are not considering your option makes it very difficult to choose it... you’re defying what is okay and what is normal by the people and activities around you.” Our conversation takes a pause and I can hear the waves in the background. I am curious why Graham is so socially influenced. I knew the university president of Graham’s alma mater had recently spoken at his graduation about the importance of not necessarily following others, doing what matters, choosing paths that will have a tangible impact on the world. So I ask: “You went to [elite university]! Why are you so worried about what everyone thinks? You made it! You got in, can’t you do whatever you want?”

Graham laughs. He can tell I’ve asked the question facetiously but he answers it seriously. “It’s funny, everyone thinks that we’ve got it but we’re all still too insecure. I certainly was, and definitely still am.” As he talks, it becomes clear Graham is trying to unravel for me his belief in the root cause behind the peer pressure. He asks me how many people I know who are trained teachers but can’t get jobs. He says he doesn’t know many teachers other than his mother, but he does know -- through college alumni networks and peer groups -- a fair few management consultants from the firm he is starting at in the Fall. He says he knows no one who looks or

sounds like they have had trouble getting future jobs. At first I think this is a joke but Graham pushes on: “There’s no reason to believe that the rest of your life is going to be easy after college, even after a place like [elite university]. Throughout college I applied for so many jobs that just never responded to me. I know so many friends who have graduated and have not gotten jobs... consulting has some of the best exit opportunities, same as finance. You’ve got to hedge your bets.”

#5. “It was so absolute”

“In a sense... hmm... I was actually pretty shocked by my mother’s reaction, I guess, initially... and like it actually made it so... we don’t argue that often, but there was a lot of discontent, at the time... and then we talked about it later... but like, I think, she just felt the logistical impossibility... no not impossibility, it would just be significantly difficult to like pursue it...” We are at the end of our interview and I can tell that Emily can’t find the words. Sitting on the floor of her childhood bedroom, she shifts her posture so the angle of the webcam bounces. Emily apologizes and blames it on tiredness -- she just submitted her senior thesis yesterday, a web application that uses machine learning to analyse sound bites, a merging of her two interests in linguistics and computer science. I tell her there’s nothing to apologize about -- that we have the time and space to help her find the words.

Until this point, the conversation has been winding but not incomprehensible. Emily speaks quickly and in short utterances; the sentences come together, but only after piecing together the various fragments of thought. She often asks me, “Does that make sense?” and I take a moment to respond to give us both time to think through all that she’s said. So far, Emily has described a range of interests that developed early in high school (language learning, robotics, dance, music) that progressively coalesced into her current interest (human computer

interactions) and led her to her post-graduation plans (software engineer at a large technology firm). A moment ago, she flagged that for “a tiny hot second” she considered teaching after college instead of taking the technology firm’s offer. “What made you think about teaching in the first place?” I ask.

With a father working in curriculum development in early childhood, and a mother working as a Mandarin teacher, Emily knew that teaching was a career, she just had never thought about it. Even when asked explicitly, she could never remember anyone suggesting or even asking her about teaching (“While growing up did anyone ask you whether you had thought about becoming a teacher?” “No.” “How about a doctor?” “Yes, loads.”). And then came a pivotal conversation with her roommate in sophomore year: “My roommate had just come back from the Education Career Fair and she had a bunch of pamphlets and we were going through them, talking about them... We always talked about careers and stuff, whenever anyone came back home and was thinking about something or wanted to talk about some internship or job they found... Anyway, we were talking and she just said something along the lines of: ‘I think you would be really great, and there are not enough people who teach STEM so please do it!’”

While divergent from the stories above, I am not surprised by Emily’s roommate’s emphatic support of teaching. I know that her elite college has recently embarked on raising the visibility of public service careers like teaching, having recently approved an education studies minor and a fifth-year teacher certification program. Emily acknowledges this herself early on in our conversation: “I remember at [college orientation] the Dean said something which I remember... people have this perspective that like you’re either a doctor, lawyer, accountant, engineer, or a loser. There’s really like nothing else you can be.” Having attended the Dean’s orientation speech, I also remember the line. It was followed by an entreaty to reject social

pressures that associate occupational prestige with meaning and impact. The lines were part of a broader hope that the incoming class would explore diverse career options, take academic and extracurricular risks, and not just follow their predecessors into the well-worn post-graduation paths of consulting, finance, medicine, and technology.

“And so when was the moment you brought up teaching with your parents?” I follow-up. Emily umms and ahhs. She shifts her seating position and the camera angle bounces a little again, pulling her in and out of focus. “So a week after that, because my parents live nearby and so we just eat and chill for a bit every week or so. I brought it up and... it was not well received. Which is why it was, like, shocking because they don’t frequently have that kind of reaction to things that I say... Most times they say things like ‘keep exploring’, you have time to make up your mind. This was the first time it was so absolute.” Emily pauses again.

After some silence I try to see if she might elaborate: “What do you mean by shocking? What sorts of things was your mom saying?” Emily responds quickly, spouting fragments as if she’s just remembered so much and yet can’t describe it in words: “She was like shocked in a negative way, mostly because she felt like she knew it would be extremely challenging... and especially having experienced, to some degree, teaching herself... she just... especially given the environment I grew up in in suburban Massachusetts, it’s not necessarily meant I’ve had... like the experiences where I would be equipped to work in urban schools... Even if I would hope to be able to make an impact, I shouldn’t do it for growth because it’s not like what I’ve been prepared for... because like maybe there are others who are more or better prepared for it... she just, she just had a pretty negative reaction.”

We sit in silence for a few more beats. Emily nods, “yeah that was pretty much it.” I nod as well. Looking at Emily, I can tell that it wasn’t so much exactly what her mother had said but

the speed and emphasis of the rejection that was impactful. Put into the broader context of supportive interactions encouraging her to explore various career options, the swiftness and negativity must indeed, have been shocking. “And so you never really thought about teaching after that?” “Yeah, I never did again.”

Discussion and Conclusion: What can we learn?

What do these portraits tell us about how high-achievers learn that they should not become teachers? First and foremost, Rory, Ashley, Micah, Ife, Graham, and Emily’s stories demonstrate how high-achievers’ ideas about teaching and careers are *learned*. A naive model for why high-achievers have poor conceptions of teaching as a career may run as follows: (1) teachers are paid poorly and not respected; (2) high-achievers want careers that are paid well and well respected; and therefore, (3) high-achievers do not want to become teachers. Read collectively, the above stories add depth, breadth, and texture to this simple model. High-achievers are not born knowing that teaching is not an appropriate career path for someone like them. Rather, they learn it and relearn it through the signals built into their interactions with family and friends. They learn it through the expectations of others in their social circles. They learn it through observing the career trajectories of their role models and their peers.

Second, these stories emphasize that while for some the signal to not pursue teaching is explicit (e.g. Ashley’s academic advisor; Emily’s parents), for most high-achievers, the signals are implicit -- they learn through implication, observation, and by example. Ashley learned that teaching was not respected through observing her peers’ and mentors’ variable responses to her career interests. Micah had reinforced that teaching was not respected by listening to stories from his family members and observing students while volunteering in local schools. Ife learned through observing others who graduated from her university that teaching is not treated so much

as a career but an “in-between” before doing something else like law or medical school. These stories emphasize that high-achievers’ negative conceptions of teaching are fostered and reinforced primarily not through what they hear others *say*, but through what they see others *do*. As Graham remarks, “it wasn’t what people said but the ways they spent their time”.

Third, these stories highlight the particularly strong influences of parents in high-achievers’ developing conceptions of teaching. Although the trope of “following in the family business” is very familiar in other sectors, it has been underexplored as it relates to the career exploration and decision-making process of prospective teachers. What makes this an interesting avenue for future inquiry is that the direction and magnitude of influences seem variable. For some, parental influences can be negative towards teaching (e.g. Emily’s shock in discovering her parents’ strong disapproval of her nascent interest in teaching; Micah’s admission that his family’s challenging experiences as teachers set a frame for how he came to perceiving teaching); for others, positive (e.g. Graham’s mother encouraged him as a student to observe the educational inequity surrounding him).

Fourth, these stories complicate the narrative for how high-achievers develop poor conceptions of the teaching profession. Previous research has emphasized the material conditions of teaching, like poor pay, working conditions, and limited career opportunities. The stories above suggest that most high-achievers do not learn about these material conditions directly -- rather they learn about them through social interactions. Emily did not learn that teaching could be poorly paid from teachers’ union agreements or employment contracts; rather, she learned them through her mother’s objections. Ife did not learn about the long-term career options of teachers through school system workforce planning documents; rather, she learned about them by talking to others who chose to teach and then left to pursue other careers.

And fifth, these stories enlarge the scope of what contributes to poor conceptions of the teaching profession. Ashley was told that she should not become a teacher because her grades were too high. Graham observed that the career opportunities for those interested in education were not deemed attractive enough by those at his college. Both Ashley and Graham were routed away from education studies programs at their respective college because they felt that others perceived it as not ‘rigorous’ enough. These instances suggest that we should think about how poor perceptions of teaching are not just caused by material conditions of *already practicing teachers*; we should also think about the influence of conditions for those who are *prospectively considering teaching* (e.g. the prevalence and accessibility of socially-sanctioned pathways to becoming a teacher).

So what do these lessons suggest for how we can disrupt negative conceptions of teaching amongst high-achievers? They suggest that, like any other social institution, high-achievers’ negative conceptions of teaching are learned and that they are learned in specific sites and in specific ways all along high-achievers’ broader career exploration and decision-making process. It is beyond the scope of this paper’s findings to conclude what must be done to disrupt all this learning. However, the findings indicate some possible levers for future investigation. For example, given how high-achievers career exploration mostly occurred during college, elite institutions could further efforts to raise the visibility and desirability of exploring public impact careers like teaching. Like Ife’s Dean during orientation, it behooves institutions to explicitly address students’ and parents’ underlying conceptions of public service careers like teaching as socially-legitimate pathways. Similarly, given how high-achievers mostly learned about teaching second- or third-hand, K-12 school systems can offer structured opportunities for high-achievers to gain access to schools and learn directly about the promises and possibilities of a career in

teaching. In teacher education, it is common to say that “it takes a village to train a teacher”. My findings suggest it also takes a village to recruit one.

In identifying these potential policy implications, it is important to note the limits of generalizability given the specific higher education contexts from which study participants were samples. The above six high-achievers attended four different selective higher education institutions (three Top 20 National Universities, and one Top 20 Regional University). As such, there remains an open question for future research into how high-achievers who attend less selective institutions like large state schools are socialized. Could it be that these high-achievers receive less negative signals about teaching because they attend institutions where teaching is a common career path? Or could it be that they receive more negative signals about teaching in efforts to differentiate them from the ‘average student’ in their institution?

It is also important to restate that my purpose in this article has not been to causally link negative conceptions of teaching to high-achievers’ unwillingness to become teachers, nor to highlight that some socialization processes may be more or less likely to lead to high-achievers choosing to become teachers. Indeed, this is why aside from Ashley, I purposefully do not identify that two other high-achievers (Rory and Graham) chose to become teachers in the year following our interview. Rather, my aim was to present the range of ways that study participants learned that teaching was inappropriate for high-achievers like them. I look to future research to provide us with better empirical evidence of the strength of this relationship.

A few weeks after conducting interviews, I ask a soon-to-retire colleague about what they think it will take in order for us to convince parents that their high-achieving angels should become teachers. She looks me up and down; she is a battle-worn veteran general looking down at a young uncertain officer. For over twenty years she has been director of a highly-regarded

teacher preparation program and a professor of education studies at an elite college. We've talked often of students who come into our offices and say, "Well, I'd love to be a teacher but my parents don't want me to waste my [elite college] education". Her response is always: "I go ballistic! What could be more important? Let me talk to your parents."

She recounts a story from a Junior Parents weekend she spoke at a few years ago. "We were in the lecture theater and I had them tape a card under every seat. I said to the parents, 'I want you to look under your seat. There's a card. I want you to write down the name of an important teacher.' Everybody does it in a heartbeat. A to Z, everyone that's there, everyone's got a great teacher. And they don't just have the name, they have stories, they really want to tell you about them. They're desperate to tell you about them."

"Then I said, 'Okay, so you've got this name of this person who was an important teacher to you. I'd like you to imagine that 20 years from now, we do the same thing and people pull up this card and they write down the name of your son or your daughter.' The room was dead silent, and I didn't say anything more. You can call it a guilt trip or whatever, but just how would you feel? You know me, I don't take any prisoners. I just let it sit."

References

- Auguste, B., Kihn, P., & Miller, M. (2010). *Closing the talent gap: Attracting and retaining top-third graduates to careers in teaching: An international and market research-based perspective*. McKinsey & Company.
- Barber, M., & Mourshed, M. (2007). *How the world's best performing school systems come out on top*. McKinsey & Company.
- Battilana, J., Leca, B., & Boxenbaum, E. (2009) How Actors Change Institutions: Towards a Theory of Institutional Entrepreneurship. *The Academy of Management Annals*, 3(1), 65-107. <https://doi.org/10.1080/19416520903053598>
- Binder, A. J., Davis, D. B., & Bloom, N. (2016). Career funneling: How elite students learn to define and desire “prestigious” jobs. *Sociology of Education*, 89(1), 20–39. <https://doi.org/10.1177/0038040715610883>
- Cohen, D. K. (2011). *Teaching and Its Predicaments*. Harvard University Press.
- Cowan, J., Goldhaber, D., Hayes, K., & Theobald, R. (2016). Missing elements in the discussion of teacher shortages. *Educational Researcher*, 45(8), 460–462. <https://doi.org/10.3102/0013189X16679145>
- Darling-Hammond, L. (1984). *Beyond the commission reports: The coming crisis in teaching*. RAND Corporation. <https://www.rand.org/pubs/reports/R3177.html>
- Duncan, A. (2011). *Working Toward “Wow”: A Vision for a New Teaching Profession*. <https://www.ed.gov/news/speeches/working-toward-wow-vision-new-teaching-profession>
- Etzioni, A. (1969). *The semi-professions and their organization: Teachers, nurses, social workers*. Free Press.

- Farkas, S., Johnson, J., & Foleno, T. (2000). *A sense of calling: Who teaches and why*. Public Agenda.
- Featherstone, Joseph (1989). To make the spirit whole. *Harvard Educational Review*, 59(3), 367-378. <https://doi.org/10.17763/haer.59.3.0392h53811473716>
- Glaser, B. G., & Strauss, A. L. (1967) *The discovery of grounded theory: Strategies for qualitative research*. Aldine Publishing Company.
- Goldhaber, D., & Walch, J. (2014). Gains in teacher quality. *Education Next*, 14(1), 38–45.
- Goldring, R., Taie, S., and Riddles, M. (2014). Teacher attrition and mobility: Results from the 2012–13 teacher follow-up survey (NCES 2014-077). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Ho, K. (2009). *Liquidated: An ethnography of Wall Street*. Duke University Press.
- Holmes Group (1986). *Tomorrow's teachers*. <https://eric.ed.gov/?id=ED270454>
- Hoyle, E. (2001). Teaching: Prestige, status and esteem. *Educational Management & Administration*, 29(2), 139–152. <https://doi.org/10.1177/0263211x010292001>
- Ingersoll, R., Merrill, L., & Stuckey, D. (2014). *Seven trends: The transformation of the teaching force* (No. #RR-80). Philadelphia, PA.: Consortium for Policy Research in Education.
- Kennedy, M. M. (2006). *Inside teaching: How classroom life undermines reform*. Harvard University Press.
- Labaree, D. (2004). *The trouble with ed schools*. Yale University Press.
- Lawrence, T., Suddaby, R., & Leca, B. (2011). Institutional work: Refocusing institutional studies of organization. *Journal of Management Inquiry*, 20(1), 52–58. <https://doi.org/10.1177/1056492610387222>
- Lawrence-Lightfoot, S., & Davis, J. H. (1997). *The art and science of portraiture*. Jossey-Bass.

- Lortie, D. (1975). *Schoolteacher: A sociological study*. University of Chicago Press.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Sage Publications, Inc.
- Rivera, L. (2015). *Pedigree: Elite reproduction in hiring*. Princeton University Press.
- Rowsey, R. E., & Ley, T. C. (1986). Perceptions of teachers' salaries and non-salary benefits. *Journal of Teacher Education*, 37(2), 42–45.
<https://doi.org/10.1177/002248718603700208>
- Schjeldahl, Peter (2019, September 23). Beheld. *The New Yorker*.
<https://www.newyorker.com/magazine/2019/09/23/the-amy-sherald-effect>
- Schneider, B., & Stevenson, D. (2000). *The ambitious generation: America's teenagers, motivated but directionless*. Yale University Press.
- Smith, T. W., Davern, M., Freese, J., & Morgan, S. L. (2019). Appendix F. *General Social Surveys, 1972-2018: Cumulative codebook* (No. 25; National Data Program for the Social Sciences Series). National Opinion Research Center.
- Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the US*. Learning Policy Institute.
- Weingarten, R. (2019). *The Freedom to Teach*. National Press Club, Washington, DC.
<https://www.aft.org/freedomtoteach>

Paper #3

Impact Evaluations of Teacher Preparation Practices: Challenges and Opportunities for More Rigorous Research

Abstract

Many teacher education researchers have expressed concerns with the lack of rigorous impact evaluations of teacher preparation practices. I summarize these various concerns as they relate to issues of internal validity, external validity, and measurement. I then assess the prevalence of these issues by reviewing 166 impact evaluations of teacher preparation practices published in peer-reviewed journals between 2002-2019. Although I find that very few studies address issues of internal validity, external validity and measurement, I highlight some innovative approaches and present a checklist of considerations to assist future researchers in designing more rigorous impact evaluations.

Introduction

Researchers and policymakers have long sought to identify the most effective teacher preparation practices, i.e., the “approaches, activities and processes that can be undertaken by teacher educators to develop pre-service teachers’ knowledge, skill, and dispositions” (Hill, Mancenido, & Loeb, 2021). A wide variety of teacher preparation practices have been studied, including general education, content area, and methods courses, as well as modules or workshops run within or in parallel to coursework (e.g., Floden & Meniketti, 2005); teacher educator pedagogies such as practice-based approaches (e.g., Forzani, 2014); and, ways of structuring elements of teacher education programs like paired placements during teaching practicum (e.g., Nokes et al., 2008).

Much has been learned from studying these teacher preparation practices through the most prevalent research approaches in the field, including interpretive research, practitioner research, and design research (Borko, Liston & Whitcomb, 2007). Some teacher educator researchers, however, have noted the lack of rigorously designed impact evaluations, i.e., studies able to identify the effects of specific teacher preparation practices on pre-service teachers and their (future) students (Cochran-Smith & Villegas, 2015; Cochran-Smith & Zeichner, 2005; Grossman, 2008; Grossman & McDonald, 2008; Sleeter, 2014; Wilson, Floden, & Ferrini-Mundy, 2002; Zientek et al., 2008). This lack of impact evaluations is a problem because it means researchers are limited in the guidance they can provide teacher educators and policymakers about what works for whom and in what contexts.

The aim of this article is to support teacher education researchers in designing more rigorous impact evaluations of teacher preparation practices. I do this by first identifying the primary issues of internal validity, measurement, and external validity that have hampered the

ability of researchers to make causal inferences about teacher preparation practices over the past two decades. Then I provide examples and guidance for researchers on how they may be able to overcome these issues in future studies.

I begin by defining impact evaluations and justifying the need for improving their quantity and quality in the field of teacher education. Then I review the research designs and measurement approaches of 166 impact evaluations published in peer-reviewed journals between 2002-2019. Finally, I develop a checklist of considerations for designing and reporting on impact evaluations of teacher preparation practices. This checklist aims to support teacher education researchers to move closer towards developing a shared evaluation paradigm (Grossman & McDonald, 2008) by meeting calls for more “clearly articulated criteria for high quality studies” (Grossman, 2008), and for more explicit guidance for peer reviewers “about the kinds of things that should be present in empirical research for it to be published” (Zeichner, 2005).

The Need for Impact Evaluations

This article focuses on *impact* evaluations of teacher preparation practices -- elsewhere described as “effects of teacher education” research (Borko, Liston, & Whitcomb, 2007) and “effectiveness research” (Hill, Mancenido, & Loeb, 2021). Impact evaluations investigate the effects of specific teacher preparation practices on pre-service teachers’ knowledge, skills, and/or dispositions, and/or on their (future) students’ learning and development. This is different from *process* evaluations (also referred to as *program monitoring* or *implementation analysis*), which generally aims to investigate how certain teacher preparation practices are implemented, and/or the extent to which they are implemented as designed (Humphrey et al., 2016).

According to modern standards for making causal claims in social science research (King, Keohane, Verba, 1994; Mackie, 1974; Shadish, Cook, & Campbell, 2002), teacher

education researchers undertaking impact evaluations need to: (1) clearly identify a relationship between a teacher preparation practice and specific outcomes (i.e., when X changes, Y also changes); (2) ensure that the practice precedes the expected changes in the outcomes (i.e., X happens before Y); and, (3) reject any other reasonable explanations for changes in the outcome (i.e., nothing else can explain the changes in Y except X). Unfortunately, reviews of teacher preparation research over the past two decades consistently lament the lack of research that meet these standards, with most research in the field not designed to isolate the effects of specific teacher preparation practices from the many varied influences on pre-service teacher development (e.g., Wilson, Floden, & Ferrini-Mundy, 2002; Grossman, 2008).

This is problematic for both political and practical reasons. From a political standpoint, teacher preparation programs are increasingly challenged to demonstrate their impact on their graduates' outcomes using causal methods (Cochran-Smith, et al. 2017; Teacher Preparation Issues, 2016; Wineburg, 2006)³. As many “insider” critics of the field have expressed (Grossman, 2008; Zeichner, 2005), if teacher education researchers continue to provide few compelling responses to these challenges, we as a field will continue to lose “professional jurisdiction” over our work. This is because lack of causal evidence may continue to be interpreted as lack of impact and, therefore, utility of teacher education (rather than a lack of investment in generating causal evidence); and/or, because policymakers will continue to pay greater attention to (and we as a field will continue to be at the mercy of) research conducted by

³ In some ways, these recent political pressures can be interpreted as a result of hotly-contested debates in the late 1990s / early 2000s over whether policymakers could better improve overall teaching quality by prioritising improving the quality of teacher preparation or the quality of incoming teacher candidates (e.g., Ballou & Podgursky, 2000; Darling-Hammond, et al., 2005). These debates suffered from a lack of rigorous causal evidence on the impact of teacher preparation, leading many states towards increased accountability pressures to incentivise programs to generate such evidence.

those outside our field (e.g., psychologists and economists)⁴. As Wilson, Floden, & Ferrini-Mundy (2002) warn: “unless we—as teacher educators and researchers—produce sound, robust measures of impact, others—policy makers and critics—will produce other, less appropriate measures” (p. 201).

