The Dissertation Committee for Lauren Helena Boyle Certifies that this is the approved version of the following Dissertation:

An Exploration of First-Year Teacher Participation in Mentoring Programs and Risk for Occupational Stress

Committee:

Christopher McCarthy, Supervisor

Paul Fitchett, Co-Supervisor

Michael Parent

Tiffany Whittaker

An Exploration of First-Year Teacher Participation in Mentoring Programs and Risk for Occupational Stress

by

Lauren Helena Boyle

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2021

Dedication

To Mom and Dad. You were there, steady and full of love, throughout my entire journey. Thank you.

To my siblings, Josh, Rebecca, and Monica. I love you all. (Josh and Becca, pro tip: if you just read the abstract, I'll never know if you haven't actually read the entire document).

To Ronnie. Thank you for encouraging me when I needed it the most, for cooking countless amazing meals while I was camped out in the living room working, and for being so proud of me. I love you.

Acknowledgements

I am incredibly grateful to my advisor, Dr. Chris McCarthy, and my dissertation committee members for their time, energy, feedback, and support. A special thank you to Paul Fitchett and Rich Lambert, whose statistical expertise and humor have helped keep me afloat time and time again. Thank you also to my colleagues in the Coping and Stress in Education Research Team. Kristen Mosley, I miss sharing the data closet with you! Special thanks are owed to Maytal Eyal and Travis Bauer, who fill my heart with joy and joined me in countless grad school Book Chex. Emma Hamilton, thank you for your tables expertise.

Abstract

An Exploration of First-Year Teacher Participation in Mentoring Programs and Risk for Occupational Stress

Lauren Helena Boyle, Ph.D.

The University of Texas at Austin, 2021

Supervisor: Christopher McCarthy, Paul Fitchett

Given the significant time and capital invested annually into the development and execution of formal mentoring programs, there is a vested interest among policymakers and administrators to better understand what types of mentoring supports teachers are receiving, and what effects mentoring programs have on teacher stress risk. Although mentoring is increasingly hailed as one of the most critical components of teacher induction programs, the corresponding research base has failed to provide conclusive support for the effectiveness of teacher mentoring. Cross-sectional data from the nationally-representative 2015-16 National Teacher and Principal Survey (NTPS) was used to empirically evaluate the relationship between school-based mentoring programs and first-year teachers' risk for occupational stress, while addressing some of the common limitations found in research relating to teacher mentoring and occupational stress. Results cautiously suggest that formal mentoring programs play an important role in beginning teacher wellness. Implications and future research directions are discussed.

Table of Contents

List of Tables

List of Figures

Introduction

Over the past several decades, mentoring programs for beginning teachers have surged in popularity across the United States. School-based mentoring programs, wherein novice teachers (mentees) are assigned to and supported by more advanced teachers (mentors), have become a favorite of policymakers and administrators alike. Increasingly, mentoring is employed as the first line of defense against a multitude of challenges faced by beginning teachers, with the ultimate aim of reducing turnover among this group (Ingersoll & Strong, 2011; Little, 1990).

Indeed, beginning teachers typically face an abundance of potential stressors: they often are assigned to the most challenging classrooms; are frequently tasked with unwanted, extracurricular responsibilities of more seasoned teachers; and routinely suffer from dashed career expectations (DePaul, 2000; Grissom, Kalogrides, & Loeb, 2015; Herbert & Worthy, 2001; Weasmer & Woods, 1998). Perhaps unsurprisingly, it is beginning teachers who are the most likely to leave their schools or quit the profession. Recent longitudinal data suggests that approximately 44 percent of public and private school teachers leave the field within their first 5 years (Ingersoll et al., 2018). This trend of beginning teacher turnover has been on the rise since the late 1980s, with no indication of slowing down. Concerningly, the potential consequences of turnover are thought to be substantial and wide-reaching; these include the loss of financial capital, depletion of human resources, erosion of staff cohesion and community, and reduction in student achievement (Ronfeldt, Loeb, & Wyckoff, 2013).

Recent research points to occupational stress as one potential driver of teacher turnover (e.g., Farber, 1991; R. Goddard, O'Brien, & Goddard, 2006; Grayson & Alvarez, 2008; McCarthy, Fitchett, Lambert, & Boyle, 2019). Given the caustic influence of high turnover on teachers, schools, and students (Guin, 2004; Hong, 2012; Ronfeldt, Loeb, & Wyckoff, 2013), it is critically important to gain a more thorough understanding of beginning teachers' risk for occupational stress, and to identify factors which might impact the likelihood of a new teacher being at risk for stress in their workplace. Whether and how mentoring programs impact beginning teachers' risk for stress remains unclear.