There is also a practical imperative for increasing the number and improving the quality of impact evaluations of teacher preparation practices. Without much evidence on the teacher preparation practices that lead to more effective teachers, teacher educators have little concrete guidance when deciding how to best prepare their pre-service teachers (Grossman, 2008), and policymakers have little evidence upon which to base decision-making about accreditation requirements, accountability regimes, and grant programs. Recently there has been some guidance coming from program-level evaluations showing certain program characteristics may be particularly effective (e.g., Boyd et al., 2009). However, as Del Schalock et al. (2006) note, these studies “carry little explanatory power as to why, how, or what within teaching or teacher preparation account for relationships found and, thus, have limited utility in guiding or refining policy, practice, or research” (p.110).

The Challenges of Undertaking Impact Evaluations

Despite the political and practical imperatives, many teacher education researchers have acknowledged the difficulties of undertaking impact evaluations. Reasons include: lack of funding and allocated research time for scholars of teacher education (Cochran-Smith, 2004; Cochran-Smith et al., 2012; Feuer, et al., 2013); heavy teaching loads and mostly part-time or

⁴ This is not to say that important research cannot or should not be undertaken by non-teacher educators. Rather, it is simply to acknowledge that teacher educators operate in a highly regulated and politically volatile environment where decisions about our work are often not guided by our own research (Mayer, 2021). Further, as Wilson, Floden, & Ferrini-Mundy (2002) warn: “unless we—as teacher educators and researchers—produce sound, robust measures of impact, others—policy makers and critics—will produce other, less appropriate measures” (p. 201).

adjunct status of practicing teacher educators (Borko, Liston, & Whitcomb, 2007; Nuttall et al., 2006; Zeichner, 2005); historical separation between research and practice amongst education school faculty (Goodlad, 1990; Labaree, 2004); limited doctoral training on evaluative methods given the dominance of interpretive and design-based research over the past few decades (Borko, Liston, Whitcomb, 2007; Grossman, 2008; Wilson, 2006; Zientek et al., 2008); the lack of teacher education research published in top-ranked peer-reviewed journals (Grossman, 2008; Wilson, Floden, & Ferrini-Mundy, 2002; Zeichner, 2005); lack of capacity to engage in multi-site and/or multi-disciplinary partnerships so as to undertake large-scale evaluations (Borko, Liston, & Whitcomb, 2007; Del Schalock, et al. 2006); the politicisation of teacher education that leads much research and practice to be reactive to political imperatives and the local teaching market (Grossman & McDonald, 2008; Cochran-Smith & Villegas, 2015).

Although these challenges are significant, I share Nuttall et al.'s concern that:

these explanations have become something of a 'default setting', and do not go far enough in explicating the complex set of circumstances with which teacher education researchers must engage in order to attain the research quality and credibility necessary to influence initial teacher education policy, as well as teacher education practice (2006: p. 326).

As such, the aim of this article is to explicate -- and support teacher education researchers to address -- one such circumstance that hampers the "research quality and credibility" of impact evaluations in the field: the lack of research design and measurement approaches that meet modern standards for making causal claims.

Issues of Research Design and Measurement

Reviews of teacher education research stretching over the past two decades have raised concerns about the research design and measurement approaches of impact evaluations of teacher preparation practices (Grossman, 2008). For example, in the first comprehensive review

of research on effective teacher preparation, Wilson, Floden, & Ferrini-Mundy (2002) noted the limited empirical evidence base for many common teacher preparation practices and made the first major call for more rigorous impact evaluations in the field. Similarly, Cochran-Smith (2005) writes in her summary of the findings and recommendations of the *AERA Panel on Research and Teacher Education*: “In particular, we need more and better research on the outcomes of teacher education... This requires better data collection and analysis tools for studying outcomes and consistent use of these tools across individual studies” (p. 302). More recently, Sleeter (2014) reviewed the 196 articles published in four top-ranked teacher education journals in 2012, finding: “Only about 1% of the articles reported large-scale mixed-methods studies, only 6% examined the impact of teacher education on teaching practice and/or student learning, and only one did both” (p. 6).

While many teacher education researchers have raised concerns, no single review has summarized these concerns about research design and measurement in one place. As such, in what follows, I bring together themes from teacher education researchers over the past two decades⁵ to identify the key issues of internal validity, measurement, external validity that hamper impact evaluations of teacher preparation practices.

Internal validity

Most studies of teacher preparation practices use small-scale, single-case study designs to investigate how a group of pre-service teachers attending a single teacher preparation program experience a specific teacher preparation practice unique to their context (Cochran-Smith &

⁵ The reviews and research commentaries consulted to summarise these issues were: Borko, Liston, & Whitcomb (2007); Cochran-Smith & Zeichner (2005); Cochran-Smith & Villegas (2015); Cochran-Smith et al. (2012); Cochran-Smith et al. (2015); Cochran-Smith et al. (2016); Crawford & Tan (2018); Grossman (2008); Grossman & McDonald (2008); Hill, Mancenido, & Loeb (2021); Livingstone & Flores (2017); Mayer (2021); Menter et al. (2010); Nuttall et al. (2006); Sleeter (2014); Wilson (2006); Wilson, Floden, & Ferrini-Mundy (2002); Zientek et al. (2008).

Zeichner, 2005; Cochran-Smith et al., 2016; Grossman, 2008; Menter et al., 2010; Wilson, Floden, & Ferrini-Mundy, 2002). This single-case approach often allows researchers to develop a deep and nuanced understanding of the dynamic interplay between a teacher preparation practice and pre-service teacher experiences; i.e., how are particular practices in this context enacted, received, perceived, and therefore, learned from?

However, this single-case approach does not allow researchers to identify the causal effects of teacher preparation practices. This is because single-case designs leave open many possible alternative explanations for observed changes in participant outcomes -- i.e., threats to internal validity (Murnane & Willett, 2010; Shadish, Cook, & Campbell, 2002). Three in particular have been flagged by prior reviews as rarely accounted for in impact evaluations: (1) pre-service teachers' natural learning over time (*maturation effects*); (2) pre-service teachers opting into (or out of) receiving a teacher preparation practice that is particularly suited towards them (or not) (*selection effects*); and, (3) improvements caused by other teacher preparation practices that occur at the same time (*history effects*)⁶. Without using a research design that can account for these threats to internal validity, impact evaluations cannot isolate whether study findings are due to the teacher preparation practice or any of these other explanations.

As Zeichner (2005) describes in his summary of the findings and recommendations of the *AERA Panel*, what is needed are research designs that allow for “*systematic analyses of distinct alternatives*” (p. 745), such as those that compare the relative effectiveness of two different teacher preparation practices, or those that compare a novel teacher preparation practice with ‘business as usual’ (‘control’). Doing this will allow researchers to better “disentangle the

⁶ See Shadish, Cook, & Campbell (2002) for a full typology of threats to internal validity in causal research, not all of which have been noted by prior reviews of the field.

influence of teacher characteristics from those of their preparation programs” (Zeichner & Conklin, 2005: p. 663). While practically challenging to implement in the context of teacher education, this includes using research designs that account for natural learning over time, such as through comparing the outcomes of a group of study participants who received the preparation practice with a group who did not. It also means using research designs that ensure participants are not able to choose to be part of (or not part of) the treatment group, and that the only difference in the experiences of participants during the study is that one group receives the teacher preparation practice.

Measurement

The second set of issues raised by prior reviews relate to measurement. One commonly raised is that impact evaluations of teacher preparation practices often do not use replicable measurement approaches, such as clearly defined and consistent rubric or rating schemes when measuring participant outcomes (Cochran-Smith et al., 2015; Floden & Meniketti, 2005; Grossman, 2008). Instead, teacher education researchers tend to prefer idiosyncratic measures usually developed through subjective inductive analysis of collected data. These measurement procedures are often not reported in detail. As Zeichner (2005) remarks: “many studies reviewed [by the *AERA Panel*] provide no information about how instruments used for data collection were developed and, validated, or how their reliability was assessed” (p.741). This makes it difficult to assess the validity of researchers’ inferences during the peer review process, and/or design future replication studies to help build the evidence base.

A related issue raised by prior reviews is the lack of commonly used measures across impact evaluations. This limits the accumulation of knowledge across the field, because when most studies use idiosyncratic measures, researchers undertaking meta-analyses or literature

reviews cannot confidently compare the relative effectiveness of different teacher preparation practices, nor can they evaluate whether any differences in study findings across contexts are simply due to different measurement approaches (Zeichner, 2005). One contributing factor to this issue is that not enough researchers in the field are producing and openly providing quality measures for use by others (Grossman, 2008; Grossman & McDonald, 2008), particularly at the grain size needed for evaluations of teacher preparation practices (Zeichner, 2005). As Zientek et al. (2008) note, “In the present set of studies, surveys were the dominant mode for collecting data, but only about half (52%) of the researcher-developed surveys were made available” (p.212).

Another issue raised by prior reviews is that impact evaluations of teacher preparation practices rarely investigate effects on pre-service teachers’ practice and student learning, preferring more proximal outcomes like pre-service teacher perceptions and understanding⁷. For example, in their review of research on preparing PSTs to teach science, Cochran-Smith, et al. (2016) note most studies “focused on relatively short-term changes in teacher candidates’ understandings and beliefs in the context of science methods courses”. Similarly, Menter et al. (2010) find in their review of teacher education research in the UK that 60.1% of the 446 studies reviewed use “reflection” as their primary method of data collection. However, as Wilson, Floden, & Ferrini-Mundy (2002) state: “Although it is important to know how teachers feel about the benefits of field experiences, attitude surveys do not answer questions about what

⁷ Cochran-Smith (2004) notes that much of this can be attributed to field norms: “over the past two decades... Very little of [teacher preparation] research was designed to establish empirical linkages to pupils’ learning, partly because teachers’ knowledge, learning, and beliefs were assumed to be important outcomes of teacher preparation in and of themselves and partly because it was considered self-evident that teachers who knew more, taught better” (p.113). More recently, research has empirically demonstrated: (a) that there is a relationship between teacher knowledge, teacher practice, and student outcomes (e.g., Hill & Chin, 2018; Piasta et al., 2009), and (b) that teachers self-reported perceptions of the effectiveness of teacher professional development often do not match with observed changes in their knowledge, practice, and student outcomes (e.g., Copur-Gencturk & Thacker, 2020; Jacob, Hill, & Corey, 2017).

prospective teachers actually learn”. In other words, just because pre-service teachers learn it in teacher preparation, doesn’t mean that they will do it later on when they are teaching, or that their K-12 students will learn from how it is implemented (Clift & Brady, 2005; Diez, 2010; Gainsburg, 2012). For researchers to identify these longitudinal effects, they need to use measurement approaches that can allow for multiple outcome measures over time -- approaches that allow researchers to track pre-service teachers into their classrooms.

Finally, prior reviews have raised concerns about the extent to which measurement approaches mitigate participant and researcher bias. Two issues are particularly salient: first, the issue of *stakes* -- i.e., using outcome measures, such as course grades or licensure tests, that have stakes attached for participants. Measures like these may show inflated effects because participants may try harder than they would otherwise, and/or researchers may unknowingly bias their grading towards identifying effects. Second, the issue of *reactivity bias* -- i.e., participants changing their behavior to meet their or others’ expectations about the effectiveness of the teacher preparation practice. Though admittedly practically challenging to do in the context of teacher education, not accounting for reactivity bias (e.g., by blinding participants to their treatment condition) can lead to potential placebo effects, and/or findings that may primarily be due to participants changing their behaviour as a result of being studied or their “desires to please the instructor” (Clift & Brady, 2005: p. 303).

External validity

The third set of issues raised by prior reviews relate to external validity -- i.e., the extent to which study findings may generalize to other teacher preparation programs and/or other pre-service teachers. This is a particular challenge for the field given research is often conducted within the specific context of a single teacher preparation program (Cochran-Smith et al., 2016;

Sleeter, 2014). That said, designing and reporting on external validity is critically important given there are thousands of diverse and dynamic teacher preparation programs across the US, and policymakers and teacher educators need to know whether and how study findings can scale to their/other contexts (Conaway & Goldhaber, 2020). Researchers also need to understand what works for whom and in what contexts in order to build a shared evidence base (Wilson, Floden, & Ferrini-Mundy, 2002).

When it comes to external validity, prior reviews of the field have primarily raised concerns about the lack of reporting for generalizability (Cochran-Smith et al., 2012; Grossman, 2008; Zeichner, 2005; Zientek et al., 2008). Two issues are particularly salient: (a) the lack of reporting on the representativeness of study participants (*participant generalizability*) -- e.g., by comparing the demographics of study participants to either the demographics of the teacher preparation program they are sampled from and/or the demographics of pre-service teachers in the US; and (b) the lack of discussion on the representativeness of study site(s) (*site generalizability*) -- e.g., by naming study site(s) and/or describing their location, size, program type, and certification level such that others can determine the relevance of findings to other teacher preparation contexts. It is important to note that more is not necessarily better when it comes to reporting on participants and sites. Rather, as Floden & Meniketti (2005) caution, teacher education researchers need to report study details in more standardized ways that enable us to compare across studies and build a shared evidence base:

when courses [in teacher preparation programs] were described, the descriptions were idiosyncratic. Authors would mention particular texts or describe the academic tasks they posed to students, in ways that offered detail, but gave scant opportunity for linking to other studies. (p. 286)

Prior reviews have also raised issues regarding the extent to which studies are designed for scalability. Impact evaluations of teacher preparation practices are often conducted by their

developers, who are often highly enthusiastic about, skilled in, and/or suited to the practice. This can limit external validity as study findings may not replicate when teacher preparation practices are undertaken by other teacher educators. Unless study authors mitigate these potential *instructor effects* in their research design (e.g., by training instructors to deliver the teacher preparation practice), others cannot assess whether the practice can be scaled to other contexts. A related issue raised by prior reviews is the potential bias of researchers who study their own practices as teacher educators (*researcher bias*). Researchers may unknowingly make and/or overlook reporting on various research design and measurement choices that limit the generalizability of study findings to other contexts. As Wilson, Floden, & Ferrini-Mundy (2002) note: “Although local teacher educator researchers have valuable knowledge of the phenomenon under investigation, critics have the right to raise questions about the conflict of interest apparent for teacher educators doing research that validates the need for teacher education” (p.194).

The Present Study

To aid future researchers in addressing these issues of internal validity, measurement, and external validity, I systematically reviewed the research designs and measurement approaches of impact evaluations of teacher preparation practices conducted over the past two decades. I conducted an online and hand search of studies of teacher preparation practices to identify impact evaluations published in peer-reviewed journals between 2002-2019. I then coded studies based on whether they address the issues of internal validity, measurement, and external validity identified above. My aim was to both empirically assess the prevalence of these issues to help identify priorities and opportunities for the field, as well as highlight innovative and promising approaches that can be emulated by future researchers.

While no other study has focused exclusively on the research design and measurement

approaches of impact evaluations of teacher preparation practices, three studies have undertaken a similar methodology of assessing the state of the field to provide guidance for future researchers. First, Zientek et al. (2008) reviewed the reporting practices of all quantitative research studies cited in *Studying Teacher Education: The Report of the AERA Panel on Research and Teacher Education* (Cochran-Smith & Zeichner, 2005). The authors coded studies based on their statistical methods and the extent to which they report various tests or statistics as recommended by the AERA Standards. Second, Sleeter (2014) investigated the extent to which teacher education research is “designed to influence policy” by reviewing 196 articles published in four top-ranked journals in the field in 2012. Sleeter (2014) categorized studies based on research design and the extent to which they connected teacher preparation policies and practices to impacts on teachers. Finally, Crawford & Tan (2019) investigated the extent to which teacher education research uses mixed methods by reviewing articles published in *Journal of Teacher Education* and *Teaching and Teacher Education* between 2010-2016. The authors coded empirical studies based on whether they used qualitative, quantitative, or mixed methods.

In all three of these studies, the aim in assessing the state of the field was to highlight specific opportunities to advance research -- whether that be more standardised reporting practices in quantitative studies, more research designed to influence policy, or more mixed-methods studies. In this paper, I seek to do this same, but with a focus on supporting teacher education researchers to conduct more rigorous impact evaluations of teacher preparation practices⁸.

Method

⁸ While there is a small overlap in the sample of studies reviewed in this paper and those reviewed by Zientek et al. (2008), Sleeter (2014), and Crawford & Tan (2019), my review analyses these studies differently as my focus is on research design and measurement approaches.

Guided by general principles for undertaking systematic literature reviews (Kennedy, 2007; Polanin, Maynard & Dell, 2017; Slavin, 1986), I first defined my inclusion criteria, and then undertook a four-step process for identifying and assessing studies. This process is summarized in Figure 2.1 and detailed below.

Inclusion Criteria

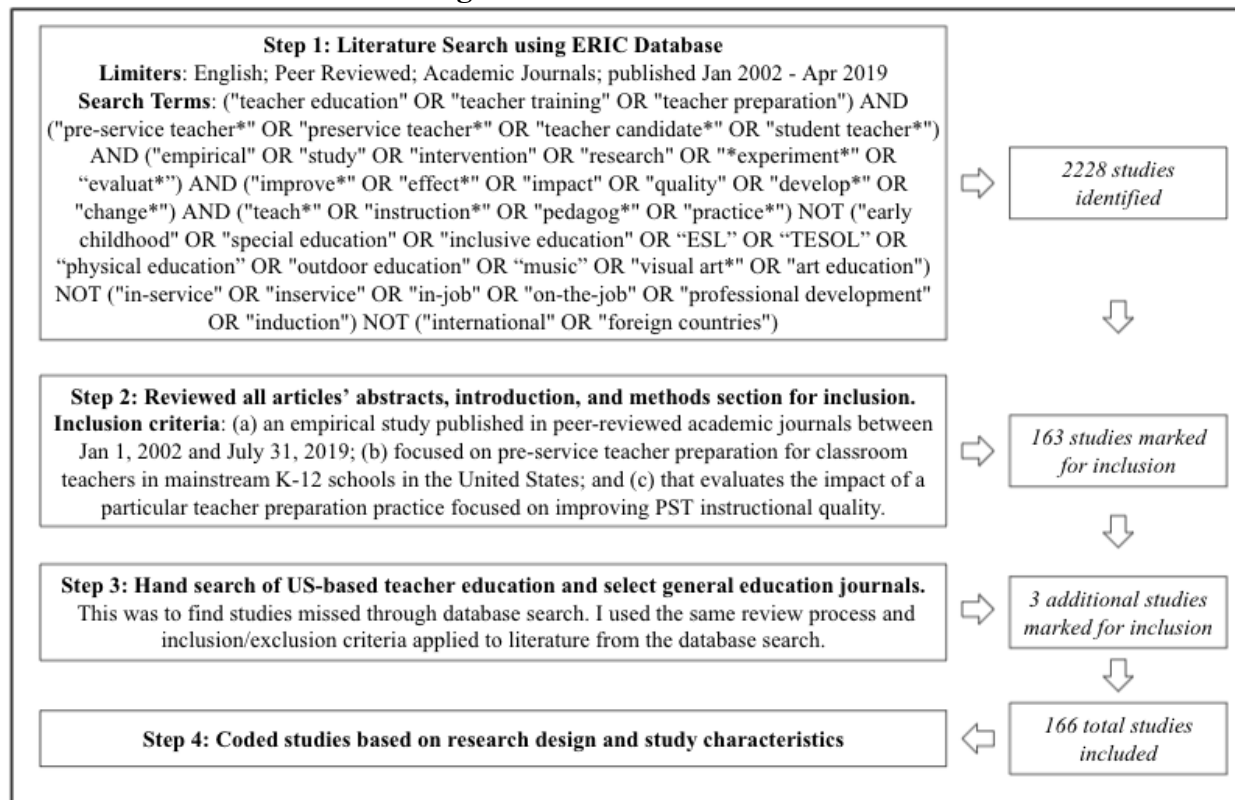
I defined the scope of my review as *impact evaluations of teacher education practices* focused on *improving K-12 pre-service teachers' instructional quality* that were *published in peer-reviewed journals* between *January, 2002 and April, 2019*. I chose to begin from 2002 because this was the year of the first major call for more rigorous evaluative research in the field (Wilson, Floden, & Ferrini-Mundy, 2002). I also chose 2002 because it was the beginning of increased support for more rigorous impact evaluations in education through the establishment of the Institute of Education Sciences (Whitehurst, 2003). I also only included studies conducted within the US and that had been published in peer reviewed journals. I did this in order to investigate the development of the field since 2002, and to ensure future reviews can more easily replicate my approach and/or compare findings.

Impact Evaluations

I defined *impact evaluations* as studies intending to investigate the *causal effects* of a specific teacher preparation practice on one or more outcomes. I included studies where authors used causal language like “evaluate”, “impact”, “effective[ness]”, “change”, “influence”, or “develop” when describing the purpose of the study in their introduction, rationale, research questions, or methods sections. I excluded studies that simply described the implementation of a teacher preparation practice and/or studies that only sought to understand how pre-service teachers experienced a teacher preparation practice. There was no minimum sample size

necessary for inclusion; and both quantitative and qualitative studies were included.

Figure 2.1. Review Procedure



Teacher Preparation Practices

I defined teacher preparation practices as either: (1) a course or set of courses organized around a particular theme within a teacher preparation program (e.g. a redesigned science methods course focused on inquiry; a new course for elementary teachers on sustainability); (2) an activity, pedagogy, or process implemented within a course in a teacher preparation program (e.g. the use of peer feedback on assignments; a short module on assessment within a math methods course); (3) a way of organizing a program component like the practicum (e.g. being assigned a certain type of mentors during practicum; being assigned to practicum placements in pairs); or (4) an experience organized by teacher educators for the purpose of developing pre-service teachers' skills, knowledge, or dispositions (e.g. undertaking a workshop on behavior

management; a penpal arrangement between pre-service teachers and Grade 4 students). Based on this definition, I excluded studies that evaluated the effects of whole programs, and those that evaluated a bundle of practices related to a single philosophical approach.

Instructional Quality

Given the wide range of teacher preparation practices studied in the field and my limited resources, I followed the lead of Wilson, Floden, & Ferrini-Mundy (2002) and chose to focus only on teacher preparation practices directly aimed at improving pre-service teachers' *subject matter knowledge* and *pedagogical skills*, as well as teacher preparation practices conducted during or to improve the *clinical experience*. This meant that I included studies of teacher preparation practices that developed pre-service teachers' disciplinary content or pedagogical knowledge or skills, as well as their knowledge and skills in using certain general instructional moves (e.g., for classroom management); assessing and providing feedback to students; developing unit plans and instructional materials; and/or addressing diverse academic and behavioural needs (e.g., supporting English language learners and students with Individualized Education Plans)⁹. Studies did not need to directly measure instructional quality to be included (e.g., they could use changes in pre-service teacher knowledge or teacher self-efficacy as the primary outcome).

Step 1 & 2: Online Search and Inclusion

I conducted an online search in May 2019 of the primary US education research database (ERIC) using the limiters and search terms described in Figure 2.1. This returned an initial pool

⁹ This focus meant I excluded teacher preparation practices focused on domains more distal to instructional quality such as: family and community engagement, general teaching self-efficacy, skills and self-efficacy in using technology in the classroom, attitudes towards diversity, equity, and inclusion. Although research has shown that these domains are related to instructional quality, I chose to exclude these studies to ensure a manageable sample. I posit that we may be able to assume results from my sample of studies may generalize to studies focused on these other domains. I leave to future reviews to test whether this is empirically true.

of 2228 studies. The abstract, introduction, research questions, and methods sections of the studies were reviewed based on the criteria described above to determine whether the study should be included in the review.

To ensure reproducibility, three research assistants were trained on and suggested clarifications to the criteria until independent percent agreement reached above 80%. 600 studies were double-coded in this manner. The remaining studies were coded for inclusion by me. To safeguard against accidental exclusion, two research assistants double-coded 100 randomly selected studies that had been coded only by me. No studies were found to be accidentally excluded.

Based on this review, 2065 of these studies were excluded and 163 were marked for inclusion. This rate of exclusion may seem large but is similar to other systematic reviews of teacher preparation research. For example, Sleeter (2014) analysed the 196 articles published in 2012 in four leading teacher education journals internationally and found “only 6% to examine the impact of teacher education on teaching practice and/or student learning”.

Step 3: Hand Search

Despite the wide online search parameters, I was concerned that key studies may have been excluded. As such, two research assistants undertook hand searches of the *Journal of Teacher Education*, *Teachers and Teaching*, *Teaching and Teacher Education*, *Educational Researcher*, *American Educational Research Journal*, and *Educational Evaluation and Policy Analysis* to identify any studies that were missed. Any studies flagged for inclusion were discussed with me and a consensus decision made. An additional 3 studies were identified through this process.

Table 2.1. Characteristics of Included Studies

Code	% (n)
<i>Type of teacher preparation practice evaluated</i>	
• A course or set of courses	29.52% (49)
• A pedagogy or assessment within a course	54.22% (90)
• A structuring of a program component	9.64% (16)
• A short program, workshop, or learning experience	6.63% (11)
<i>Study sample</i>	
• Number of participants	Mean = 56.5, Median = 32, Min = 3, Max = 542
• Studies that did not report number of participants	2.41% (4)
<i>Level of teacher preparation program attended by participants</i>	
• Undergraduate	65.06% (108)
• Postgraduate	27.71% (46)
• Not reported	19.88% (33)
<i>Type of teacher preparation program attended by participants</i>	
• Traditional	57.83% (96)
• Alternative	6.02% (10)
• Not reported	39.76% (66)
<i>Certification level of participants</i>	
• Elementary	66.27% (110)
• Middle/high school	36.14% (60)
• Not reported	9.04% (15)
<i>Types of data collected</i>	
• Survey	65.66% (109)
• Teacher educator's observations or reflections	41.57% (69)
• Written reflections or journals	34.94% (58)
• Interviews	34.94% (58)
• Course assessment (e.g., test; assignment)	27.71% (46)
• Observations of the pre-service teacher's instruction (e.g., in practicum settings and simulated environments)	15.06% (25)
• External task (e.g., EdTPA)	3.01% (5)
• Other	6.63% (11)
<i>Quantitative analysis</i>	
• Uses statistical analysis	39.76% (66)

Characteristics of included studies

A total of 166 studies were included in this review. A reference list is presented in Online Appendix C, and a summary of study characteristics is presented in Table 2.1. Briefly, most studies evaluated a pedagogy or assessment within a course (54.22%; n=90) or a particular course or set of courses (29.52% n=49). The average number of study participants was 56.5, with 65.4% of studies having 50 or fewer participants, and 85.19% of studies having 100 or fewer participants. The modal study was of pre-service teachers at the undergraduate level (65.05%; n=108), attending traditional preparation programs (57.83%; n=96), and seeking certification at the elementary level (66.27%; n=110). The most common form of data collected was surveys (65.66%; n=109), followed by teacher educators' observations or reflections of pre-service teachers' participation in the teacher preparation practice (41.57%; n=69), pre-service teachers' written reflections or journals (34.94%; n=58), and interviews (34.94%; n=58). Finally, 39.76% (n=66) of studies used statistical analyses on at least one outcome measure.

Step 4: Coding

Coding Scheme

I developed my coding scheme in partnership with another researcher with experience conducting systematic literature reviews of impact evaluations of teacher professional development. As no prior review had already identified a set of issues with research design and measurement that could be used as a coding scheme, we developed codes by first establishing a shared understanding of issues raised by prior reviews of the field (as discussed above). We then used our shared understanding to curate and simplify the taxonomy of issues related to internal validity, measurement, and external validity described in Campbell & Stanley (1963) and elaborated by Shadish, Cook, & Campbell (2002). We then added to these codes based on our

broad reading of research in program evaluation, causal inference, and research design (e.g. King, Keohane, & Verba, 1994; Murnane & Willett, 2010; Institute of Education Sciences, 2020) as well as our understanding of the contexts of teacher education as researchers and practitioners. We then iteratively refined and clarified this coding scheme using a random sample of included studies.

It is important to note that these codes were not developed to reify a single research design, such as a randomized controlled trial (RCT). I do not subscribe to the view that RCTs are the only way to generate evidence of causation nor that a particular research design is “rigorous” and others are not. This is why I began this review by discussing issues with impact evaluations raised by prior reviews rather than discussing the prevalence of any one particular design. I wanted to ensure that I approached my review from the problems we seek to address (i.e., the issues of internal validity, measurement, and external validity that teacher education researchers need to overcome), rather than from some preconceived belief that any one particular research design or any particular measurement approach -- qualitative or quantitative -- is the best.

The full coding scheme is presented in Appendix D, with key codes relevant to the present study described in the following sections and summarized in Table 2.2.

Internal Validity. I used three codes to assess whether studies addressed issues of internal validity raised by prior reviews: (a) whether the study used a comparison groups design to account for natural learning over time or ‘business as usual’ (*maturation effects*)¹⁰; (b) whether participants were able to choose to be part of the treatment group (or not) (*selection effects*); and

¹⁰ Comparison groups designs are the most prevalent but not the only approach to addressing maturation effects. Other approaches include single-case designs (Kratochwill et al., 2010) and -- as discussed in the results section -- using unrelated dependent variables (e.g., Morris & Hiebert, 2017). For a comprehensive discussion of alternative designs, see Shadish, Cook, & Campbell (2002). I chose to focus on comparison groups designs as they are the most simple and common way to account for maturation effects in education research.

(c) whether the only difference between comparison groups was that one group received the teacher preparation practice (*history effects*).

Measurement. I used three sets of codes to assess whether studies addressed issues of measurement. The first set of codes related to whether the study used replicable measures. There were two codes: (1a) whether participant outcomes were measured using a clearly defined and consistent rubric or rating scheme (*replicable measures*); and (1b) whether measures had previously been developed and used by other researchers to enable comparison of results with other studies (*established measures*). The second set of codes related to whether the study measured teacher practice and/or student outcomes. There were two codes: (2a) whether the study collects outcome measures of pre-service teachers' practice and/or student learning (i.e., not just pre-service teachers' perceptions or participation records) (*practice measures*); and (2b) whether the study collects multiple outcome measures over time to assess the persistence of effects (*longitudinal measures*). The third set of codes related to whether the study addressed issues of participant and researcher bias. There were two codes: (3a) whether the study did not use outcome measures that had stakes attached for participants (*low stakes*); and (3b) whether study authors blinded participants to their treatment condition and/or discuss the effects of participants changing their behavior based on their expectations (*mitigated reactivity bias*).

External Validity. I used two sets of codes to assess whether studies addressed issues of external validity. The first set of codes related to whether the study discussed the representativeness of the study sample. There were two codes: (1a) whether the demographics of study participants were compared to either the demographics of the teacher preparation program they are sampled from and/or the demographics of pre-service teachers in the US (*participant generalizability*); and (1b) whether study authors named the study site(s) (i.e., the teacher

preparation program in which the study took place), and/or described the location, size, program type, and certification level of study site(s) (*site generalizability*)¹¹. The second set of codes related to whether the study was designed for scalability. There were two codes: (2a) whether multiple instructors were trained to deliver the teacher preparation practice and/or study authors discussed potential issues of having instructors highly enthusiastic about, skilled in, or suited to the teacher preparation practice (*instructor effects*); and (2b) whether the study was undertaken in one of the study authors' courses (*researcher bias*).

Coding Process

All included studies were independently coded by myself and at least one other trained research assistant. Any discrepant codes were discussed and consensus reached. Following all coding, I calculated summary statistics in order to describe the study characteristics, research designs, and measurement approaches of included studies. I ran correlational analyses and trend analyses to determine any relationships between codes, and whether any types of studies or methods had become more or less prevalent over time. I found nothing meaningful and so do not present these results below, however, they are available upon request. I then went back to the individual codes to identify research designs and measurement approaches that addressed particular issues in innovative ways and could be used as guidance for future researchers. Finally, I synthesised the coding framework, the summary statistics, and the identified noteworthy studies to develop a common checklist of considerations for designing and reporting on impact evaluations of teacher preparation practices.

¹¹ Another factor “*selectivity*” was considered as part of this code, however, it was not included as very few studies reported on it and there is no standard way for reporting across undergraduate and graduate teacher preparation programs (e.g., acceptance rate, average achievement on standardized tests like the SAT/ACT/GRE, undergraduate GPA).

Table 2.2. Percent of Impact Evaluations Addressing Issues of Internal Validity, External Validity, and Measurement

Code	% (n)
<u>Internal validity</u>	
<i>Study authors used a research design that analyses distinct alternatives</i>	
A comparison groups design was used to account for natural learning over time or ‘business as usual’ (<i>Maturation effects</i>)	25.9% (43)
Participants are not able to choose to be part of the treatment group (or choose not to be a part of it). (<i>Selection effects</i>)	Of the 43 studies that used comparison groups, 65.12% (28)
The only difference between comparison groups was that one group received the teacher preparation practice (<i>History effects</i>)	Of the 43 studies that use comparison groups, 55.81% (24)
<u>Measurement</u>	
<i>Study authors used replicable measures</i>	
Participant outcomes were measured using a clearly defined and consistent rubric or rating scheme (<i>Replicable measures</i>)	45.18% (75)
Study authors used measures developed by other researchers to enable comparison of results with other studies (<i>Established measures</i>)	18.67% (31)
<i>Study authors measured teacher practice and/or student outcomes</i>	
Study authors collected outcome measures of pre-service teachers’ practice and/or student learning (<i>Practice measures</i>)	25.3% (42)
Study authors collected multiple outcome measures over time to assess the persistence of effects (<i>Longitudinal measures</i>)	6.02% (10)
<i>Study authors mitigated participant and researcher bias</i>	
Study authors do not use outcome measures that have stakes attached for participants (e.g., course grades, licensure tests) (<i>Low stakes</i>)	64.46% (107)
Study authors blinded participants to their treatment condition and/or discussed potential biases from participants changing their behavior based on their expectations (<i>Mitigated reactivity bias</i>)	4.22% (7)
<u>External validity</u>	
<i>Study authors discussed the representativeness of the study sample</i>	
The demographics of study participants were compared to <i>either</i> the demographics of the teacher preparation program they are sampled from and/or the demographics of pre-service teachers in the US. (<i>Participant generalizability</i>)	13.25% (22)
The study site(s) were named and/or their location, size, program type, and certification level were described. (<i>Site generalizability</i>)	21.69% (36)
<i>Study authors designed the study for scalability</i>	
Multiple instructors were trained to deliver the teacher preparation practice and/or study authors discussed potential issues of having instructors highly enthusiastic about, skilled in, or suited to it. (<i>Instructor effects</i>)	Of the 158 studies that evaluated a practice delivered by an instructor, 10.13% (16)
The study was not undertaken in one of the study authors’ courses. (<i>Researcher bias</i>)	Of the 92 studies undertaken in courses and that reported course instructors, 7.61% (7)

Results

In this section, I present my findings on the extent to which the impact evaluations reviewed addressed issues of internal validity, measurement, and external validity. My general finding is that all issues raised by prior reviews remain prevalent in the field, but that there are bright spots in the literature. I highlight these studies as guidance for future researchers, and then present a checklist of considerations for those seeking to undertake more rigorous impact evaluations of teacher preparation practices.

Internal Validity

Systematic Analyses of Distinct Alternatives

Of the 166 studies reviewed, only 26.67% (n=43) addressed maturation effects by using a comparison groups design. Though not initially accounted for in my coding scheme, an additional 3 studies addressed maturation effects by using within-person comparisons of intended vs unrelated outcomes (1.81%; n=3; design described below). Given only comparison group designs can account for selection or history effects, I restricted analysis of those issues to only the 43 studies that used a comparison group design. Most of the 43 studies addressed selection effects by not allowing participants to choose whether they received the teacher preparation practice, either through randomization (46.51%; n=20), or by comparing cohorts of pre-service teachers over multiple semesters or years (18.60%; n=8). The remaining studies either allowed pre-service teachers to select which group to be in (11.63%; n=5), or did not report how pre-service teachers were assigned to groups (23.26%; n=10). Finally, of the 43 studies that used comparison groups, only 55.81% (n=24) addressed history effects by ensuring that the only difference between comparison groups was that one group received the teacher preparation practice. A summary of research designs used by studies is presented in Table 2.3.

These results empirically affirm concerns raised by prior reviews that very few studies are designed to rigorously evaluate the causal effects of teacher preparation practices.

Table 2.3. Types of Research Designs Used

Code	% (n)
<i>Type of Research Design</i>	
• Posttest-only single-case	28.31% (47)
• Pretest-posttest single-case	43.98% (73)
• Posttest-only comparison group	10.24% (17)
• Pretest-posttest comparison group	15.66% (23)
• Within-person comparison (i.e., intended vs unrelated outcome)	1.81% (3)
<i>Method of assigning groups for studies using a comparison group design (n=43)</i>	
• Random assignment	46.51% (20)
• Cohort-/time-based	18.60% (8)
• Participants chose group	11.63% (5)
• Did not report	23.26% (10)

Many teacher education researchers have acknowledged the practical challenges of systematically analysing distinct alternatives given the context of teacher education: i.e., teacher preparation programs are and operate within highly institutionalized organizations, and so researchers often have little control over what, when and how pre-service teachers learn (Boyd et al., 2008; Labaree, 2004). Despite this, some studies reviewed still managed to address issues of maturation effects, selection bias, and history effects by leveraging program structures. For example, Kopcha & Alger (2011) leveraged the fact that pre-service teachers in their program were “placed randomly into one of four cohorts at the beginning of the year” and so assigned treatment at the cohort level. Similarly, Mahalingappa, Hughes, & Polat (2018) randomly assigned the teacher preparation practice being evaluated to one of two sections of a course. The authors reported a minimal likely impact of selection bias given: (a) students had no foreknowledge of which section would be assigned the teacher preparation practice; (b) students

enroll in sections primarily based on their schedule; and (c) researchers picked the section to be treated without knowledge of who was enrolled (see also Hanuscin & Zangori, 2016 for a similar design).

Some studies creatively analysed distinct alternatives by comparing groups of pre-service teachers across years or semesters, leveraging changes outside the control of study participants. For example, Matthews and Seaman (2007) evaluated the effects of a required specialised math content course by comparing cohorts the year before and after the course was introduced (see also Mizell & Cates, 2004; Smith et al., 2012). Similarly, Yadav et al. (2014) evaluated the effects of a one-week computational thinking module implemented in a required educational psychology course by comparing outcomes with students who took the same course (sans module) in the previous semester.

Some researchers argue that given the high-stakes nature of teacher preparation, it may be unethical to use comparison group designs if some pre-service teachers are not receiving a teacher preparation practice that may be beneficial (Sikes, 2006). One way some studies dealt with this issue was to use a counterbalanced design, where one group receives one teacher preparation practice while another group receives a different one; then after a period of time, they switch. For example, Bulunuz & Jarrett (2009) randomly assigned one group of students to one teacher preparation practice (learning stations) for three topics (A, B, C) and a different practice (scaffolded reading) for three different topics (D, E, F); the other group of students was assigned the opposite (i.e., learning stations for topics D, E, F and scaffolded reading for topics A, B, C). Participant outcomes were measured after each set of topics.

Another way studies dealt with ethical concerns while still addressing issues of internal validity was by using within-person comparisons of intended vs unrelated outcomes. Pioneered

by Morris & Hiebert (2017) and then replicated by Suppa, DiNapoli, & Mixell (2018) and Hiebert et al. (2019), these studies measure the effects of teacher preparation mathematics courses years after graduation. The studies do this by comparing graduate teachers' skills and knowledge across multiple mathematical content areas, some of which were explicitly taught during teacher preparation courses and some which were not. For example, Hiebert et al. (2019) test graduates on their conceptual understanding and ability to assess students' work in the topics of multiplying whole numbers, subtracting fractions, and dividing fractions (topics taught during their first-year mathematics courses) as well as finding the mean (a topic not taught throughout the teacher preparation program). To mitigate history effects (i.e., the effects of any other teacher preparation practices or other experiences that could lead to differences in knowledge and skills across different mathematics topics), Hiebert et al. (2019) collect data from graduates about other educational experiences, and then systematically test whether those experiences are driving effects (e.g., other experiences during the preparation program; in-school professional development; the foci of school curricula).

Measurement

Reproducible Measures

Of the 166 studies reviewed, only 45.18% (n=75) measured participant outcomes -- whether qualitatively or quantitatively -- using a clearly defined and consistent rubric or rating scheme. For example, Isabelle & de Groot (2008) established a three-level rubric to analyse the depth of students' written open-ended responses to questions requiring scientific explanations. Bartell et al. (2013) developed a rubric for measuring pre-service teachers' conceptual understanding of mathematical content and their evaluations of children's mathematical understanding. Foley et al. (2017) and Soh, Fowler, & Zygielbaum (2007) coded the complexity

of pre-service teachers' concept maps of their disciplinary content knowledge. The remaining studies reviewed (54.82%; n=91) used idiosyncratic measures developed through subjective inductive analysis of collected data.

Of studies reviewed, only 18.67% (n=31) used measures developed by other researchers to enable comparison of results across studies. Some study authors used measures wholesale (e.g., Mathematics Teaching Efficacy Beliefs Instrument (MTEBI); Science Teaching Efficacy Beliefs Instrument (STEBI)); while others adapted or augmented existing measures to accommodate additional outcomes. Using previously established measures allowed some researchers to demonstrate how their study was building on previous findings and developing a shared evidence base. For example, Wilkins & Brand (2004) used the Mathematics Belief Instrument (MBI) developed by Hart (2002). The authors noted how their study findings “provide additional support for the reliability of the initial quantitative findings related to the MBI presented by Hart (2002)”, and that this replication was important given their study “was conducted with a larger sample of preservice teachers”. Similarly, Lambert & Bleicher (2013) used an established measure of people's perspectives on climate change, and then compared pre-service teachers' knowledge and beliefs about climate change with the broader US adult population. This comparison allowed them to discuss the magnitude of their study findings in relation to the broader US population.

In most cases, however, study authors did not draw connections between studies despite using measures developed by other researchers. Sometimes this was because no other studies had used the same measure in the context of teacher preparation, as noted by Kopcha & Alger (2011) and McGee & Colby (2014). However, in most cases study authors used established measures such as MTEBI or STEBI but overlooked the opportunity to compare the magnitude of impacts

and/or to shed light on similarities across study populations.

Teacher Practice and/or Student Learning

Of the 166 studies reviewed, only 25.3% (n=42) evaluated the effect of a teacher preparation practice on pre-service teachers' practice and/or their students' outcomes. The remaining studies (74.7%; n=124) only collected and analysed data on participants' perceptions of the teacher preparation practice and/or on tests and surveys of their skills, knowledge, or beliefs.

Forty studies collected and analysed data on pre-service teachers' practice. Of these, 37.5% (n=15) used pre-service teachers' self-reports, 42.5% (n=17) used direct observations of participants' instruction, and 20% (n=8) used both. The primary way studies collected data on teacher practice was through pre-service teachers' school placements, whether concurrent to or soon after the teacher preparation practice. For example, Girod & Girod (2006) timed their Web-based simulation experience in between already-scheduled mentor lesson evaluations of pre-service teachers. This allowed them to evaluate the effects of the simulation using data on pre-service teachers' practice. Some study authors triangulated practice data with other artifacts, such as lesson plans and instructional materials. These artifacts allowed researchers to track impact from teacher preparation practice, to pre-service teacher learning, to pre-service teacher enactment. For example, Bangel et al. (2010) analysed both lesson plans and lesson observations to observe whether what was taught ended up implemented in both instructional planning and instructional practice (see also Welsh & Schaffer, 2017).