Given the significant time and capital invested annually into the development and execution of formal mentoring programs, there is a vested interest among policymakers and administrators to better understand what types of mentoring teachers are receiving, and what effects these mentoring supports have on teacher stress risk. To this point, although mentoring is increasingly hailed as one of the most – if not the most – critical components of teacher induction programs, the corresponding research base is still working towards providing conclusive support for the effectiveness of mentoring.

Although the literature seems to very broadly suggest that mentoring is effective in mitigating turnover, is perceived as helpful amongst mentees, and leads to improvements in beginning teacher occupational health and performance (Ingersoll & Kralik, 2004; Ingersoll & Strong, 2011; Spooner-Lane, 2017), many mentoring studies are undermined by appreciable limitations and/or flaws. These include small sample size, limited generalizability, lack of a comparison group, inadequate control variables, failure to examine the individual components of mentoring, examination of samples other than first-year teachers, and the evaluation of stress from non-transactional frameworks.

The current study used cross-sectional data from the nationally representative 2015-16 National Teacher and Principal Survey (NTPS) to empirically evaluate the effects of school-based mentoring programs on first year teachers' risk for occupational stress, while addressing some of the common limitations found in research relating to teacher mentoring and occupational stress. Specifically, binomial logistic regression was used to explore the link between participation in various components of mentoring programs and the likelihood of being classified as Demanded (i.e., at risk for occupational stress).

Literature Review

Career Phases of Teachers

Models of teachers' career phases tend to group teachers according to their perceptions of workplace climate (e.g., successes and difficulties), placing them at different stages according to said perceptions (McCarthy, Fitchett, Lambert, & Boyle, 2019). Early on, developmental models tended to focus on teachers who remained in the field for extended periods, rather than examining teachers with different career lengths (Burke, et al., 2015; Fessler, 1992; Katz, 1972). These early stage theorists posited that beginning teachers tend to achieve mastery of their craft relatively quickly. According to Katz' (1972) stage theory, teachers move through a five year cycle of professional mastery and commitment, progressing from survival to a period of renewal and, finally, professional maturity. However, as more variable career lengths have become commonplace in the field, models of teachers' professional phases have changed. For example, Hargreaves (2005) developed a three-level model of career development, wherein the following stages are defined: early career (0-5 years of experience), mid-career (6-19 years) and later-career (20 or more years). Given that in recent years, a large proportion of teachers fail to remain in the field beyond the first five years of their career, this three-stage model is better-aligned with current trends in the educational workforce. Other stage models define early-career teachers (ECTs) as being in their first three years; however, the essence of these frameworks are very similar. Hargreaves (2005) paints the early-career stage as a challenging, critical period for identity development, positing that beginners are attempting to "establish their basic confidence

and competence as professionals" (p. 970) while confronting the many demands present in the job.

Indeed, research suggests the beginning years of teaching are particularly formative; a teacher's earliest experiences can have lasting effects on future levels of occupational commitment, job satisfaction, and retention (Doney, 2013; Goddard & Foster, 2001; Hoy & Spero, 2005; Lortie, 1975; Rust, 1994). A teacher's first year – a time of role transition, professional identity formation, knowledge accumulation, and confrontation of reality – is regarded as especially difficult (Hebert & Worthy, 2001; Jenlink, 2014; Lortie, 1975). Influenced by their collective experiences as student teachers, beginners enter the field with a set of expectations – often quite idealistic – regarding their roles, responsibilities, and abilities (Gaede, 1978; Lortie, 1975). In reality, there is often a stark disconnect between these expectations and the actual experiences of a newly-qualified teacher, prompting many novices to label their first year a "painful beginning" (Huberman, 1989) or "reality shock" (Hebert & Worthy, 2001; Veenman, 1984). While experiences certainly vary, many first-year teachers find themselves assigned to the most challenging classrooms, and frequently inherit the unwanted, extracurricular responsibilities of their more experienced colleagues (DePaul, 2000; Grissom, Kern, & Rodriguez, 2015; Weasmer & Woods, 1998). Furthermore, beginners routinely report feeling bewildered by unexpected gaps between academic theories of teaching and the practical realities of the job (Gaede, 1978).

Trends in the Workforce

The landscape of the U.S. elementary and secondary teaching workforce has undergone dramatic changes in recent years. Most germane to the current discussion are changes in teacher experience level and rates of beginning teacher turnover.

Greening of the Field

One of the most notable trends in recent years relates to teacher demographics; specifically, the proportion of beginning teachers has increased. Ingersoll and colleagues (2018) refer to this phenomenon as the 'greening' of the field. Greening is primarily the result of a substantial increase in the hiring of new teachers, a trend which picked up in the late 1980s, temporarily decreased during the 2008-2012 financial crisis, and resumed in 2013. By the 2015-16 school year, the modal teacher was an early-career teacher in their first three years on the job (Ingersoll, 2018). The large pool of beginning teachers would be rather unremarkable if not for the next trend to be discussed: high rates of beginning teacher turnover.