One challenge with collecting teacher practice data is that it is resource intensive; teacher education researchers often do not have the research funding to hire raters to conduct classroom observations, nor the time to do it themselves. Some studies overcame this challenge by drawing

on data already collected through state licensure requirement such as EdTPA (e.g., Santagata & Yeh, 2014). Other studies tasked pre-service teachers to collect data themselves, usually through course structures like assignments or online discussions. For example, Welsh & Schaffer (2017) used data from four course assignments spread across a semester; each assignment required students to provide a written lesson plan, a 10–15 minute video segment of their teaching the lesson, and a two- or three-page written reflection. Similarly Caughlan et al. (2013) used data from a course discussion board where pre-service teachers were required to post and reflect on multiple 5-minute videos of them teaching. However, Caughlan et al. (2013) admitted that because videos were chosen by pre-service teachers, “any measured change over time is partly an artifact of why each candidate chose each video”.

Another challenge with collecting teacher practice data from practicum is that the quality of mentor teachers and placement school may significantly influence pre-service teachers’ practice. This means that any findings could actually reflect mentor teacher or placement school quality rather than the effects of the teacher preparation practice. One study, Kopcha & Alger (2011), addressed this issue by collecting two survey measures of practicum support (*Learning to Teach Scale*; *Learning Climate Questionnaire*), and then testing for any relationship between these variables and the outcome.

Only five studies reviewed (3%) collected data on student outcomes. A summary of these studies is presented in Table 2.4; they all leveraged distinct course structures that allowed for collection and analysis of student outcomes during practicum placements. Three studies collected perception data (i.e., pre-service teachers’ observations, or students’ self-reports); two studies collected student learning or behavioral data. For example, Stover, Yearta, & Sease (2014) collected and analysed blog posts that students wrote as part of an electronic pen pal arrangement

with pre-service teachers (which was the teacher preparation practice being evaluated). Allen & Blackston (2003) collected multiple direct observations of students' target behaviors. Although pre-service teachers collected and quantitatively coded this observational data, the study authors sent independent observers to each student on three occasions to test for rater reliability.

Table 2.4. Summary of Studies Collecting Student Outcome Measures

Study	Teacher preparation practice evaluated	Student outcome measure
Allen & Blackston (2003)	Training pre-service teachers in “the use of collaboratively developed intervention scripts... to address academic or behavioral concerns about a target student”	Graphical analysis of trends in students' target behaviours. Student target behaviours were observed and quantitatively coded by pre-service teachers. Independent observers rated a sample of students' target behaviours to assess pre-service teachers' rater reliability.
Bullough et al. (2003)	A peer teaching model, where two student teachers work with one mentor during a student teaching placement	Thematic analysis of data from focus groups of students from participants' classrooms. Students were “asked questions about the advantages and disadvantages of having multiple student teachers”.
Stover et al. (2014)	Pre-service teachers blogging with fifth graders (“electronic pen pals”) about commonly read texts	Thematic analysis of student blog posts.
Thomas & Sondergeld (2015)	A “scaffolded approach” to teaching pre-service teachers to provide effective feedback on student work	Quantitative analysis of students' perceptions of feedback before and after working on a project with pre-service teachers (who provided them with written feedback at multiple timepoints). Thematic analysis of students' written responses regarding their feelings about receiving feedback on their work and how teachers could make their feedback to students more meaningful.
Weinburgh (2007)	A multi-week investigation of mealworms in an elementary science methods course. As part of this investigation, pre-service teachers brought mealworms to their school placement and taught children about them.	Thematic analysis of pre-service teachers' observations of any changes observed in the organisms and reactions of the children in their school placement.

Finally, of the studies reviewed only 6.02% (n=10) collected outcome measures over multiple time points to assess the persistence of effects (*longitudinal measures*). Three studies of these ten studies followed pre-service teachers across multiple semesters to assess the impact of content or methods courses on pre-service teaching practice during their practicum. For example,

in addition to the data they collected during and directly after the science methods course being evaluated, Windschitl, Thompson, Braaten (2008) also conducted multiple lesson observations and debriefs with study participants in their practicum placements a year later (see also Cartwright, Smith, & Hallar, 2014; Ragland, 2016). The remaining seven studies ran the same posttest outcome measure (or a parallel form) multiple times (Hiebert et al., 2019; Isabelle & de Groot, 2008; Morris & Hiebert, 2017; Rodriguez-Valls & Ponce, 2013; Sayeski et al., 2015; Sayeski et al., 2017; Suppa, DiNapoli, & Mixell, 2018). This allowed study authors to assess whether pre-service teachers retained the knowledge and skills that they learned from the teacher preparation practice over time. In four studies, the timespan was relatively short (2-4 weeks) and only occurred once (to guard against participants simply remembering the answers, Sayeski et al. (2015) mixed up the order of test items). However, three studies collected data over multiple years, testing teachers on their pedagogical content knowledge yearly for 2-4 years following graduation (Hiebert et al., 2019; Morris & Hiebert, 2017; Suppa, DiNapoli, & Mixell, 2018).

Participant and Researcher Bias

Of studies reviewed, 35.54% (n=59) used outcome measures that had stakes attached for participants, such as course assessments, course evaluations, participation grades, or high-stakes credential tests. As noted above, these measures are prone to bias as participants may have incentives to inflate their scores, or researchers may unknowingly bias grades in favor of finding effects. That said, using these measures is sometimes necessary given that pre-service teachers have limited capacity to undertake extra work outside class, and collecting additional data is resource intensive. Some researchers mitigated bias when using outcome measures with stakes by simply re-analysing the raw data using a different observation rubric and/or with a different focus. For example, Santagata & Yeh (2014) re-analyse videos of participants teaching that were

initially collected and submitted for the Performance Assessment for California Teachers Teaching Event (PACT), a high-stakes performance assessment tied to licensure requirements in California. Regardless of outcome measure, many study authors took extra measures to minimize the issue of stakes. For example, some researchers ensured that those rating and analysing the participant data were blind to treatment status (e.g., James & Scharmann, 2007; Schussler et al., 2017; Sayeski et al., 2017); and other researchers who doubled as course instructors explicitly informed participants that any responses to surveys or interviews would have no influence on grades (e.g., Morrison and McDuffie, 2009).

Of the 166 studies reviewed, only 4.22% (n=7) mitigated reactivity bias by blinding participants to their treatment condition and/or discussing potential biases from participants changing their behavior based on their or others' expectations. Studies did this differently depending on the practice being evaluated. Fives & Barnes (2017) evaluated the impact of providing pre-service teachers with additional guidance when developing assessments. As participants entered the room where the study was being conducted, they randomly received a package of papers: some included the additional guidance, some did not. As participants completed the task individually and were not told the purpose of the study, they were blind to the fact that some participants were receiving additional guidance. Giebelhaus & Bowman (2002) randomly assigned the practicum mentors of one group of pre-service teachers to additional training. Pre-service teachers were not advised about whether their practicum mentor was given the additional training. Crespo & Sinclair (2008) and Olson et al. (2016) evaluated the impact of various within-course activities on pre-service teachers' knowledge and skills. To address issues of reactivity, participants in both studies were only advised after the course concluded that they were part of a study and asked for consent. Finally, the remaining three studies used the within-

person comparison design described above, where researchers purposefully did not tell participants about the research design (i.e., to ensure that they did not respond differently based on the knowledge that they had only been taught some of the content being tested).

External Validity

Representativeness

Although nearly all studies reviewed provided some details about participant demographics and study site contextual factors, few discussed the representativeness of their study participants or sites so as to enable others to determine for whom and in what contexts findings may (or may not) generalize. Of the 166 studies reviewed, only 13.25% (n=22) compared study participant demographics to *either* the demographics of the teacher preparation program they were sampled from and/or the demographics of pre-service teachers in the US. Most discussions of participant representativeness were fairly straightforward. For example, Foley et al. (2017) evaluated a science methods course by randomly selecting 234 of the total 687 pre-service teachers who took the required course sometime between 2012-14. They then compared the sample with the broader population in terms of race and gender. Some studies, however, addressed the issue of representativeness through purposive sampling. For example, to ensure that their evaluation investigated effects across various levels of knowledge and abilities, Santagata & Yeh (2014) randomly selected four participants who scored near the average on a standardized teacher performance assessment (PACT), four who scored 1 SD above the mean, and four who scored 1 SD below (see also Matkins & Bell, 2007).

21.69% (n=36) of studies reviewed either named the study site(s) and/or described their location, size, type, and certification level to enable others to determine the relevance of study findings to other teacher preparation contexts. Addressing this issue is challenging given ethical

concerns and institutional policies can limit disclosure of contextual information lest participants become identifiable (Sikes, 2006). This is perhaps why studies with larger samples and/or conducted across multiple different institutions were more likely to report teacher preparation program characteristics or to identify the programs in which the study was conducted.

Scalability

158 studies (95.18%) evaluated teacher preparation practices that were administered by instructors (e.g., courses, pedagogies, or experiences). As noted above, this raises the question of scalability: unless studies are designed to address issues of external validity like instructor effects and researcher bias, then findings may not generalize. Of the 158 studies, only 10.13% (n=16) either trained multiple instructors to deliver the teacher preparation practice and/or discussed issues of using instructors that are highly enthusiastic about, skilled in, or suited to it. In most cases, these studies were multi-site evaluations where multiple instructors implemented the same practice across different course sections and/or teacher preparation programs. For example, Olson et al. (2016) evaluated how three instructors used various video cases across nine course sections of elementary science methods courses over the span of four semesters. To reduce instructor effects, study authors ensured all instructors taught at least one section across all study conditions, and that sections “used the same course syllabi, objectives, goals, and reading materials to ensure continuity in course content and sequence”. Some studies with multiple instructors collected additional implementation data to ensure fidelity (e.g., Bell, Matkins, & Gansneder, 2011). Further, some studies ensured at least one study author did not have any interactions with participants to ensure an “outsider” perspective for data collection and analysis (e.g., Bennet & Hart, 2014; Hart & Bennett, 2013; Hoppey & Mickelson, 2017).

Of the 139 studies that were undertaken within courses, 61.15% (n=85) reported that at

least one study author was also a course instructor. As noted above, evaluating one's own practices is often fraught as study authors may unknowingly bias data collection and analysis towards identifying positive findings. That said, evaluating one's own practices is often necessary given the funding, release time, and resources available to teacher education researchers. In these cases, some researchers guarded against issues of researcher bias by ensuring that data collection and analysis was not conducted by course instructors. For example, Hanuscin & Zangori (2016) note that "the consent process was conducted by another individual and the instructor did not know whether or not students consented until the conclusion of the semester". Similarly, Laframboise & Shea (2009) and Duncan, Pilitsis, & Piegaraol. (2010) collaborated with graduate students who conducted interviews with participants and did not share data or preliminary findings with the first author/course instructor until after the course concluded.

Checklist of Considerations

To aid future researchers in developing and reporting on impact evaluations of teacher preparation practices, I present a checklist of considerations in Table 2.5. This checklist largely reflects the issues with research design and measurement described above¹². For each consideration on the checklist, researchers are provided with a simple yes/no question to assess whether they are addressing the issue in their impact evaluation. If the answer is no, then researchers are presented with a follow-up open-ended question to assist them in mitigating the underlying problem. Teacher education researchers can use this checklist at any point during the development, implementation, and write-up of their impact evaluations. Given the complexity of

¹² Some considerations listed have not been discussed in this review; this is because they have not been raised by the field as a general issue, however are still important considerations about research design and measurement.

impact evaluations, the aim is to support researchers by presenting the wide range of considerations related to internal validity, measurement, and external validity in a concise and consolidated way (see also Hill, Mancenido, & Loeb, 2021). Peer reviewers, research funders, and policymakers can also use this checklist to demand greater transparency and accountability for more rigorous causal research methods in the field (Zeichner, 2005; Zientek et al., 2008).

Discussion and Conclusion

If we want high quality instruction for all students, we need well prepared teachers. To have well prepared teachers, we need to identify the most effective teacher preparation practices and get policymakers and teacher educators to implement them at scale. To identify the most effective teacher preparation practices and provide clear guidance to policymakers and teacher educators about what works, we need to increase the quantity and rigor of our impact evaluations. To increase the quantity and rigor of our impact evaluations, we need to make tractable and identify ways of addressing issues of research design and measurement that have been raised by teacher education researchers over the past two decades. This has been the aim of this review.

Like prior reviews, I find that we still have a lot of work to do as a field. Of the 166 studies reviewed, few used research designs that addressed issues of internal validity like maturation effects (27.71%; n=46), selection effects (12.04%; n=20), and history effects (14.46%; n=24). Further, few studies measured the impact of teacher preparation practices on pre-service teachers' instruction, whether observed or self-reported (24.01%; n=40), and even fewer measured the impact on student outcomes (3.01%; n=5). Less than half of the studies (45.18%; n=75) measured outcomes using reproducible measures, and only 18.67% (n=31) used established measures allowing for comparison of results across the field. And finally, few studies

discussed the representativeness of their study participants (13.25%; n=25) or study site(s) (21.69%; n=36) to support others in identifying for whom and in what contexts study findings may generalize.

These findings may be skewed by my narrow inclusion criteria (i.e., impact evaluations of teacher preparation practices focused on improving instructional quality) or search procedure (e.g., online search only through ERIC; hand search of only four top-rated journals). For example, I excluded studies that made causal claims in their discussion or conclusion sections (e.g., “these findings show that this practice improves pre-service teachers’ teaching skills”), but did not explicitly use causal language in their framing (e.g., “the aim of this study is to understand how pre-service teachers experience this practice”). Although I did this because I did not want to mischaracterize studies, it means that the number of studies reviewed underrepresents how many impact evaluations have been undertaken *in practice*. It also means that my findings may be inflated as studies that do not characterise themselves as impact evaluations but then make causal claims may be more likely to exhibit issues of internal validity, measurement, and external validity.

Improving the rigor of impact evaluations is challenging but not impossible. Above, I have highlighted a number of studies that have demonstrated innovative and promising approaches for addressing various issues of research design and measurement. I have also presented a checklist of considerations in Table 2.5 to provide teacher education researchers with more specific guidance on conducting impact evaluations of teacher education. In the sections below, I make more general recommendations about how we can improve as a field.

Focus on Causal Research

Despite being warned by prior reviews about the limited number of impact evaluations, and

despite the wide-ranging search terms and thousands of studies screened, I remain struck that only 166 studies were identified for inclusion in this review. As others have noted, this dearth of studies can be partly attributed to teacher education researchers' lack of dedicated funding and time (Cochran-Smith, 2004; Zeichner, 2005; Sleeter, 2012). However, calls for more impact evaluations of teacher preparation practices have been ongoing since 2002, and since then there has been a proliferation of causal research across education research broadly and cognate subfields (e.g., teacher professional development) as well as a significant increase in professional support to undertake this work (e.g., funding through the IES and NSF; establishment of societies and organisations like the *Society for Research on Educational Effectiveness*). Given these advancements elsewhere yet limited development within teacher education, we must own up to our deprivileging of causal work. We must acknowledge that over and above the many challenges of doing this work, many teacher education researchers have chosen not to undertake impact evaluations. Again, this is not to say that the dominant research traditions of the field over the past two decades have not provided useful insights (Borko, Liston, & Whitcomb, 2007; Cochran-Smith et al., 2016). Rather, it is simply to acknowledge that there continue to remain a number of answerable yet unanswered questions regarding the most effective teacher preparation practices, and that policymakers, teacher educators, and researchers will benefit if we shift our collective focus towards them.

Addressing this at a field-level requires clarifying our theories of action about how various types of research improve teacher preparation, and then re-aligning incentives so that teacher education researchers are more likely to choose to undertake this challenging yet important work. Teacher educators *are* interested in improving their practice. However, faced with the urgency and necessity of the day-to-day of teacher preparation, it is easier and more

efficient for teacher education researchers to undertake localised descriptive investigations of one's own practices, rather than more systematic and reproducible impact evaluations of distinct alternatives. But these localised descriptive investigations cannot be the basis for policymakers' decision-making on system-wide policies because they do not give insights into what practices are effective for whom and in what contexts. Only impact evaluations can help policymakers decide whether mandating certain program structures through accreditation requirements, or investing in grant programs that incentivize teacher educators to implement certain pedagogies will improve pre-service teacher learning at scale. However, while policymakers are interested in generating more causal evidence to inform their decision-making, they cannot undertake impact evaluations themselves because they do not run teacher preparation programs and cannot control what teacher educators do in their classrooms.

On a practical level, then, one way forward is to focus on ensuring that doctoral students and early career teacher education researchers have adequate training and support in not just qualitative and quantitative research methods, but also in research design and measurement (Borko Liston, & Whitcomb, 2007; Grossman, 2008; Wilson, 2006). As others have noted: "That many teacher educators have specialized in research methods most attuned for interpretive and practitioner studies means that as a field, teacher education has less ability to design studies that speak to policymakers' concerns and reflect teacher educators' deep knowledge about learning to teach" (Borko, Liston, & Whitcomb, 2007: p.9). Addressing this gap in training involves going beyond the standard prescription that 'research methods should be tailored to specific research questions', and providing dedicated courses that help researchers-in-training to identify causal questions and design studies that effectively answer them.

Measure Teacher Practice and Student Outcomes

Another recommendation stemming from this review is the need to develop more measures of teaching practice and student outcomes validated for use in the context of teacher education. This includes developing measures that go beyond general assessments of instructional quality, and can detect improvements in the specific domains or parts of instruction targeted by teacher preparation practices. Many researchers are already working on this, including through developing teaching simulations that require observational measures of discrete elements of instruction (Cohen et al., 2020; Shaughnessy & Boerst, 2018); or developing rubrics and rating schemes for assessing teaching artifacts, such as science teaching portfolios (Kloser et al., 2017; Martínez et al., 2012). Importantly, many of these measures are also starting to be publicly shared through repositories like EdInstruments (<http://edinstruments.com/>) to make it easier for researchers to find and use. Despite these developments, more needs to be done. More measures need to be developed, particularly in non-STEM teaching areas; and teacher education researchers need more encouragement to adopt and use established measures.

That said, even if we had better measures, teacher education researchers must still overcome the practical challenge of actually collecting data on teacher practice and student outcomes. Researchers could provide teachers with recording equipment and task them to collect video of themselves teaching; or alternatively, when teachers are asked to provide samples of instructional materials, they could also provide copies of strong, average, and weak student responses collected through their practicum. If this is done across multiple time points, researchers could assess improvements in teachers' practice and associated impact on student learning. Work like this has already gained some traction in the literature: Waggoner, et al. (2015) discuss how their program tasks all pre-service teachers to collect student-level pre- and post-test data for a unit they teach during practicum; and Brady, et al. (2016) describe how

student data from practica can be used to develop “value-added” measures that generate evidence of program effectiveness.

Another possibility is developing standardized instruments and procedures that can be used to evaluate teacher preparation practices that are not discipline-specific. For example, to get around the challenge of study participants being prepared for different subject areas and grade levels, Metcalf & Cruikshank (1991) developed thirty 15-minute “Reflective Teaching” lesson plans which could be prepared for and implemented by participants and then scored by trained raters using a common rubric. Similarly, Jacoby-Senghor, Sinclair, & Shelton (2016) provided participants with materials on Byzantine history and gave them 20-minutes to prepare a seven-minute lesson. Participants taught their lesson to someone who then took a content test. This enabled the authors to measure both instructional quality and student learning. Studies like this offer the prospect of being able to evaluate teacher preparation practices across the whole causal pathway from what teacher educators do, to what pre-service teachers’ learn, to how these teachers teach, to what their students learn.

Clear Guidance for Reporting on Studies

Like prior reviews (Cochran-Smith et al., 2012; Zientek et al., 2008; Zeichner, 2005), I find that studies generally underreport information about study participants, study site(s), and other contextual factors that influence how teacher preparation practices are implemented. This makes it difficult to determine to whom and in what contexts study findings generalize. I recommend, like others have before, that the field invest in developing shared vocabulary and norms for reporting (Borko, Liston & Whitcomb, 2007; Grossman & McDonald, 2008; Zeichner, 2005). This must go beyond the general prescriptions from the AERA Standards (2006; e.g., “The social, historical, or cultural context of the phenomena studied should also be described”),

and provide clarity to future researchers about what exactly to report and how. This will be challenging given the diversity of teacher education contexts and the myriad influences that contribute to pre-service teacher development. That said, without clear guidelines for reporting, teacher education researchers will continue to provide either too little (or in some cases, too much) detail, making it difficult for policymakers, teacher educators, and other researchers to compare findings across studies, and determine the relevance of findings across contexts.

Developing shared vocabulary and norms is a significant task, but can be made more tractable if teacher education researchers begin to undertake more linked small-scale studies (Grossman, 2008). As Sleeter (2014) describes: “researchers in different geographic locations... with careful planning and coordinating, could carry out linked small-scale studies that ask the same questions and use the same methodology”. This sort of collaboration will require developing a shared vocabulary and norms for describing the teacher preparation practice (so it can be implemented with fidelity) and reporting on site-specific contextual factors (so that researchers across institutions can identify and/or control for potential influences). From these efforts, other researchers could then conduct replication studies, adding more or less context depending on what they find is important for understanding the generalizability of findings across contexts. These sorts of “chains of inquiry around particular questions and consistently defined outcomes” (Zeichner, 2005: p.742) will enable teacher education researchers to achieve our collective goal of developing a shared evidence base.

Table 2.5
Considerations for designing and reporting on
impact evaluations of teacher preparation practices

It is challenging but necessary for teacher education researchers to evaluate the causal impacts of teacher preparation practices on pre-service teacher learning and/or student outcomes. The following list of considerations aims to support teacher education researchers in these efforts. Given the diverse and contingent contextual realities of teacher education programs, it is understandable that not all considerations may be able to be fully addressed by every study. As such, considerations have been listed in a loose hierarchy to help teacher education researchers prioritise their efforts.