Beginning Teacher Turnover

While teacher turnover is a widely recognized problem within the entire teaching field, it is particularly marked amongst beginning teachers. This is unsurprising due to the first years of teaching – particularly the first – being particularly challenging to navigate and endure. As previously mentioned, beginning teachers comprise the largest group of U.S. public school teachers, and it is this group which leaves the field at the highest rates. Over 40% of teachers leave the field within the first five years of their career (Ingersoll et al., 2018). Turnover is a form of occupational instability wherein teachers move out of their schools or districts, or leave the field entirely. The term 'attrition' refers specifically

to those teachers who leave teaching altogether. Regardless of the type of turnover – moving or leaving – this instability is thought to have negative effects on teacher morale and student outcomes (Guin, 2004; Kraft & Papay, 2014). Indeed, Ronfeldt et al. (2013) suggest that this occupational instability disrupts entire school organizations, negatively impacting both teachers and students who remain in the school that a teacher has left. While some turnover is expected and even beneficial – weeding out teachers who are not capable of remaining in the profession – an excess of attrition and moving between schools and districts is clearly detrimental to the field (Ingersoll et al., 2018).

Formal Support Programs for Beginning Teachers

In an effort to combat the occupational issues associated with being a new teacher (e.g., high risk of turnover, experiences of disillusionment), many states and districts across the nation have employed school-based induction programs (of which mentoring is typically a major component). In the following section, I will provide an overview of induction, followed by a deeper dive into mentoring programs.

Induction Programs

Teacher induction programs are amongst the most popular tools employed by policymakers and administrators for improving the quality of classroom instruction and reducing turnover among beginning teachers (Kapadia et al., 2007). Kapadia, Coca, and Easton (2007) describe induction as “a means to orient, assist, and guide beginning teachers so they remain in the profession and grow into capable practitioners” (p.4), with the indirect goal, of course, of ultimately improving student outcomes (e.g., personal growth, academic achievement). Induction must be differentiated from both pre-service and in-service programs; pre-service refers to programs aimed at preparing teaching

candidates prior to their entering the field, while in-service programs are a form of continuing education offered over the entire course of a teacher's career (Feiman-Nemser, 2001). Induction is unique in that it is specifically geared toward supporting novices who are transitioning into their new roles as classroom teachers (Goldrick, 2009; Smith & Ingersoll, 2004). Typically part of the first 2 or 3 years of a teacher's career, induction is designed to facilitate and ease the transition from student to teacher, and move beyond simply obtaining knowledge/skills into actively applying this information in the classroom (Feiman-Nemser, 2001). In other words, induction denotes a process intended to support, orient, develop, and guide novice teachers, and is often packaged in the form of a formal program (Ingersoll & Strong, 2011; Wong, 2004). The importance of induction is made clear by Goldrick (2009), who notes, "The learning that takes place in a beginning teacher's first few years on the job is different from their preparation experience and is different from their subsequent professional learning. There is no period as formational to a teacher's career as the initial years in the classroom" (p. 3).

The specific goals and components of teacher induction programs tend to vary widely. For example, in addition to the common goals of improved teacher performance and retention, programs may emphasize teacher socialization, role adjustment, professional development, and/or evaluation (Ingersoll & Strong, 2011). The components of an induction program can include additional classroom assistance, workshops, collaboration with faculty, lighter workloads, and – most commonly – mentoring (Ingersoll & Strong, 2011).

Mentoring

Over the past three decades, mentoring has emerged as the leading form of teacher induction (Fideler & Haselkorn, 1999, as cited in Ingersoll and Kralik, 2004). In fact, mentoring is such a standard component of teacher induction programs that the terms ‘mentoring’ and ‘induction’ are often used synonymously (Smith & Ingersoll, 2004). However, induction is the overarching process, while mentoring is typically a key facet of induction. In the context of education, a mentoring relationship typically consists of a novice teacher (mentee) paired with a more experienced teacher (mentor) (Hobson, Ashby, Malderez, & Tomlinson, 2009). For the purposes of the current study, I will use ‘mentoring’ to describe one-on-one, school-based support for beginning teachers; this excludes group mentoring (also referred to as peer mentoring), which involves mentorship of more than one mentee (e.g., Cornu, 2005; Mitchell, 1999).