<u>Consideration</u>	<u>(Y/N)</u>	<u>If N, then...</u>
<i>Internal validity</i>		
<i>1. The research design systematically analyses distinct alternatives</i>		
• A comparison groups design is used to account for natural learning over time or ‘business as usual’		If N, how will you know what would have happened to participants had they not received the teacher preparation practice? For example, how will you know that improvements aren’t just due to natural learning over time?
• Participants are not able to choose to be part of the treatment group (or choose not to be a part of it).		If N, how will you know that the effects aren’t caused by participants choosing to receive the treatment (or not) because it is better for them?
• During the study, the only difference between the treatment group and the control group is that the former receive the teacher preparation practice being evaluated.		If N, how will you know that the effects are due to the treatment and no other teacher preparation practice?
<i>Measurement</i>		
<i>1. Outcome measures are replicable</i>		
• Participant outcomes are measured using a clearly defined and consistent rubric or rating scheme		If N, how can you ensure that other researchers analysing the data would come to the same conclusions about participant outcomes?
• At least one measure used for this study has also been used in other studies		If N, how will you compare your study findings with other researchers’? How will you know how big or small any detected effects are?
<i>2. Outcome measures are fit-for-purpose</i>		
• Outcome measures accurately assess the study’s theoretically-defined constructs of interest.		If N, how will you know that any detected effects reflect changes in the target skill, knowledge, or disposition (rather than changes in some other related but not desired construct)?

Table 2.5 (Continued)

● Outcome measures can meaningfully differentiate between study participants' performance.	If N, how will you identify differences in outcomes between participants and ensure that any lack of study findings aren't due to noise?
● Outcome measures are fine-grained enough to detect study participants' growth (e.g., changes in the outcome due to the teacher preparation practice).	If N, how will you know that your outcome measures are able to detect any hypothesised effects?
<i>3. Effects on teacher practice and/or student learning are measured</i>	
● Outcome measures of pre-service teachers' practice and/or student learning are collected	If N, how will you know that the teacher preparation practice has effects on teacher practice and/or student learning?
● Outcome measures are collected over time	If N, how will you know whether effects persist?
<i>4. Participant and researcher bias is mitigated</i>	
● Outcome measures and how they are scored mitigates potential researcher or participant bias (e.g., multiple raters blind to treatment status are used)	If N, how will you know that findings are not influenced by the context in which outcomes are collected or the biases held by participants or researchers? For example, how will you know that any self-report outcomes are not influenced by participants' desirability or recall bias? Or that researchers are not scoring outcomes in a way that are biased towards finding an effect?
● Participants are blind to treatment status and/or not given undue expectations about the effects of the teacher preparation practice (e.g. they are not told to expect a particular outcome, or given cues to behave differently than they would ordinarily).	If N, how will you know that any effects detected are not due to participants changing their behaviour as a result of being studied? For example, how will you know that any effects detected are not placebo effects?
<i>External validity</i>	
<i>1. The representativeness of the study sample is described</i>	
● The demographics of study participants are compared to either the demographics of the teacher preparation program they are sampled from and/or the demographics of pre-service teachers in the US.	If N, how will you identify the broader population that study findings may generalize to?
● The study site(s) were named and/or their location, size, program type, and certification level were described.	If N, how will you identify potential contextual factors that may be influencing your results so that you and others can determine whether your study findings are applicable to other contexts?

Table 2.5 (Continued)

<i>2. The study is designed for scalability</i>	
● Multiple instructors are trained to deliver the teacher preparation practice.	If N, how will you know that study findings can be replicated by other teacher educators in other contexts?
● The teacher preparation practice being evaluated is not being delivered by one of the study authors in one of their courses.	If N, how will you know that study findings are not due to the specific behaviours or context of the study authors? For example, how will you know that findings may not just be limited to someone highly knowledgeable, enthusiastic, and/or suited to the teacher preparation practice?

References

- Allen, S. J., & Blackston, A. R. (2003). Training preservice teachers in collaborative problem solving: An investigation of the impact on teacher and student behavior change in real-world settings. *School Psychology Quarterly*, 18(1), 22–51. <https://doi.org/10.1521/scpq.18.1.22.20878>
- American Educational Research Association. (2006). Standards for Reporting on Empirical Social Science Research in AERA Publications. *Educational Researcher*, 35(6), 33–40. <https://doi.org/10.3102/0013189X035006033>
- Ballou, D., & Podgursky, M. (2000). Reforming teacher preparation and licensing: What is the evidence? *Teachers College Record*, 102(1), 5–27.
- Bangel, N. J., Moon, S. M., & Capobianco, B. M. (2010). Preservice Teachers' Perceptions and Experiences in a Gifted Education Training Model. *Gifted Child Quarterly*, 54(3), 209–221. <https://doi.org/10.1177/0016986210369257>
- Bartell, T. G., Webel, C., Bowen, B., & Dyson, N. (2013). Prospective teacher learning: Recognizing evidence of conceptual understanding. *Journal of Mathematics Teacher Education*, 16(1), 57–79. <https://doi.org/10.1007/s10857-012-9205-4>
- Bennett, S., & Hart, S. (2014). Addressing the 'Shift': Preparing Preservice Secondary Teachers for the Common Core. *Reading Horizons: A Journal of Literacy and Language Arts*, 53(4).
- Borko, H., Liston, D., & Whitcomb, J. A. (2007). Genres of Empirical Research in Teacher Education. *Journal of Teacher Education*, 58(1), 3–11. <https://doi.org/10.1177/0022487106296220>
- Boyd, D., Grossman, P. L., Hammerness, K., Lankford, R. H., Loeb, S., McDonald, M., Reininger, M., Ronfeldt, M., & Wyckoff, J. (2008). Surveying the Landscape of Teacher Education in New York City: Constrained Variation and the Challenge of Innovation. *Educational Evaluation and Policy Analysis*, 30(4), 319–343. <https://doi.org/10.3102/0162373708322737>
- Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher Preparation and Student Achievement. *Educational Evaluation and Policy Analysis*, 31(4), 416–440.
- Brady, M. P., Miller, K., McCormick, J., & Heiser, L. A. (2018). A Rational and Manageable Value-Added Model for Teacher Preparation Programs. *Educational Policy*, 32(5), 728–750. <https://doi.org/10.1177/0895904816673741>
- Bullough, R. V., Young, J., Birrell, J. R., Cecil Clark, D., Winston Egan, M., Erickson, L., Frankovich, M., Brunetti, J., & Welling, M. (2003). Teaching with a peer: A comparison of two models of student teaching. *Teaching and Teacher Education*, 19(1), 57–73. [https://doi.org/10.1016/S0742-051X\(02\)00094-X](https://doi.org/10.1016/S0742-051X(02)00094-X)

- Bulunuz, N., & Jarrett, O. S. (2009). Understanding of Earth and Space Science Concepts: Strategies for Concept-Building in Elementary Teacher Preparation. *School Science and Mathematics*, 109(5), 276–289. <https://doi.org/10.1111/j.1949-8594.2009.tb18092.x>
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally & Company.
- Cartwright, T., Smith, S., & Hallar, B. (2014). Confronting Barriers to Teaching Elementary Science: After-School Science Teaching Experiences for Preservice Teachers. *Teacher Education & Practice*, 27(2/3), 464–487.
- Caughlan, S., Juzwik, M. M., Borsheim-Black, C., Kelly, S., & Fine, J. G. (2013). English Teacher Candidates Developing Dialogically Organized Instructional Practices. *Research in the Teaching of English*, 47(3), 212–246.
- Clift, R. T., & Brady, P. (2005). Research on Methods Courses and Field Experiences. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying Teacher Education: The report of the AERA Panel on Research and Teacher Education* (pp. 309–424). Lawrence Erlbaum Associates, Inc.
- Cochran-Smith, M. (2004). Ask a Different Question, Get a Different Answer: The Research Base for Teacher Education. *Journal of Teacher Education*, 55(2), 111–115. <https://doi.org/10.1177/0022487104262971>
- Cochran-Smith, M. (2005). Studying Teacher Education: What We Know and Need to Know. *Journal of Teacher Education*, 56(4), 301–306. <https://doi.org/10.1177/0022487105280116>
- Cochran-Smith, M., Baker, M., Burton, S., Chang, W.-C., Carney, M. C., Fernández, M. B., Keefe, E. S., Miller, A. F., & Sánchez, J. G. (2017). The accountability era in US teacher education: Looking back, looking forward. *European Journal of Teacher Education*, 40(5), 572–588. <https://doi.org/10.1080/02619768.2017.1385061>
- Cochran-Smith, M., Cannady, M., Meachern, K. P., Mitchell, K., & Piazza, P. (2012). Teachers' Education and Outcomes: Mapping the Research Terrain. *Teachers College Record*, 114(10), 1–49.
- Cochran-Smith, M., & Villegas, A. M. (2015). Framing Teacher Preparation Research: An Overview of the Field, Part I. *Journal of Teacher Education*, 66(1), 7–20. <https://doi.org/10.1177/0022487114549072>
- Cochran-Smith, M., Villegas, A. M., Abrams, L., Chavez-Moreno, L., Mills, T., & Stern, R. (2015). Critiquing Teacher Preparation Research: An Overview of the Field, Part II. *Journal of Teacher Education*, 66(2), 109–121. <https://doi.org/10.1177/0022487114558268>
- Cochran-Smith, M., Villegas, A. M., Abrams, L. W., Chavez-Moreno, L. C., Mills, T., & Stern, R. (2016). Research on Teacher Preparation: Charting the Landscape of a Sprawling Field. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of Research on Teaching* (Fifth, pp. 439–547).

American Educational Research Association. https://doi.org/10.3102/978-0-935302-48-6_7

Cochran-Smith, M., & Zeichner, K. M. (2005). *Studying teacher education: The report of the AERA Panel on research and teacher education*. Lawrence Erlbaum Associates, Inc. <https://doi.org/10.4324/9780203864043>

Cohen, J., Wong, V., Krishnamachari, A., & Berlin, R. (2020). Teacher Coaching in a Simulated Environment. *Educational Evaluation and Policy Analysis*, 42(2), 208–231. <https://doi.org/10.3102/0162373720906217>

Conaway, C., & Goldhaber, D. (2020). Appropriate Standards of Evidence for Education Policy Decision Making. *Education Finance and Policy*, 15(2), 383–396. https://doi.org/10.1162/edfp_a_00301

Copur-Gencturk, Y., Cimpian, J. R., Lubienski, S. T., & Thacker, I. (2020). Teachers' Bias Against the Mathematical Ability of Female, Black, and Hispanic Students. *Educational Researcher*, 49(1), 30–43. <https://doi.org/10.3102/0013189X19890577>

Crawford, R., & Tan, H. (2019). Responding to evidence-based practice: An examination of mixed methods research in teacher education. *The Australian Educational Researcher*, 46(5), 775–797. <https://doi.org/10.1007/s13384-019-00310-w>

Crespo, S., & Sinclair, N. (2008). What makes a problem mathematically interesting? Inviting prospective teachers to pose better problems. *Journal of Mathematics Teacher Education*, 11(5), 395–415. <https://doi.org/10.1007/s10857-008-9081-0>

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Vasquez Heilig, J. (2005). Does Teacher Preparation Matter? Evidence about Teacher Certification, Teach for America, and Teacher Effectiveness. *Education Policy Analysis Archives*, 13(42). <https://doi.org/10.14507/epaa.v13n42.2005>

Del Schalock, H., Schalock, M. D., & Ayres, R. (2006). Scaling Up Research in Teacher Education: New Demands on Theory, Measurement, and Design. *Journal of Teacher Education*, 57(2), 102–119. <https://doi.org/10.1177/0022487105285615>

Diez, M. E. (2010). It Is Complicated: Unpacking the Flow of Teacher Education's Impact on Student Learning. *Journal of Teacher Education*, 61(5), 441–450. <https://doi.org/10.1177/0022487110372927>

Duncan, R. G., Pilitsis, V., & Piegaro, M. (2010). Development of Preservice Teachers' Ability to Critique and Adapt Inquiry-based Instructional Materials. *Journal of Science Teacher Education*, 21(1), 81–102. <https://doi.org/10.1007/s10972-009-9153-8>

Feuer, M. J., Floden, R. E., Chudowsky, N., & Ahn, J. (2013). *Evaluation of Teacher Preparation Programs: Purposes, Methods, and Policy Options*. National Academy of Education. <https://files.eric.ed.gov/fulltext/ED565694.pdf>

- Fives, H., & Barnes, N. (2017). Informed and Uninformed Naïve Assessment Constructors' Strategies for Item Selection. *Journal of Teacher Education*, 68(1), 85–101. <https://doi.org/10.1177/0022487116668019>
- Floden, R., & Meniketti, M. (2005). Research on the Effects of Coursework in the Arts and Sciences and in the Foundations of Education. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying Teacher Education: The report of the AERA Panel on Research and Teacher Education* (pp. 261–308). Lawrence Erlbaum Associates, Inc.
- Foley, R. W., Archambault, L. M., Hale, A. E., & Dong, H.-K. (2017). Learning Outcomes in Sustainability Education Among Future Elementary School Teachers. *Journal of Education for Sustainable Development*, 11(1), 33–51. <https://doi.org/10.1177/0973408217725861>
- Forzani, F. M. (2014). Understanding “Core Practices” and “Practice-Based” Teacher Education: Learning From the Past. *Journal of Teacher Education*, 65(4), 357–368. <https://doi.org/10.1177/0022487114533800>
- Gainsburg, J. (2012). Why new mathematics teachers do or don't use practices emphasized in their credential program. *Journal of Mathematics Teacher Education*, 15(5), 359–379. <https://doi.org/10.1007/s10857-012-9208-1>
- Giebelhaus, C. R., & Bowman, C. L. (2002). Teaching Mentors: Is It Worth the Effort? *The Journal of Educational Research*, 95(4), 246–254. <https://doi.org/10.1080/00220670209596597>
- Girod, M., & Girod, G. (2006). Exploring the Efficacy of the Cook School District Simulation. *Journal of Teacher Education*, 57(5), 481–497. <https://doi.org/10.1177/0022487106293742>
- Goodlad, J. I. (1990). *Teachers for our nation's schools* (1st ed.). Jossey-Bass.
- Grossman, P. (2008). Responding to Our Critics: From Crisis to Opportunity in Research on Teacher Education. *Journal of Teacher Education*, 59(1), 10–23. <https://doi.org/10.1177/0022487107310748>
- Grossman, P., & McDonald, M. (2008). Back to the Future: Directions for Research in Teaching and Teacher Education. *American Educational Research Journal*, 45(1), 184–205. <https://doi.org/10.3102/0002831207312906>
- Hanuscin, D. L., & Zangori, L. (2016). Developing Practical Knowledge of the Next Generation Science Standards in Elementary Science Teacher Education. *Journal of Science Teacher Education*, 27(8), 799–818. <https://doi.org/10.1007/s10972-016-9489-9>
- Hart, L. C. (2002). Preservice Teachers' Beliefs and Practice After Participating in an Integrated Content/Methods Course. *School Science and Mathematics*, 102(1), 4–14. <https://doi.org/10.1111/j.1949-8594.2002.tb18191.x>

- Hart, S. M., & Bennett, S. M. (2013). Disciplinary Literacy Pedagogy Development of STEM Preservice Teachers. *Teacher Education & Practice*, 26(2), 221–241.
- Hiebert, J., Berk, D., Miller, E., Gallivan, H., & Meikle, E. (2019). Relationships Between Opportunity to Learn Mathematics in Teacher Preparation and Graduates' Knowledge for Teaching Mathematics. *Journal for Research in Mathematics Education*, 50(1), 23–50. <https://doi.org/10.5951/jresematheduc.50.1.0023>
- Hill, H. C., & Chin, M. (2018). Connections Between Teachers' Knowledge of Students, Instruction, and Achievement Outcomes. *American Educational Research Journal*, 55(5), 1076–1112. <https://doi.org/10.3102/0002831218769614>
- Hill, H. C., Mancenido, Z., & Loeb, S. (2021). Effectiveness Research for Teacher Education. (EdWorkingPaper: 20-252). In *EdWorkingPapers.com*. <https://doi.org/10.26300/zhhb-j781>
- Hoppey, D., & Mickelson, A. M. (2017). Partnership and Coteaching: Preparing Preservice Teachers to Improve Outcomes for Students with Disabilities. *Action in Teacher Education*, 39(2), 187–202. <https://doi.org/10.1080/01626620.2016.1273149>
- Institute of Education Sciences. (2020). *What Works Clearinghouse™ Standards Handbook Version 4.0*. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf
- Isabelle, A. D., & de Groot, C. (2008). Alternate Conceptions of Preservice Elementary Teachers: The Itakura Method. *Journal of Science Teacher Education*, 19(5), 417–435. <https://doi.org/10.1007/s10972-008-9105-8>
- Jacob, R., Hill, H., & Corey, D. (2017). The Impact of a Professional Development Program on Teachers' Mathematical Knowledge for Teaching, Instruction, and Student Achievement. *Journal of Research on Educational Effectiveness*, 10(2), 379–407. <https://doi.org/10.1080/19345747.2016.1273411>
- Jacoby-Senghor, D. S., Sinclair, S., & Shelton, J. N. (2016). A lesson in bias: The relationship between implicit racial bias and performance in pedagogical contexts. *Journal of Experimental Social Psychology*, 63, 50–55. <https://doi.org/10.1016/j.jesp.2015.10.010>
- James, M. C., & Scharmann, L. C. (2007). Using analogies to improve the teaching performance of preservice teachers. *Journal of Research in Science Teaching*, 44(4), 565–585. <https://doi.org/10.1002/tea.20167>
- Kennedy, M. M. (2007). Defining a Literature. *Educational Researcher*, 36(3), 139–147. <https://doi.org/10.3102/0013189X07299197>
- King, G., Keohane, R. O., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton University Press.

Kloser, M., Borko, H., Martínez, J. F., Stecher, B., & Luskin, R. (2017). Evidence of Middle School Science Assessment Practice From Classroom-Based Portfolios. *Science Education*, 101(2), 209–231. <https://doi.org/10.1002/sce.21256>

Kopcha, T. J., & Alger, C. (2011). The impact of technology-enhanced student teacher supervision on student teacher knowledge, performance, and self-efficacy during the field experience. *Journal of Educational Computing Research*, 45(1), 49–73. <https://doi.org/10.2190/EC.45.1.c>

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. *What Works Clearinghouse*. <https://files.eric.ed.gov/fulltext/ED510743.pdf>

Labaree, D. F. (2004). *The Trouble with Ed Schools*. Yale University Press.

Laframboise, K. L., & Shea, K. (2009). Developing Understanding of Research-based Pedagogy with Preservice Teachers: An Instrumental Case Study. *The Qualitative Report*, 14(1), 105–128.

Lambert, J. L., & Bleicher, R. E. (2013). Climate Change in the Preservice Teacher's Mind. *Journal of Science Teacher Education*, 24(6), 999–1022. <https://doi.org/10.1007/s10972-013-9344-1>

Livingston, K., & Flores, M. A. (2017). Trends in teacher education: A review of papers published in the European journal of teacher education over 40 years. *European Journal of Teacher Education*, 40(5), 551–560. <https://doi.org/10.1080/02619768.2017.1387970>

Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Clarendon Press.

Mahalingappa, L., Hughes, E. M., & Polat, N. (2018). Developing Preservice Teachers' Self-Efficacy and Knowledge through Online Experiences with English Language Learners. *Language and Education*, 32(2), 127–146.

Martínez, J. F., Borko, H., & Stecher, B. M. (2012). Measuring instructional practice in science using classroom artifacts: Lessons learned from two validation studies. *Journal of Research in Science Teaching*, 49(1), 38–67. <https://doi.org/10.1002/tea.20447>

Matkins, J. J., & Bell, R. L. (2007). Awakening the Scientist Inside: Global Climate Change and the Nature of Science in an Elementary Science Methods Course. *Journal of Science Teacher Education*, 18(2), 137–163. <https://doi.org/10.1007/s10972-006-9033-4>

Matthews, M. E., & Seaman, W. I. (2007). The Effects of Different Undergraduate Mathematics Courses on the Content Knowledge and Attitude towards Mathematics of Preservice Elementary Teachers. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 1.

Mayer, D. (2021). The connections and disconnections between teacher education policy and research: Reframing evidence. *Oxford Review of Education*, 47(1), 120–134.

<https://doi.org/10.1080/03054985.2020.1842179>

McGee, J., & Colby, S. (2014). Impact of an Assessment Course on Teacher Candidates' Assessment Literacy. *Action in Teacher Education*, 36(5–6), 522–532. <https://doi.org/10.1080/01626620.2014.977753>

Menter, I., Hulme, M., Murray, J., Campbell, A., Hextall, I., Jones, M., Mahony, P., Procter, R., & Wall, K. (2010). Teacher education research in the UK: The state of the art. *Swiss Journal of Educational Research*, 32(1), 121–142. <https://doi.org/10.24452/sjer.32.1.4829>

Metcalf, K. K., & Cruickshank, D. R. (1991). Can Teachers Be Trained to Make Clear Presentations? *The Journal of Educational Research*, 85(2), 107–116. <https://doi.org/10.1080/00220671.1991.10702820>

Mizell, J. A., & Cates, J. (2004). The Impact of Additional Content Courses on Teacher Candidates' Beliefs regarding Mathematics Content and Pedagogy. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 4.

Morris, A. K., & Hiebert, J. (2017). Effects of Teacher Preparation Courses: Do Graduates Use What They Learned to Plan Mathematics Lessons? *American Educational Research Journal*, 54(3), 524–567. <https://doi.org/10.3102/0002831217695217>

Morrison, J., & McDuffie, A. R. (2009). Connecting Science and Mathematics: Using Inquiry Investigations to Learn About Data Collection, Analysis, and Display. *School Science and Mathematics*, 109(1), 31–44. <https://doi.org/10.1111/j.1949-8594.2009.tb17860.x>

Murnane, R. J., & Willett, J. B. (2010). *Methods Matter: Improving Causal Inference in Educational and Social Science Research*. Oxford University Press.

Nokes, J. D., Bullough, R. V., Egan, W. M., Birrell, J. R., & Merrell Hansen, J. (2008). The paired-placement of student teachers: An alternative to traditional placements in secondary schools. *Teaching and Teacher Education*, 24(8), 2168–2177. <https://doi.org/10.1016/j.tate.2008.05.001>

Nuttall, J., Murray, S., Seddon, T., & Mitchell, J. (2006). Changing Research Contexts in Teacher Education in Australia: Charting new directions. *Asia-Pacific Journal of Teacher Education*, 34(3), 321–332. <https://doi.org/10.1080/13598660600927224>

Olson, J. K., Bruxvoort, C. N., & Vande Haar, A. J. (2016). The impact of video case content on preservice elementary teachers' decision-making and conceptions of effective science teaching. *Journal of Research in Science Teaching*, 53(10), 1500–1523. <https://doi.org/10.1002/tea.21335>

Piasta, S. B., Connor, C. M., Fishman, B. J., & Morrison, F. J. (2009). Teachers' Knowledge of Literacy Concepts, Classroom Practices, and Student Reading Growth. *Scientific Studies of Reading*, 13(3), 224–248. <https://doi.org/10.1080/10888430902851364>

- Polanin, J. R., Maynard, B. R., & Dell, N. A. (2017). Overviews in Education Research: A Systematic Review and Analysis. *Review of Educational Research*, 87(1), 172–203. <https://doi.org/10.3102/0034654316631117>
- Ragland, R. (2016). Implementing an Evidence-Based Reflective Teaching Cycle: Using Scholarly Research in Curriculum Design. *Mid-Western Educational Researcher*, 28(3), 196–217.
- Rodriguez-Valls, F., & Ponce, G. (2013). Classroom, the We Space: Developing Student-Centered Practices for Second Language Learner (SLL) Students. *Education Policy Analysis Archives*, 21, 55. <https://doi.org/10.14507/epaa.v21n55.2013>
- Santagata, R., & Yeh, C. (2014). Learning to teach mathematics and to analyze teaching effectiveness: Evidence from a video- and practice-based approach. *Journal of Mathematics Teacher Education*, 17(6), 491–514. <https://doi.org/10.1007/s10857-013-9263-2>
- Sayeski, K. L., Earle, G. A., Eslinger, R. P., & Whintenton, J. N. (2017). Teacher candidates' mastery of phoneme-grapheme correspondence: Massed versus distributed practice in teacher education. *Annals of Dyslexia*, 67(1), 26–41. <https://doi.org/10.1007/s11881-016-0126-2>
- Sayeski, K. L., Kennedy, M. J., Irala, S. de, Clinton, E., Hamel, M., & Thomas, K. (2015). The Efficacy of Multimedia Modules for Teaching Basic Literacy-Related Concepts. *Exceptionality*, 23(4), 237–257. <https://doi.org/10.1080/09362835.2015.1064414>
- Schussler, D., Frank, J., Lee, T.-K., & Mahfouz, J. (2017). Using Virtual Role-Play to Enhance Teacher Candidates' Skills in Responding to Bullying. *Journal of Technology and Teacher Education*, 25(1), 30.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Houghton Mifflin.
- Shaughnessy, M., & Boerst, T. A. (2018). Uncovering the Skills That Preservice Teachers Bring to Teacher Education: The Practice of Eliciting a Student's Thinking. *Journal of Teacher Education*, 69(1), 40–55. <https://doi.org/10.1177/0022487117702574>
- Sikes, P. (2006). On dodgy ground? Problematics and ethics in educational research. *International Journal of Research & Method in Education*, 29(1), 105–117. <https://doi.org/10.1080/01406720500537502>
- Slavin, R. E. (1986). Best-Evidence Synthesis: An Alternative to Meta-Analytic and Traditional Reviews. *Educational Researcher*, 15(9), 5–11. <https://doi.org/10.3102/0013189X015009005>
- Sleeter, C. (2014). Toward Teacher Education Research That Informs Policy. *Educational Researcher*, 43(3), 146–153. <https://doi.org/10.3102/0013189X14528752>
- Smith, M. E., Swars, S. L., Smith, S. Z., Hart, L. C., & Haardorfer, R. (2012). Effects of an

Additional Mathematics Content Course on Elementary Teachers' Mathematical Beliefs and Knowledge for Teaching. *Action in Teacher Education*, 34(4), 336–348. eric.

Soh, L.-K., Fowler, D., & Zygielbaum, A. I. (2007). The Impact of the Affinity Learning Authoring Tool on Student Learning. *Journal of Educational Technology Systems*, 36(1), 29–62. <https://doi.org/10.2190/ET.36.1.d>

Stover, K., Yearta, L. S., & Sease, R. (2014). “Experience Is the Best Tool for Teachers”: Blogging to Provide Preservice Educators with Authentic Teaching Opportunities. *Journal of Language and Literacy Education*, 10(2), 99–117.

Sun, J., & van Es, E. A. (2015). An Exploratory Study of the Influence That Analyzing Teaching Has on Preservice Teachers' Classroom Practice. *Journal of Teacher Education*, 66(3), 201–214.

Suppa, S., DiNapoli, J., & Mixell, R. (2018). Teacher Preparation “Does” Matter: Relationships between Elementary Mathematics Content Courses and Graduates' Analyses of Teaching. *Mathematics Teacher Education and Development*, 20(2), 25–57.

Thomas, A. F., & Sondergeld, T. (2015). Investigating the Impact of Feedback Instruction: Partnering Preservice Teachers with Middle School Students to Provide Digital, Scaffolded Feedback. *Journal of the Scholarship of Teaching and Learning*, 15(4), 83–109.

Waggoner, J., Carroll, J. B., Merk, H., & Weitzel, B. N. (2015). Critics and Critical Analysis: Lessons from 19,000 P-12 Students in Candidates' Classrooms. *Teacher Education Quarterly*, 42(1), 97–108.

Weinburgh, M. (2007). The Effect of Tenebrio obscurus on Elementary Preservice Teachers' Content Knowledge, Attitudes, and Self-efficacy. *Journal of Science Teacher Education*, 18(6), 801–815. <https://doi.org/10.1007/s10972-007-9073-4>

Welsh, K. A., & Schaffer, C. (2017). Developing the Effective Teaching Skills of Teacher Candidates during Early Field Experiences. *Educational Forum*, 81(3), 301–321.

Whitehurst, G. J. R. (2003). The Institute of Education Sciences: New Wine, New Bottles. *American Educational Research Association 2003 Annual Meeting Presidential Invited Session*.

Wilkins, J. L. M., & Brand, B. R. (2004). Change in Preservice Teachers' Beliefs: An Evaluation of a Mathematics Methods Course. *School Science and Mathematics*, 104(5), 226–232. <https://doi.org/10.1111/j.1949-8594.2004.tb18245.x>

Wilson, S. M. (2006). Finding a Canon and Core: Meditations on the Preparation of Teacher Educator-Researchers. *Journal of Teacher Education*, 57(3), 315–325. <https://doi.org/10.1177/0022487105285895>

Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2002). Teacher Preparation Research: An Insider's View from the Outside. *Journal of Teacher Education*, 53(3), 190–204.

<https://doi.org/10.1177/0022487102053003002>

Windschitl, M., Thompson, J., & Braaten, M. (2008). How Novice Science Teachers Appropriate Epistemic Discourses Around Model-Based Inquiry for Use in Classrooms. *Cognition and Instruction*, 26(3), 310–378. <https://doi.org/10.1080/07370000802177193>

Wineburg, M. S. (2006). Evidence in Teacher Preparation: Establishing a Framework for Accountability. *Journal of Teacher Education*, 57(1), 51–64. <https://doi.org/10.1177/0022487105284475>

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *Trans. Comput. Educ.*, 14(1), 5:1-5:16. <https://doi.org/10.1145/2576872>

Zeichner, K. (2005). A research agenda for teacher education. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying Teacher Education: The report of the AERA Panel on Research and Teacher Education* (pp. 737–759). Lawrence Erlbaum Associates, Inc.

Zeichner, K., & Conklin, H. G. (2005). Teacher Education Programs. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying Teacher Education: The report of the AERA Panel on Research and Teacher Education* (pp. 645–735). Lawrence Erlbaum Associates, Inc.

Zientek, L. R., Capraro, M. M., & Capraro, R. M. (2008). Reporting Practices in Quantitative Teacher Education Research: One Look at the Evidence Cited in the AERA Panel Report. *Educational Researcher*, 37(4), 208–216. <https://doi.org/10.3102/0013189X08319762>

Conclusion

The three articles presented in this dissertation aim to help policymakers and teacher educators in improving teacher quality: whether through attracting more academic high achievers into the teaching profession, or through identifying effective teacher preparation practices.

In Paper 1, I find that academic high achievers who chose to teach were intrinsically/altruistically motivated *and* had beliefs and plans that addressed their extrinsic concerns about the low pay, prestige, and long-term career options of teaching. In Paper 2, I find that academic high achievers are socialised through both explicit and implicit signals (primarily from parents and peers) to believe that teaching is not an appropriate career. Together these papers suggest the possibility of attracting more academic high achievers into teaching through policies and programs that target them early in the career exploration process (i.e., in high school and early college) and that focus on developing their intrinsic/altruistic motivations and/or fostering beliefs and plans that directly address prevalent extrinsic concerns. In Paper 3, I review impact evaluations of teacher preparation practices over the past two decades to better understand one of the primary challenges faced by the field: research design and measurement. Although I find that very few studies address issues of internal validity, external validity and measurement, I highlight some innovative approaches and present a checklist of considerations to assist future researchers in designing more rigorous impact evaluations.

Across these three papers has been an underlying commitment to identifying *scalable practices*, actions that can be easily learned from and/or implemented by individuals to improve how we attract and prepare teachers. These individuals can be *practitioners* – e.g., teachers encouraging their high-achieving students to learn more about and become interested in teaching as a career; teacher educators adapting and implementing practices that have been shown to have

an impact elsewhere. These individuals can also be *policymakers* – e.g., state education departments incentivizing (through grants) or mandating (through accreditation) teacher educators to implement certain practices that have a demonstrated impact on pre-service teacher learning; school systems creating mentoring programs and scholarships that funnel academic high-achievers towards teaching.

This work is complex. Finding ways to better attract and prepare teachers is challenging given that the structures and practices of teacher preparation programs across the US are so varied, and given that the contexts within which they operate are so diverse. That said, just because our structures, practices, and contexts are different, it does not mean that we cannot learn from each other by systematically collecting and analysing data in ways that enable us to identify generalisable insights. This is the promise of rigorous educational research – and I look forward to continuing to contribute to this community of practice in my future work.

Appendix A

Example codes and indicative quotes for Paper #1

Influence of family and friends

I definitely thought maybe my parents would think, "Oh, we didn't pay \$250,000 for you to go make \$40,000 a year." I have a very complicated relationship with them but they didn't really say anything negative to me. My dad was excited because he's a professor. He was like, "Teaching is such a noble profession." - Jessica

Yes my family or close friends ask me why I'm choosing to become a teacher but I will say that maybe it's different because they know I want to go to law school, so I think they may see it a little differently. It's like, I think if I, maybe if I propose like, "I'm just going to stay in the classroom" or say "no I want to be a teacher", I think I could imagine some of my family members questioning like, "Is that really what you want to do? You went to Penn, and spent all this money to go to Penn and you didn't choose to become a lawyer or doctor or whatnot?" - Tabitha

I understand the difficulties of coming from Penn. Everyone's making six figures when they come out and you're scraping by, and people are wondering why you do it. There are also people who are like, "You're so smart. Why are you doing this?" That's not easy. That's not super easy, but I have enough of the support system. I've always been very strong in my opinions and my ways. If I want to do something, I'm going to do it. - Rory

[Family] would be like, "You love kids. Do you want to go into pediatrics?" No one ever said like, "You love kids. Do you want to be a teacher?" - Johanna

Influence of teachers and professors

I think them sharing how lost they were when they were my age kind of really helped me. I just needed someone who was older, who has been through what I'm going through to tell me that it's going to be fine basically. - Li

I was only there an hour a week so you don't get a great picture of what it'd be like to be a teacher full time... but I feel like it was very disrespectful for the teacher, even with someone else like me present, the kids didn't want to act appropriately. Just the way students were acting, the things they would say... The teachers were always struggling to get them to listen and to actually be interested in learning. And it wasn't all the students... you just have a couple in the class who act that way, and then it kind of takes away from the entire class. - Micah

My teachers, they had the same responses for me. They were like, "We don't want to pigeonhole you as a teacher because we feel like you can do other things." - Charles

Don't want to teach because teachers are paid poorly

I was like, "Why am I trying to do something that's not fulfilling to me? Why am I trying to, I don't know, push myself to go into a profession--" Because when I was like, "Okay, why do I want to do it," it was because of money. It was because I don't want to struggle. My parents did. I don't want my children to have to struggle. - Rory

I come from Michigan which has just completely defunded public education, but also I know teachers, I talk to teachers-- they don't get paid what they deserve to be paid. It's just hard to make ends meet as a teacher if you have a family, particularly as a single breadwinner. I feel like a fancy [elite college] degree or two fancy [elite college] degrees are not going to protect you from that fate. - Nicholas

[Mentor] was like, "What do you want to do in life, basically." I was like, "Honestly, I want to become a teacher." He was like, "You can do that because you're a woman. My wife is a teacher. She doesn't make 100K a year, but since I make a lot of money it works and it's fine." Then in my head I was just like, "See I don't want that. I don't want to do that." - Divya

How much teachers are paid doesn't matter

Living on a teacher's salary for a few years is better than me doing other Americorps programs where there's absolutely no money. Like, you're doing a lot of great things, a lot of great impact, a lot of leadership development but you're not paid enough to live. At least on a teacher's salary, I can live and be comfortable and have a little spending money. - Faith

I'm obviously not doing teaching for money. I'm doing it so that a kid will understand something that they didn't know before doing it, so that a kid can see themselves growing intellectually doing it, so that a kid can go to college. - Kyle

At a career event I went to there was someone who had gone to Chicago and become a teacher.

He told the story about how he was rooming with someone who went into investment banking or something like that. He said that one day his roommate came home, just looked at him, and asked, "What is it like to enjoy what you do?" I think about that a lot. I know that I probably won't be making as much money as a lot of the people around me, and that won't precisely be fun, but it's more important that I'd be doing something that I enjoy and that I can feel good about." - Francis

Wanting to do something that has an impact

I was doing tech thinking in Silicon Valley and it was a really good job and obviously it paid well and I loved it. There wasn't a deeper purpose in my opinion, I'm sure you know what I mean. I didn't feel like I was making anyone better off that was in a position to need help necessarily. I didn't feel like I was making America a better place in terms of socio-economic inequality anything like that. I didn't feel a higher purpose attached to the work I was doing. - Ryan

To be a privileged guy coming from where I come from without really a care in the world in comparison to others, and being able to make a lot of money because I happen to be well educated and pretty quick to be able to write a sentence. I think any one in my position has to think very hard about your duties and obligations as a person in society given all the things that you've been given. - Nicholas

I just felt this thrilled feeling like, "Yes, I'm going to Los Vegas. I got high school. There's so much I can do there. There is so much impact I can have and I get to try something new. - Jessica

Elite campus career pressures

I don't think I was ever explicitly pressured to do anything in particular or to have a specific idea of what I wanted to do. But people came back to school after their junior summer and so many had a really good idea that they were going to do finance or tech or med, or were already talking about offers they had or were going to get. I think that the explicit pressure on me was never very strong, but the implicit pressure of having the example of other people have their career set up definitely made me feel behind the curve. - Isaac

Prestige was huge... It's sad but I think it's also partially, because you kind of visualize the LinkedIn profile update and how that will be received. Honestly, when I signed with [elite investing firm] at sophomore year and I told people what I was doing that summer, they were like, "You must be really smart." That starts to get to you where you realise that if you go and get this thing, then this actually affects how people perceive you. I just remember going to some NYU Stern dorm party, and then the kids being like, "Oh, my god, you're interning there?" Which is like so dumb, but you do get habituated to it. Then you're like, "Oh fuck, once I graduate [elite university], no one's going to be like, "Oh, shit, you go to [elite university]," anymore, what's next?" - Blair

I think it's more like there are internal and external pressures on students to get the high prestige job out of college. It's not so much, "I want to do these jobs because it is going to be the ultimate good for society," but rather "Getting this job is how you're going to prove that you're better than your peers. That you've worked hard in college, this is how you're going to continue to

demonstrate that you're a high-achiever" - Ryan

Appendix B

Example summary memos of interviews transcripts for Paper #1

Alexandra - Not Interested in Teaching

Alexandra moved from Spain to the US her sophomore year of high school, and one of the first things she noticed about the education system here was a shocking amount of inequity. She started tutoring and realized she loved teaching, but she didn't tell anyone because "people look down on you wanting to be a teacher just because they think you can do more with your education." There's the first barrier to pursuing teaching that Alexandra identifies when you come from a higher level institution, like an Ivy League: prestige. Alexandra describes an anecdote which resonated with her: "I remember I had a teacher, my English teacher went to Stanford and just I would get constant comments from my parents like, 'Oh my gosh, she went to Stanford and she's just a teacher.'"

When she arrived at [elite college], Alexandra's goal was to study economics and education. Her junior year, she taught at a High School summer program, which solidified a love of teaching. She said the environment and the people were both great, and she felt really comfortable in that space. Then, she ran into the second barrier to teaching from an elite institution: parental pushback. Alexandra's mother had been a teacher in Spain, where the salary is much better than in America. And since Alexandra had loved math in high school, it seemed expected that she would become an engineer, rather than a math teacher. Her parents assumed that because she was high-achieving and hard working, she could (or should) get a better job than just being a teacher. Her intent to study economics satisfied her parents' expectations for

her.

The third barrier to teaching, teacher pay, impacted Alexandra's decision making as well. Her parents wanted her to be able to live comfortably and have enough money to cover any emergencies or medical expenses, quite a valid concern for the United States. Especially coming from Spain, the differences are stark:

“The teacher salary is actually much better in Spain and healthcare is free. It's just an easier life. My mom was a teacher in Spain and so it's a life where you can have a good living and you can have a house, if something happens it'll be okay because healthcare screening, all of that. I think it's just the idea of coming to the US and realizing how different it is and how expensive it is if something bad happens to you.”

All of these factors compounded and ultimately led to her taking a job in economics research postgrad, to determine whether or not she might want a PhD. Alexandra said that she still really wanted to go into education, but knew that she didn't want to teach long term, and the research job satisfied her parents' concerns as well. For now, research is where she'll land post-graduation, but she still has a deep fondness for being in the classroom.

Edwin - Interested but not committed to teaching

Edwin proudly came into college Undecided. After taking a freshman seminar in sociology, Edwin realized that he might be interested in pursuing academia. Around senior fall, when it became time to apply to postgraduate opportunities, there were several options on the table for Edwin. He evaluated each, and ultimately decided against both a fellowship to a PhD program in the UK and [elite university]. He ultimately took a research position at a philanthropy research firm, which requires both qualitative and quantitative thinking skills to conduct and synthesize interviews. This is a job which, in Edwin's words, has “opportunity for both personal

and professional development”. There’s variety and a dynamic work environment, and he’ll be able to live at home and save with a possibility of eventually transferring to their San Francisco office.

When asked why he opted not to apply to his alma mater’s fifth year masters in teaching program, Edwin responded:

“The exercise of passing down knowledge is something that I really like specifically with somebody that is younger than me, and so I think that I by no means have excluded teaching as something that I will do. However, I think that in this process I think I realized that I was at a point where I was more interested in deploying more my sociological experiences, rather than an interest and a desire to teach”

What really seemed to tip the scale for Edwin was a lack of actual teaching experience, in comparison to his familiarity with the research process. This, combined with other uncertainties surrounding both the fifth year teaching program and the process of becoming a teacher in general (placement, finances, individual school support and experiences) worked to pull Edwin towards the sociology field rather than the education field. Instead of jumping right into education, when he didn’t feel quite ready to make a long term commitment to teaching, Edwin is pursuing research so he has time to grapple more with questions pertaining to academia as a whole like: "How do you change a field? How do you change academia? Is it possible to be an academic that's also practicing somehow?"

A direct impact is clearly quite important to Edwin, which is why sociological research is not the long term goal for him either. He wants to be able to give back, coming from such a privileged space as [Elite university], to make that space as accessible as possible. And though teachers coming from Ivy League schools sometimes struggle with the prestige factor when

accepting post grad positions, Edwin responded with “the fact that I went to [elite university] has not been a reason for me to say that I don't want to teach rather for me I think about it in the sense that I'm trying to have again, for me, it's more about the impact”.

Amiyah - Committed to teaching

Amiyah’s view of education was shaped predominantly by her K-12 experience in a very urban environment, with students who were mostly black and Latinx and the teachers who were white. The language differences between students and teachers that Amiyah observed set her down a path studying linguistics, with hopes to apply her studies to activism work. Amiyah began to realize she enjoyed working with kids, and wanted to get into a school setting, which led her to Generation Teach freshman summer:

“I've had previous experience in high school doing mentorship stuff and as I've been going on this process, I was like, "I like working with kids." I applied to the program and I was accepted. I spent my summer with it. It was called Generation Teach, if that's helpful information. I really enjoyed myself. I loved working with the kids. I love the camaraderie, I dealt with the other folks who were really interested in education and pursuing careers as teachers, as principals, as counselors, all that kind of stuff.”

Her summer experience teaching had a domino effect that led to switching her major at [Elite University] to Education Studies. Upon realizing she wouldn’t be certified to teach in public schools after graduation, Amiyah began applying to graduate school programs through the Institute for Recruitment of Teachers. She looks fondly at her twin brother’s high school speech pathology teacher, who helped her recognize some of the challenges teachers face: overworked and under-supported, to name two key difficulties. Those challenges served more as inspiration than as a deterrent for Amiyah. And although attending Brown meant that the people around her suggested careers such as doctor, lawyer, or politician frequently to Amiyah, none of those stuck

with her as much as teaching.

Amiyah realized she would be content “living comfortably” and doing something that was “social-justice oriented” because of the way activism spaces at Brown had encouraged her. Amiyah wanted to make an impact, which is a typical driving force behind those who go into education as a career. However, Amiyah’s mom was a teacher, and didn’t want Amiyah to be somewhere where she was (again) overworked, undervalued and underpaid. Ultimately, though, Amiyah made it clear that teaching would make her happiest, and of course coming from a teaching background herself, Amiyah’s mom understood that feeling.

Amiyah is going to [Elite University] for her Masters in Education thanks to the IRC program. Towards the end of her interview, Amiyah summed up what the teaching experience meant to her, and why it didn’t matter that she was straying from the “Ivy League” path: “Not as many people who are interested in entering as an educator and the people who I have met have also been low-income women of color who are like, “Oh my God, there's so much here. Let me bring it back to where I'm from.”

Appendix C

List of Included Studies for Paper #3

- Abendroth, M., Golzy, J. B., & O'Connor, E. A. (2011). Self-Created Youtube Recordings of Microteachings: Their Effects upon Candidates' Readiness for Teaching and Instructors' Assessment. *Journal of Educational Technology Systems*, 40(2), 141–159.
- Adams, J. D., & Gupta, P. (2017). Informal science institutions and learning to teach: An examination of identity, agency, and affordances. *Journal of Research in Science Teaching*, 54(1), 121–138.
- Adler, R. F., & Kim, H. (2018). Enhancing future K-8 teachers' computational thinking skills through modeling and simulations. *Education and Information Technologies*, 23(4), 1501–1514.
- Adler, R. F., & Kim, H. (2018). Enhancing future K-8 teachers' computational thinking skills through modeling and simulations. *Education and Information Technologies*, 23(4), 1501–1514.
- Akcay, H., & Yager, R. (2010). Accomplishing the Visions for Teacher Education Programs Advocated in the National Science Education Standards. *Journal of Science Teacher Education*, 21(6), 643–664.
- Allen, S. J., & Blackston, A. R. (2003). Training preservice teachers in collaborative problem solving: An investigation of the impact on teacher and student behavior change in real-world settings. *School Psychology Quarterly*, 18(1), 22.
- Ambrose, R. (2004). Initiating Change in Prospective Elementary School Teachers' Orientations to Mathematics Teaching by Building on Beliefs. *Journal of Mathematics Teacher Education*, 7(2), 91–119.
- Anhalt, C. O., & Cortez, R. (2016). Developing understanding of mathematical modeling in secondary teacher preparation. *Journal of Mathematics Teacher Education*, 19(6), 523–545.
- Artzt, A. F., Sultan, A., Curcio, F. R., & Gurl, T. (2012). A capstone mathematics course for prospective secondary mathematics teachers. *Journal of Mathematics Teacher Education*, 15(3), 251–262.
- Bahr, D., Monroe, E. E., & Shaha, S. H. (2013). Examining Preservice Teacher Belief Changes in the Context of Coordinated Mathematics Methods Coursework and Classroom Experiences: Coordinated Methods and Classroom Experiences. *School Science and Mathematics*, 113(3), 144–155.
- Baker, R. S., & Milner, J. O. (2006). Complexities of Collaboration: Intensity of Mentors' Responses to Paired and Single Student Teachers. *Action in Teacher Education*, 28(3), 61–72.
- Bangel, N. J., Moon, S. M., & Capobianco, B. M. (2010). Preservice Teachers'

Perceptions and Experiences in a Gifted Education Training Model. *Gifted Child Quarterly*, 54(3), 209–221.

- Bangert, A., & Kelting-Gibson, L. (2006). Teaching Principles of Assessment Literacy through Teacher Work Sample Methodology. *Teacher Education and Practice*, 19(3), 351–364.
- Barlow, A. T., & Reddish, J. M. (2005). Teacher Candidates' Conceptual Understandings of Mathematics Concepts. *Mid-Western Educational Researcher*, 18(4), 19–24.
- Bartell, T. G., Webel, C., Bowen, B., & Dyson, N. (2013). Prospective teacher learning: Recognizing evidence of conceptual understanding. *Journal of Mathematics Teacher Education*, 16(1), 57–79.
- Baxter, B. K., Jenkins, C. C., Southerland, S. A., & Wilson, P. (2004). Using a Multilevel Assessment Scheme in Reforming Science Methods Courses. *Journal of Science Teacher Education*, 15(3), 211–232.
- Baylor, A. L. (2002). Expanding preservice teachers' metacognitive awareness of instructional planning through pedagogical agents. *Educational Technology, Research and Development*; New York, 50(2), 5.
- Baylor, A. L., & Kitsantas, A. (2005). A Comparative Analysis and Validation of Instructivist and Constructivist Self-Reflective Tools (IPSRT and CPSRT) for Novice Instructional Planners. *Journal of Technology and Teacher Education*, 13(3), 433.
- Bell, R. L., Matkins, J. J., & Gansneder, B. M. (2011). Impacts of contextual and explicit instruction on preservice elementary teachers' understandings of the nature of science. *Journal of Research in Science Teaching*, 48(4), 414–436.
- Benken, B. M., & Brown, N. (2008). Integrating Teacher Candidates' Conceptions of Mathematics, Teaching, and Learning: A Cross-University Collaboration. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 1-15
- Bennett, S., & Hart, S. (2014). Addressing the 'Shift': Preparing Preservice Secondary Teachers for the Common Core. *Reading Horizons: A Journal of Literacy and Language Arts*, 53(4).
- Bickmore, B. R., Thompson, K. R., Grandy, D. A., & Tomlin, T. (2009). Science As Storytelling for Teaching the Nature of Science and the Science-Religion Interface. *Journal of Geoscience Education*, 57(3), 178–190.
- Bintz, W., & Shake, M. (2005). From University to Classrooms: A Preservice Teachers' Writing Portfolio Program and its Impact on Instruction in Teaching Strategies for Writing Portfolios in the Classroom. *Reading Horizons*, 45(3), 18.
- Birrell, J. R., & Bullough, R. V. (2005). Teaching with a Peer: A Follow-Up Study of the 1st Year of Teaching. *Action in Teacher Education*, 27(1), 72–81.
- Bischoff, P. J. (2006). The role of knowledge structures in the ability of preservice elementary teachers to diagnose a child's understanding of molecular kinetics. *Science Education*, 90(5), 936–951.
- Bischoff, P. J., & Golden, C. F. (2003). Exploring the Role of Individual and Socially

Constructed Knowledge Mobilization Tasks in Revealing Preservice Elementary Teachers' Understandings of a Triangle Fraction Task. *School Science and Mathematics*, 103(6), 266–273.

- Boesdorfer, S. B., & Asprey, L. M. (2017). Exploratory Study of the Teaching Practices of Novice Science Teachers Who Participated in Undergraduate Science Education Research. *Electronic Journal of Science Education*, 21(3), 21–45.
- Bravo, M. A., Mosqueda, E., Solís, J. L., & Stoddart, T. (2014). Possibilities and Limits of Integrating Science and Diversity Education in Preservice Elementary Teacher Preparation. *Journal of Science Teacher Education*, 25(5), 601–619.
- Britner, S. L., & Finson, K. D. (2005). Preservice teachers' reflections on their growth in an inquiry-oriented science pedagogy course. *Journal of Elementary Science Education*, 17(1), 39.
- Buchanan, L. B. (2015). Fostering Historical Thinking toward Civil Rights Movement Counter-Narratives: Documentary Film in Elementary Social Studies. *The Social Studies*, 106(2), 47–56.
- Buck, G. A., & Cordes, J. G. (2005). An Action Research Project on Preparing Teachers to Meet the Needs of Underserved Student Populations. *Journal of Science Teacher Education*, 16(1), 43–64.
- Bullough, R. V., Young, J., Birrell, J. R., Cecil Clark, D., Winston Egan, M., Erickson, L., ... Welling, M. (2003). Teaching with a peer: A comparison of two models of student teaching. *Teaching and Teacher Education*, 19(1), 57–73.
- Bulunuz, N., & Jarrett, O. S. (2009). Understanding of Earth and Space Science Concepts: Strategies for Concept-Building in Elementary Teacher Preparation. *School Science and Mathematics*, 109(5), 276–289.
- Cardetti, F., & Truxaw, M. P. (2014). Toward Improving the Mathematics Preparation of Elementary Preservice Teachers. *School Science and Mathematics*, 114(1), 1–9.
- Carrier, S. J. (2013). Elementary Preservice Teachers' Science Vocabulary: Knowledge and Application. *Journal of Science Teacher Education*, 24(2), 405–425.
- Cartwright, T., Smith, S., & Hallar, B. (2014). Confronting Barriers to Teaching Elementary Science: After-School Science Teaching Experiences for Preservice Teachers. *Teacher Education & Practice*, 27(2/3), 464–487.
- Casey, S. A., Albert, J., & Ross, A. (2018). Developing Knowledge for Teaching Graphing of Bivariate Categorical Data. *Journal of Statistics Education*, 26(3), 197–213.
- Castro, A. M. (2006). Preparing Elementary Preservice Teachers to Use Mathematics Curriculum Materials. *Mathematics Educator*, 16(2), 14–24.
- Caughlan, S., Juzwik, M. M., Borsheim-Black, C., Kelly, S., & Fine, J. G. (2013). English Teacher Candidates Developing Dialogically Organized Instructional Practices. *Research in the Teaching of English*, 47(3), 212–246.
- Certo, J. L., Apol, L., Wibbens, E., & Hawkins, L. K. (2012). Living the Poet's Life: Using an Aesthetic Approach to Poetry to Enhance Preservice Teachers' Poetry

Experiences and Dispositions. *English Education*, 44(2), 102–146.

- Chamberlin, M. T., & Candelaria, M. S. (2018). Learning from Teaching Teachers: A Lesson Experiment in Area and Volume with Prospective Teachers. *Mathematics Teacher Education and Development*, 20(1), 86–111.
- Chamberlin, M. T., & Powers, R. A. (2007). Selecting from Three Curricula for a Preservice Elementary Teacher Geometry Course. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 4.
- Chizhik, E. W., Chizhik, A. W., Close, C., & Gallego, M. (2017). SMILE (Shared Mentoring in Instructional Learning Environments): Effectiveness of a Lesson-Study Approach to Student-Teaching Supervision on a Teacher-Education Performance Assessment. *Teacher Education Quarterly*, 44(2), 27–47.
- Clark, K. M. (2012). History of mathematics: Illuminating understanding of school mathematics concepts for prospective mathematics teachers. *Educational Studies in Mathematics*, 81(1), 67–84.
- Cohen, M. D., & Nath, J. L. (2006). Paired Placements for Early Field Experiences. *Teacher Education and Practice*, 19(1), 24–40.
- Colwell, J. (2016). Examining Preservice Teachers' Beliefs about Disciplinary Literacy in History through a Blog Project. *Action in Teacher Education*, 38(1), 34–48.
- Conner, A., Edenfield, K. W., Gleason, B. W., & Ersoz, F. A. (2011). Impact of a content and methods course sequence on prospective secondary mathematics teachers' beliefs. *Journal of Mathematics Teacher Education*, 14(6), 483–504.
- Courtney, A., & King, F. B. (2009). Online Dialogue Discussion: A Tool to Support Pre-Service Teacher Candidates' Understanding of Literacy Teaching and Practice2. *Contemporary Issues in Technology and Teacher Education*, 9(3), 226–256.
- Crespo, S., & Nicol, C. (2006). Challenging Preservice Teachers' Mathematical Understanding: The Case of Division by Zero. *School Science and Mathematics*, 106(2), 84–97.
- Crespo, S., & Sinclair, N. (2008). What makes a problem mathematically interesting? Inviting prospective teachers to pose better problems. *Journal of Mathematics Teacher Education*, 11(5), 395–415.
- Curcio, R., & Adams, A. (2019). The Development of Mentoring Partnerships: How a Shared Learning Experience Enhanced the Final Internship. *SRATE Journal*, 28(1), 8.
- Cushmann, C. A., & Kemp, A. T. (2012). The Effects of Clinical Experiences on the Understanding of Classroom Management Techniques. *Journal of Inquiry & Action in Education*, 4(3), 44–58.
- Dass, P. M. (2005). Using a Science/Technology/Society Approach To Prepare Reform-Oriented Science Teachers: 14(1), 14.
- Dee, A. (2012). Collaborative Clinical Practice: An Alternate Field Experience. *Issues in Teacher Education*, 21(2), 147–163.
- DelleBovi, B. M. (2012). Literacy instruction: From assignment to assessment. *Assessing*

Writing, 17(4), 271–292.

- DeMink-Carthew, J. (2017). Reform-Oriented Collaborative Inquiry as a Pedagogy for Student Teaching in Middle School. *Middle Grades Research Journal*, 11(1), 13–27.
- Dempsey, M. S., PytlikZillig, L. M., & Bruning, R. H. (2009). Helping preservice teachers learn to assess writing: Practice and feedback in a Web-based environment. *Assessing Writing*, 14(1), 38–61.
- Deniz, H. (2011). Examination of Changes in Prospective Elementary Teachers' Epistemological Beliefs in Science and Exploration of Factors Meditating That Change. *Journal of Science Education and Technology*, 20(6), 750–760.
- Duncan, R. G., Pilitsis, V., & Piegaro, M. (2010). Development of Preservice Teachers' Ability to Critique and Adapt Inquiry-based Instructional Materials. *Journal of Science Teacher Education*, 21(1), 81–102.
- Duran, L. B., McArthur, J., & Hook, S. V. (2004). Undergraduate Students' Perceptions of an Inquiry-Based Physics Course. *Journal of Science Teacher Education*, 15(2), 155–171.
- Edwards, C. J., Carr, S., & Siegel, W. (2006). Influences of Experiences and Training on Effective Teaching Practices to Meet the Needs of Diverse Learners in Schools. *Education*, 126(3), 580–592.
- Ely, E., Alves, K. D., Dolenc, N. R., Sebolt, S., & Walton, E. A. (2018). Classroom Simulation to Prepare Teachers to Use Evidence-Based Comprehension Practices. *Journal of Digital Learning in Teacher Education*, 34(2), 71–87.
- Fives, H., & Barnes, N. (2017). Informed and Uninformed Naïve Assessment Constructors' Strategies for Item Selection. *Journal of Teacher Education*, 68(1), 85–101.
- Foley, R. W., Archambault, L. M., Hale, A. E., & Dong, H.-K. (2017). Learning Outcomes in Sustainability Education Among Future Elementary School Teachers. *Journal of Education for Sustainable Development*, 11(1), 33–51.
- Giebelhaus, C. R., & Bowman, C. L. (2002). Teaching Mentors: Is It Worth the Effort? *The Journal of Educational Research*, 95(4), 246–254.
- Girod, M., & Girod, G. (2006). Exploring the Efficacy of the Cook School District Simulation. *Journal of Teacher Education*, 57(5), 481–497.
- Goodson-Espy, T., Cifarelli, V. V., Pugalee, D., Lynch-Davis, K., Morge, S., & Salinas, T. (2014). Applying NAEP to Improve Mathematics Content and Methods Courses for Preservice Elementary and Middle School Teachers. *School Science and Mathematics*, 114(8), 392–404.
- Groth, R. E., Bergner, J. A., & Burgess, C. R. (2016). An Exploration of Prospective Teachers' Learning of Clinical Interview Techniques. *Mathematics Teacher Education and Development*, 18(2), 48–71.
- Groves, F. H., & Pugh, A. F. (2002). Cognitive Illusions as Hindrances to Learning Complex Environmental Issues. *Journal of Science Education and Technology*, 11(4), 381–390.

- Hanuscin, D. L., & Zangori, L. (2016). Developing Practical Knowledge of the Next Generation Science Standards in Elementary Science Teacher Education. *Journal of Science Teacher Education*, 27(8), 799–818.
- Hart, L. C. (2002). Preservice Teachers' Beliefs and Practice After Participating in an Integrated Content/Methods Course. *School Science and Mathematics*, 102(1), 4–14.
- Hart, S. M. & Bennett, S. M. (2013). Disciplinary Literacy Pedagogy Development of STEM Preservice Teachers. *Teacher Education & Practice*, 26(2), 221–241.
- Hawkins, S., & Park Rogers, M. (2016). Tools for Reflection: Video-Based Reflection Within a Preservice Community of Practice. *Journal of Science Teacher Education*, 27(4), 415–437.
- Hestness, E., Randy McGinnis, J., Riedinger, K., & Marbach-Ad, G. (2011). A Study of Teacher Candidates' Experiences Investigating Global Climate Change Within an Elementary Science Methods Course. *Journal of Science Teacher Education*, 22(4), 351–369.
- Hiebert, J., Berk, D., Miller, E., Gallivan, H., & Meikle, E. (2019). Relationships Between Opportunity to Learn Mathematics in Teacher Preparation and Graduates' Knowledge for Teaching Mathematics. *Journal for Research in Mathematics Education*, 50(1), 23.
- Hoppey, D., & Mickelson, A. M. (2017). Partnership and Coteaching: Preparing Preservice Teachers to Improve Outcomes for Students with Disabilities. *Action in Teacher Education*, 39(2), 187–202.
- Isabelle, A. D., & de Groot, C. (2008). Alternate Conceptions of Preservice Elementary Teachers: The Itakura Method. *Journal of Science Teacher Education*, 19(5), 417–435.
- Iwai, Y. (2019). Culturally Responsive Teaching in a Global Era: Using the Genres of Multicultural Literature. *The Educational Forum*, 83(1), 13–27.
- Jackson, C., Mohr-Schroeder, M., & Little, D. (2014). Using Informal Learning Environments to Prepare Preservice Teachers. *Teacher Education & Practice*, 27(2/3), 445–463.
- Jacobbe, T., Ross, D. D., Caron, D. A., Barko, T., & Busi, R. (2014). Connecting Theory and Practice: Preservice Teachers' Construction of Practical Tools for Teaching Mathematics. *Teacher Education and Practice*, 27, 332–355.
- James, M. C., & Scharmann, L. C. (2007). Using analogies to improve the teaching performance of preservice teachers. *Journal of Research in Science Teaching*, 44(4), 565–585.
- Jenkins, O. F. (2010). Developing teachers' knowledge of students as learners of mathematics through structured interviews. *Journal of Mathematics Teacher Education*, 13(2), 141–154.
- Jett, C. C. (2014). Using Mathematics Literature with Prospective Secondary Mathematics Teachers. *Journal of Mathematics Education at Teachers College*, 5(2), 49–53.

- Jo, I. (2016). Future Teachers' Dispositions Toward Teaching With Geospatial Technologies. *Contemporary Issues in Technology and Teacher Education*, 16(3), 310–327.
- Jung, M. L., & Tonso, K. L. (2006). Elementary preservice teachers learning to teach science in science museums and nature centers: A novel program's impact on science knowledge, science pedagogy, and confidence teaching. *Journal of Elementary Science Education*, 18(1), 15–31.
- Kattoula, E., Verma, G., & Martin-Hansen, L. (2009). Fostering Preservice Teachers' "Nature of Science" Understandings in a Physics Course. *Journal of College Science Teaching*, 39(1), 18–26.
- Kaya, E., Newley, A., Deniz, H., Yesilyurt, E., & Newley, P. (2017). Introducing Engineering Design to a Science Teaching Methods Course Through Educational Robotics and Exploring Changes in Views of Preservice Elementary Teachers. *Journal of College Science Teaching*, 47(2), 66–75.
- Kenney, R., Shoffner, M., & Norris, D. (2014). Reflecting on the Use of Writing to Promote Mathematical Learning: An Examination of Preservice Mathematics Teachers' Perspectives. *The Teacher Educator*, 49(1), 28–43.
- Kinach, B. M. (2002). A cognitive strategy for developing pedagogical content knowledge in the secondary mathematics methods course: Toward a model of effective practice. *Teaching and Teacher Education*, 18(1), 51–71.
- Kopcha, T. J., & Alger, C. (2011). The Impact of Technology- Enhanced Student Teacher Supervision on Student Teacher Knowledge, Performance, and Self-Efficacy during the Field Experience. *Journal of Educational Computing Research*, 45(1), 49–73.
- Kucan, L., Palincsar, A. S., Busse, T., Heisey, N., Klingelhofer, R., Rimbey, M., & Schutz, K. (2011). Applying the Grossman et al. Theoretical Framework: The Case of Reading. *Teachers College Record*, 113(12), 2897–2921.
- Laframboise, K. L., & Shea, K. (2009). Developing Understanding of Research-based Pedagogy with Preservice Teachers: An Instrumental Case Study. *The Qualitative Report*, 14(1), 105–128.
- Lambert, J. L., & Bleicher, R. E. (2013). Climate Change in the Preservice Teacher's Mind. *Journal of Science Teacher Education*, 24(6), 999–1022.
- Lee, C. K.-P. (2012). An evaluation of an elementary science methods course with respect to preservice teacher's pedagogical. 13(2), 19.
- Lee, H.-J., Özgün-Koca, S. A., Meagher, M., & Edwards, M. T. (2018). Examining the Impact of a Framework to Support Prospective Secondary Teachers' Transition from "Doer" to "Teacher" of Mathematics. *Mathematics Teacher Education and Development*, 20(1), 112–131.
- Lewis, E., Dema, O., & Harshbarger, D. (2014). Preparation for Practice: Elementary Preservice Teachers Learning and Using Scientific Classroom Discourse Community Instructional Strategies. *School Science and Mathematics*, 114(4), 154–165.

- Lin, C.-Y. (2009). A comparison study of web-based and traditional instruction on preservice teachers' knowledge of fractions. *Contemporary Issues in Technology and Teacher Education*, 9(3), 257–279.
- Loverude, M. E., Gonzalez, B. L., & Nanes, R. (2011). Inquiry-based course in physics and chemistry for preservice K-8 teachers. *Physical Review Special Topics - Physics Education Research*, 7(1), 010106.
- Lubinski, C. A., & Otto, A. D. (2004). Preparing K-8 Preservice Teachers in a Content Course for Standards-Based Mathematics Pedagogy. *School Science and Mathematics*, 104(7), 336–350.
- Lux, N. J. (2013). Technology-Focused Early Field Experiences in Preservice Teacher Education. *Journal of Digital Learning in Teacher Education*, 29(3), 82–88.
- Mahalingappa, L., Hughes, E. M., & Polat, N. (2018). Developing preservice teachers' self-efficacy and knowledge through online experiences with English language learners. *Language and Education*, 32(2), 127–146.
- Matkins, J. J., & Bell, R. L. (2007). Awakening the Scientist Inside: Global Climate Change and the Nature of Science in an Elementary Science Methods Course. *Journal of Science Teacher Education*, 18(2), 137–163.
- Matthews, M. E., & Seaman, W. I. (2007). The Effects of Different Undergraduate Mathematics Courses on the Content Knowledge and Attitude towards Mathematics of Preservice Elementary Teachers. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 1.
- McCall, M. (2017). Elementary Preservice Science Teaching Efficacy and Attitude toward Science: Can a College Science Course Make a Difference? *Electronic Journal of Science Education*, 21(6), 1–11.
- McCulloch, A., Lovett, J., & Edgington, C. (2019). Designing to Provoke Disorienting Dilemmas: Transforming Preservice Teachers' Understanding of Function Using a Vending Machine Applet. *Contemporary Issues in Technology and Teacher Education*, 19(1), 4–22.
- McGee, J., & Colby, S. (2014). Impact of an Assessment Course on Teacher Candidates' Assessment Literacy. *Action in Teacher Education*, 36(5–6), 522–532.
- Merk, H., Waggoner, J., & Carroll, J. (2013). Co-Learning: Maximizing Learning in Clinical Experiences. *AILACTE Journal*, 10(1), 79–95.
- Mitchell, K., Homza, A., & Ngo, S. (2012). Reading Aloud with Bilingual Learners: A Fieldwork Project and Its Impact on Mainstream Teacher Candidates. *Action in Teacher Education*, 34(3), 276–294.
- Mizell, J. A., & Cates, J. (2004). The Impact of Additional Content Courses on Teacher Candidates' Beliefs regarding Mathematics Content and Pedagogy. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 4.
- Moreno, R., & Abercrombie, S. (2010). Promoting Awareness of Learner Diversity in Prospective Teachers: Signaling Individual and Group Differences within Virtual

- Classroom Cases. *Journal of Technology and Teacher Education*, 18(1), 111–130.
- Morris, A. K., & Hiebert, J. (2017). Effects of Teacher Preparation Courses: Do Graduates Use What They Learned to Plan Mathematics Lessons? *American Educational Research Journal*, 54(3), 524–567.
 - Morrison, J., & McDuffie, A. R. (2009). Connecting Science and Mathematics: Using Inquiry Investigations to Learn About Data Collection, Analysis, and Display. *School Science and Mathematics*, 109(1), 31–44.
 - Morrone, A. S., Harkness, S. S., D'Ambrosio, B., & Caulfield, R. (2004). Patterns of Instructional Discourse that Promote the Perception of Mastery Goals in a Social Constructivist Mathematics Course. *Educational Studies in Mathematics*, 56(1), 19–38.
 - Mostofo, J., & Zambo, R. (2015). Improving instruction in the mathematics methods classroom through action research. *Educational Action Research*, 23(4), 497–513.
 - Nasir, A., & Heineke, A. J. (2014). Teacher Candidates and Latina/o English Learners at Fenton Elementary School: The Role of Early Clinical Experiences in Urban Teacher Education. *Association of Mexican American Educators Journal*, 8(1).
 - Nesmith, S. M., Purdum-Cassidy, B., Cooper, S., & Rogers, R. D. (2017). Love It, Like It, or Leave It — Elementary Preservice Teachers' Field-Based Perspectives toward the Integration of Literature in Mathematics. *Action in Teacher Education*, 39(3), 321–339.
 - Nokes, J. D. (2010). Preparing Novice History Teachers to Meet Students' Literacy Needs. *Reading Psychology*, 31(6), 493–523.
 - Novak, E., & Wisdom, S. (2018). Effects of 3D Printing Project-based Learning on Preservice Elementary Teachers' Science Attitudes, Science Content Knowledge, and Anxiety About Teaching Science. *Journal of Science Education and Technology*, 27(5), 412–432.
 - Olson, J. K., Bruxvoort, C. N., & Vande Haar, A. J. (2016). The impact of video case content on preservice elementary teachers' decision-making and conceptions of effective science teaching. *Journal of Research in Science Teaching*, 53(10), 1500–1523.
 - Otto, C. A., Luera, G. R., & Everett, S. A. (2009). An Innovative Course Featuring Action Research Integrated with Unifying Science Themes. *Journal of Science Teacher Education*, 20(6), 537.
 - Ozogul, G., & Sullivan, H. (2009). Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Educational Technology Research and Development*, 57(3), 393–410.
 - Petrosino Jr., A., & Mann, M. (2018). Data Modeling for Preservice Teachers and Everyone Else. *Journal of College Science Teaching*, 47(3).
 - Phillippo, K., & Blosser, A. (2017). Stable roles, changed skills: Teacher candidate responses to instruction about adolescent psychosocial support practices. *Advances in School Mental Health Promotion*, 10(1), 5–25.
 - Piro, J. S., & Hutchinson, C. J. (2014). Using a Data Chat to Teach Instructional Interventions: Student Perceptions of Data Literacy in an Assessment Course. *The New*

Educator, 10(2), 95–111.

- Plummer, J. D., Zahm, V. M., & Rice, R. (2010). Inquiry and Astronomy: Preservice Teachers' Investigations of Celestial Motion. *Journal of Science Teacher Education*, 21(4), 471–493.
- Pratt, S. S. (2018). Area models to image integer and binomial multiplication. *Investigations in Mathematics Learning*, 10(2), 85–105.
- Ragland, R. (2016). Implementing an Evidence-Based Reflective Teaching Cycle: Using Scholarly Research in Curriculum Design. *Mid-Western Educational Researcher*, 28(3), 196–217.
- Rhine, S., Harrington, R., & Olszewski, B. (2015). The Role of Technology in Increasing Preservice Teachers' Anticipation of Students' Thinking in Algebra. *Contemporary Issues in Technology and Teacher Education*, 15(2), 85–105.
- Rigelman, N. M., & Ruben, B. (2012). Creating foundations for collaboration in schools: Utilizing professional learning communities to support teacher candidate learning and visions of teaching. *Teaching and Teacher Education*, 28(7), 979–989.
- Rinke, C. R., Gladstone-Brown, W., Kinlaw, C. R., & Cappiello, J. (2016). Characterizing STEM Teacher Education: Affordances and Constraints of Explicit STEM Preparation for Elementary Teachers: Characterizing STEM Teacher Education. *School Science and Mathematics*, 116(6), 300–309.
- Ro, J. M., Magiera, K., Gradel, K., & Simmons, R. (2013). Blog-Based Support for Preservice Teachers in an Afterschool Tutoring Program. *Journal of Educational Technology Systems*, 42(1), 69–84.
- Rodriguez-Valls, F., & Ponce, G. (2013). Classroom, the We Space: Developing Student-Centered Practices for Second Language Learner (SLL) Students. *Education Policy Analysis Archives*, 21, 55.
- Rosaen, C. L., Lundeberg, M., Terpstra, M., Cooper, M., Niu, R., & Fu, J. (2010). Constructing videocases to help novices learn to facilitate discussions in science and English: How does subject matter matter? *Teachers and Teaching*, 16(4), 507–524.
- Sabel, J. L., Forbes, C. T., & Zangori, L. (2015). Promoting Prospective Elementary Teachers' Learning to Use Formative Assessment for Life Science Instruction. *Journal of Science Teacher Education*, 26(4), 419–445.
- Salajan, F. D., Nyachwaya, J. M., Hoffman, J. G., & Hill, B. D. (2016). Improving Teacher Candidates' Lesson Planning Competencies Through Peer Review in a Wiki Environment. *The Teacher Educator*, 51(3), 185–210.
- Salter, I., & Atkins, L. (2013). Student-Generated Scientific Inquiry for Elementary Education Undergraduates: Course Development, Outcomes and Implications. *Journal of Science Teacher Education*, 24(1), 157–177.
- Santagata, R., & Yeh, C. (2014). Learning to teach mathematics and to analyze teaching effectiveness: Evidence from a video- and practice-based approach. *Journal of Mathematics Teacher Education*, 17(6), 491–514.

- Santau, A. O., Maerten-Rivera, J. L., Bovis, S., & Orend, J. (2014). A Mile Wide or an Inch Deep? Improving Elementary Preservice Teachers' Science Content Knowledge Within the Context of a Science Methods Course. *Journal of Science Teacher Education*, 25(8), 953–976.
- Santoyo, C., & Zhang, S. (2016). Secondary Teacher Candidates' Lesson Planning Learning. *Teacher Education Quarterly*, 43(2), 3–27.
- Sayeski, K. L., Earle, G. A., Eslinger, R. P., & Whitenton, J. N. (2017). Teacher candidates' mastery of phoneme-grapheme correspondence: Massed versus distributed practice in teacher education. *Annals of Dyslexia*, 67(1), 26–41.
- Sayeski, K. L., Kennedy, M. J., Irala, S. de, Clinton, E., Hamel, M., & Thomas, K. (2015). The Efficacy of Multimedia Modules for Teaching Basic Literacy-Related Concepts. *Exceptionality*, 23(4), 237–257.
- Schussler, D., Frank, J., Lee, T.-K., & Mahfouz, J. (2017). Using Virtual Role-Play to Enhance Teacher Candidates' Skills in Responding to Bullying. *Journal of Technology and Teacher Education*, 25(1), 30.
- Scott, C. (2016). Using Citizen Science to Engage Preservice Elementary Educators in Scientific Fieldwork. *Journal of College Science Teaching*, 46(2), 37–41.
- Seung, E., Bryan, L. A., & Butler, M. B. (2009). Improving Preservice Middle Grades Science Teachers' Understanding of the Nature of Science Using Three Instructional Approaches. *Journal of Science Teacher Education*, 20(2), 157–177.
- Shin, E.-K., Wilkins, E. A., & Ainsworth, J. (2007). The Nature and Effectiveness of Peer Feedback during an Early Clinical Experience in an Elementary Education Program. *Action in Teacher Education*, 28(4), 40–52.
- Smith, M. E., Swars, S. L., Smith, S. Z., Hart, L. C., & Haardörfer, R. (2012). Effects of an Additional Mathematics Content Course on Elementary Teachers' Mathematical Beliefs and Knowledge for Teaching. *Action in Teacher Education*, 34(4), 336–348.
- Soh, L.-K., Fowler, D., & Zygielbaum, A. I. (2007). The Impact of the Affinity Learning Authoring Tool on Student Learning. *Journal of Educational Technology Systems*, 36(1), 29–62.
- Steele, M. D., Hillen, A. F., & Smith, M. S. (2013). Developing mathematical knowledge for teaching in a methods course: The case of function. *Journal of Mathematics Teacher Education*, 16(6), 451–482.
- Stover, K., Yearta, L. S., & Sease, R. (2014). “Experience Is the Best Tool for Teachers”: Blogging to Provide Preservice Educators with Authentic Teaching Opportunities. *Journal of Language and Literacy Education*, 10(2), 99–117.
- Sunal, C. S., & Sunal, D. W. (2003). Teacher Candidates' Conceptualization of Guided Inquiry and Lesson Planning in Social Studies Following Web-Assisted Instruction. *Theory & Research in Social Education*, 31(2), 243–264.
- Suppa, S., DiNapoli, J., & Mixell, R. (2018). Teacher Preparation “Does” Matter: Relationships between Elementary Mathematics Content Courses and Graduates’

- Analyses of Teaching. *Mathematics Teacher Education and Development*, 20(2), 25–57.
- Tessier, J. (2010). An Inquiry-Based Biology Laboratory Improves Preservice Elementary Teachers' Attitudes About Science. *Journal of College Science Teaching*; Washington, 39(6), 84–90.
 - Thanheiser, E. (2015). Developing prospective teachers' conceptions with well-designed tasks: Explaining successes and analyzing conceptual difficulties. *Journal of Mathematics Teacher Education*, 18(2), 141–172.
 - Thomas, A. F., & Sondergeld, T. (2015). Investigating the Impact of Feedback Instruction: Partnering Preservice Teachers with Middle School Students to Provide Digital, Scaffolded Feedback. *Journal of the Scholarship of Teaching and Learning*, 15(4), 83–109.
 - Trundle, K. C., Atwood, R. K., & Christopher, J. E. (2006). Preservice Elementary Teachers' Knowledge of Observable Moon Phases and Pattern of Change in Phases. *Journal of Science Teacher Education*, 17(2), 87–101.
 - Uribe, S. N., & Vaughan, M. (2017). Facilitating student learning in distance education: A case study on the development and implementation of a multifaceted feedback system. *Distance Education*, 38(3), 288–301.
 - Van Eck, R. N., Guy, M., Young, T., Winger, A. T., & Brewster, S. (2015). Project NEO: A Video Game to Promote STEM Competency for Preservice Elementary Teachers. *Technology, Knowledge and Learning*, 20(3), 277–297.
 - Weinburgh, M. (2007). The Effect of *Tenebrio obscurus* on Elementary Preservice Teachers' Content Knowledge, Attitudes, and Self-efficacy. *Journal of Science Teacher Education*, 18(6), 801–815.
 - Welsh, K. A., & Schaffer, C. (2017). Developing the Effective Teaching Skills of Teacher Candidates During Early Field Experiences. *The Educational Forum*, 81(3), 301–321.
 - Westrick, J. M., & Morris, G. A. (2016). Teacher education pedagogy: Disrupting the apprenticeship of observation. *Teaching Education*, 27(2), 156–172.
 - Wilkins, J. L. M., & Brand, B. R. (2004). Change in Preservice Teachers' Beliefs: An Evaluation of a Mathematics Methods Course. *School Science and Mathematics*, 104(5), 226–232.
 - Windschitl, M., & Thompson, J. (2006). Transcending Simple Forms of School Science Investigation: The Impact of Preservice Instruction on Teachers' Understandings of Model-Based Inquiry. *American Educational Research Journal*, 43(4), 783–835.
 - Windschitl, M., Thompson, J., & Braaten, M. (2008). How Novice Science Teachers Appropriate Epistemic Discourses Around Model-Based Inquiry for Use in Classrooms. *Cognition and Instruction*, 26(3), 310–378.
 - Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *Trans. Comput. Educ.*, 14(1), 5:1–5:16.

- Yeh, C., & Santagata, R. (2015). Preservice Teachers' Learning to Generate Evidence-Based Hypotheses About the Impact of Mathematics Teaching on Learning. *Journal of Teacher Education*, 66(1), 21–34.
- Zhang, S., & Stephens, V. (2013). Learning to Teach English-Language Learners in Mainstreamed Secondary Classrooms. *Teacher Education and Practice*, 26(1), 99–116.

Appendix D

Full Coding Scheme for Paper #3

Study characteristics

Level 2 code	Level 1 code	Response options
Study sample	Number of participants	Numerical
	Did not report participants	1=Yes, 0=No
	Number of comparison groups (<i>1 if single-case</i>)	Numerical
	Number of study sites (<i>naturally occurring sites where the treatment is delivered, such as courses, TEPs and/or higher education institution (i.e. number of different places where treatment administration might be reasonably said to differ in some way because of context)</i>)	Numerical
Certification Level	Elementary	1=Yes, 0=No
	Secondary (<i>incl. middle years</i>)	1=Yes, 0=No
	Level not reported	1=Yes, 0=No
Program Type	Traditional (<i>incl. MAT programs</i>)	1=Yes, 0=No
	Alternative (<i>based in schools; whole program concurrently teacher of record and taking coursework</i>)	1=Yes, 0=No
	Type not reported	1=Yes, 0=No
Program Level	Undergraduate	1=Yes, 0=No
	Postgraduate	1=Yes, 0=No
	Level not reported	1=Yes, 0=No
Type of teacher preparation practice (select	A course or set of courses organized around a particular theme within a TEP	1=Yes, 0=No
	An activity, pedagogy, or process implemented within a course in a TEP (<i>i.e. less than 100% of the course is being analysed</i>)	1=Yes, 0=No

one)	An experience organized by teacher educators (outside a course) for PSTs for the purpose of development	1=Yes, 0=No
	A way of organizing a TEP component (<i>e.g. assigning certain types of mentors or peers in practicum</i>)	1=Yes, 0=No
	Other	1=Yes, 0=No
Outcome measure type	All treatment/comparison groups received at least some training related to the primary outcome variable (<i>treatment-dependent -- i.e., using a new curriculum or doing a very specific kind of instructional move that the comparison group did not learn/do</i>)	1=Yes, 0=No, 98=Cannot be determined (because no reporting of what the comparison group did), 99=N/A (there is no comparison group)
	Primary outcome measures are something that the treated group were doing as part of the treatment (<i>e.g., if the treatment is writing fieldnotes, and the outcome variable is fieldnotes</i>)	1=Yes, 0=No, 98=Cannot be determined
	Primary outcome measures are developed by researchers external to the project (<i>i.e. uses a measure that has not been developed by the researchers for the purpose of the evaluation; references a different researcher who has developed the outcome measure.</i>)	1=Yes, 0=No
Type of data collected	Evidence of participation in the intervention (<i>e.g. frequency of participation, the type of participation, implementation analysis</i>)	1=Yes, 0=No
	PSTs perspectives of how good the intervention was/how much they think they learned from it	1=Yes, 0=No
	Observed PSTs skill/knowledge/belief learning from the treatment (<i>e.g. survey of teacher beliefs, aptitude on assignments/tasks</i>)	1=Yes, 0=No
	Self-reported PST instructional practice (<i>e.g. actual frequency/likelihood of implementing a particular practice</i>)	1=Yes, 0=No
	Classroom observations of PST instructional practice	1=Yes, 0=No
	Teachers' self report of student learning AND/OR students' self-report of own learning	1=Yes, 0=No
	K-12 students' perceptions of PSTs	1=Yes, 0=No
	Observed K-12 student learning (e.g. test scores or samples of work)	1=Yes, 0=No
	Other	1=Yes, 0=No
Types of outcome measures	Survey of PSTs (Knowledge/Skills/Beliefs)	1=Yes, 0=No
	PSTs' Ongoing Reflections/Journals (on their learning from the treatment)	1=Yes, 0=No

	PSTs' course/classroom participation (e.g. audio/video recording of teacher education classroom, PSTs participation on online discussion forums, teacher educator notes on PSTs participation)	1=Yes, 0=No
	Interviews with PSTs	1=Yes, 0=No
	Lesson Observation of PSTs teaching (whether in a simulated environment or a K-12 classroom)	1=Yes, 0=No
	PSTs' Course Assessment/Test/Assignment	1=Yes, 0=No
	PSTs' external performance task (e.g. edTPA for licensure)	1=Yes, 0=No
Quantitative analysis	Uses statistical analyses	1=Yes, 0=No
	Presents power analysis	1=Yes, 0=No

Research design

Level 2 code	Level 1 code	Response options
Comparison	Between groups (e.g. one group who gets and one who doesn't get the treatment)	1=Yes, 0=No
	Within person (intended vs unrelated outcome)	1=Yes, 0=No
	Within person (pretest vs posttest)	1=Yes, 0=No
	Simulated/Estimated	1=Yes, 0=No
	No Comparison	1=Yes, 0=No
Assignment	Random	1=Yes, 0=No
	Score-Based	1=Yes, 0=No
	Time (cohort to cohort)	1=Yes, 0=No
	Matching (on demographics/covariates using statistical methods)	1=Yes, 0=No
	Non-random/Self-Selection/Convenience Sample	1=Yes, 0=No
Pretest	Administered multiple times (over a period of time)	1=Yes, 0=No

	Single administration	1=Yes, 0=No
	None	1=Yes, 0=No
Pretest type	One treatment group does not receive pretest	1=Yes, 0=No
	What participants are asked (to do) in pretest is different from posttest (i.e. more or different outcomes in the posttest)	1=Yes, 0=No
	What participants are asked (to do) in pretest is the same as in the posttest	1=Yes, 0=No
	Retrospective pretest (e.g. "what was your level before you took the test? what is your level now?")	1=Yes, 0=No
Posttest (<i>measures after the treatment</i>)	Administered multiple times (over a period of time) to test retention	1=Yes, 0=No
	Single	1=Yes, 0=No
	None	1=Yes, 0=No
Type of treatment	Switching/Reversed (<i>counterbalanced design</i>)	1=Yes, 0=No
	Repeated/Removed (<i>ie. there is a period where the treatment is not administered and an outcome is measured (removed); subsequently, there is a period where the treatment is re-administered (repeated)</i>)	1=Yes, 0=No
	Single (<i>treatment happened once</i>)	1=Yes, 0=No

Issues (Threats to validity)

Level 2 code	Level 1 code	Response options
Internal validity	During the study, the experiences of the treatment and comparison groups only differed in that the treatment group received the treatment. (<i>History effects</i>)	1=Yes, 0=No, 99=N/A (no comparison group)
	Whether treatment and comparison groups could interact and influence each others' behaviours during the intervention/before the posttest is discussed/reported.	1=Yes, 0=No, 99=N/A (No comparison group)
	Whether participants moved from treatment to comparison group (or vice versa, or not) is discussed/reported.	1=Yes, 0=No, 99=N/A (No comparison group)

	Attrition from the experiment is discussed (AND/OR tested for bias).	1=Yes, 0=No (Code 1 if even slightly mentioned)
Measurement	The primary outcome was measured using a rubric/rating scheme that was the same across time and groups.	1=Yes, 0=No
	Multiple raters are used (to ensure reliability of the outcome measure).	1=Yes, 0=No, 99=N/A (Outcome measure isn't subjectively rated)
	The influence of stakes being attached to performance on the outcome (e.g. course grades, licensure, etc.) is discussed/reported.	1=Yes, 0=No
	The influence of participants knowing the theory of change of the treatment (i.e. having expectations of its effect) is discussed/reported.	1=Yes, 0=No
	Treatment is blinded (participants do not know if they are part of the treatment or control condition; code 1 if reasonably sure from how they describe the intervention)	1=Yes, 0=No or not reported, 99=N/A (No control condition)
External validity	How study participants are similar/different from the broader TEP population is described. (If all PSTs in a TEP are in the study or if mandatory to participate, then 1 -- does not count if they just say 'PSTs who volunteered were included')	1=Yes, 0=No
	Details of the site of the study (usually the TEP) are reported such that a "typical" site can be identified. (NOTE: 1 if the TEP is named; if not named, all the following must be reported: Location of program (state and urbanicity); Size of program (approximate number of students), Type of program (alternative/traditional), Level of program (undergrad/postgrad)).	1=Yes, 0=No
	The effects of the instructor as different to the effects of the treatment are discussed/reported/dealt with (<i>EITHER researchers discuss the issue of the treatment being delivered by an instructor who may be highly enthusiastic, skilled, or suited to the treatment; OR the treatment is delivered by a number of trained facilitators</i>).	1=Yes, 0=No, 99=N/A (The treatment isn't delivered by an instructor)
	The research was undertaken in the authors' own teacher education course.	1=Yes, 0=No, 99=Unsure/Not reported
	The analysis is clustered by site.	1=Yes, 0=No, 99=N/A (Only 1 site)