Organizational literature typically delineates three primary categories of mentoring support: psychosocial support, career support, and role modeling (Chun, Sosik, & Yun, 2012). Psychosocial support is associated with the interpersonal relationship between mentor and mentee, and can include validating mentee behavior, providing companionship, helping mentees adjust to the profession, and discussing occupational challenges and experiences (Chun et al., 2012; SREB, 2018). Indeed, supportive mentoring often involves sharing interpersonal takeaways related to negative relationships teachers sometimes encounter when working with colleagues, students, and parents (Prilleltensky, Neff, & Bessell, 2016). Mentors providing career support might help mentees navigate organizational complexities, provide assignments meant to challenge mentees and foster their occupational growth, or offer job-related coaching

(Chun et al., 2012). Many mentors specifically provide support with regard to not only surviving, but thriving, the often turbulent, complex, and demanding workplace climate found in schools (Smith & Ingersoll, 2004). The third type of support, role modeling, is an opportunity for the mentee to expose the mentor to the work-related values, perspectives, and behavior they (the mentor) have cultivated over the course of their career (Chun et al., 2012).

As with induction programs, there is tremendous variation in how mentoring programs are designed and executed (Ingersoll & Strong, 2011). While the overall objective of providing guidance and support to novice teachers is relatively ubiquitous, mentoring programs often differ in terms of their specific goals, duration, strategies, components, execution, and timing of implementation (Hobson et al., 2009; Ingersoll & Strong, 2011). For example, program length can vary from one stand-alone mentor/mentee meeting to regularly scheduled sessions (Ingersoll & Strong, 2011). Furthermore, programs can differ in terms of mentor compensation, mentor selection, mentor training, emphasis placed on matching mentors and mentees by subject area of expertise, amount and type of mentee support, and overall program efficacy (Ingersoll & Strong, 2011).

Due to the large degree of variation that currently exists across mentoring programs, there is a clear need to empirically examine the types of mentoring teachers are receiving, and to determine how effective these various components are. Policymakers have a vested interest in the findings from such studies, as their funding of such programs has increased dramatically since the 1980s (Hobson et al., 2009). Approximately one-third of U.S. states have mandated beginning teacher participation in

formal induction programs (Roehrig, Bohn, Turner, & Pressley, 2008), and the percentage of new teachers taking part in induction has skyrocketed from 41% in 1990 to nearly 79% in 2000 (Ingersoll & Smith, 2004). In spite of induction – and particularly mentoring – being heralded as a possible panacea for early career teacher turnover, there are indications that these programs have not consistently met the goals and/or requirements of their accompanying mandates (Bradley & Gordon, 1994, as cited in Roehrig et al., 2008). For example, despite formal mandates in one U.S. state, only 75% of their school districts reported direct observation of a mentee's classroom teaching by a mentor, and a mere 70% instituted mentor-mentee workshops prior to the beginning of the school year (Roehrig et al., 2008). That being said, even when comprehensive mentoring programs are successfully executed, the helpfulness of the individual program components/activities remain unclear. In the next section, I will provide an overview of the research on mentoring programs, and will conclude with a summary of the major limitations and gaps in the research base.

Empirical Research on Mentoring Programs for Beginning Teachers

Kapadia, Coca, and Easton (2007) used self-report measures to cross-sectionally examine mentoring activities with 1,737 beginning (first or second year) teachers within Chicago Public Schools (CPS). Specifically, they described the quality and intensity of mentoring supports, and examined the effect of mentoring activities on new teachers' job satisfaction and occupational commitment. In terms of frequency of mentoring, approximately 70% of the CPS novice teachers – where 'novice' describes teachers in their first or second year of teaching – reported meeting with their assigned mentors at least once every two weeks, with the overwhelming majority of

respondents feeling satisfied with the amount of face-to-face contact with their mentor. Novice elementary and secondary teachers found their mentors helpful in terms of teaching strategies, classroom management, and classroom observation, but found them less helpful with regards to parent communication, student assessment, and analyzing student work. Beginning elementary teachers reported receiving a greater quantity of mentoring than beginning high school teachers. Hierarchical generalized linear models with teachers nested in schools were used to examine three self-reported outcomes: the quality of a teacher's experience in their first year, intention to remain in the field, and intention to remain in the same school. The authors controlled for background characteristics of teachers, classrooms, and schools, incorporating working conditions variables that could be related to the three outcomes. Teacher-level control variables included gender, race/ethnicity, prior work experience, alternative certification, and highest level of education. Classroom-level controls included class size, percentage of students who lack knowledge and skills, percentage of students enrolled in a bilingual program, and percentage of students who create serious behavior problems. School-level controls included racial/ethnic composition of the school, school-level concentration of poverty, SES of students' communities, percentage of new teachers on staff, teacher retention, school leadership, teacher socialization, school size, and school-level achievement. SES was not applied to high school models, as it was highly correlated with the racial/ethnic composition of the school. Elementary teachers whose mentors observed their classroom teaching and discussed classroom management techniques were most likely to report a positive teaching experience. Mentoring activities most strongly and positively related to teachers' intention to remain in the same school were

developing teaching strategies, discussing classroom management, and assisting with student assessment. Only one mentoring activity – analyzing student work – was linked with high school teachers' intention to remain in the same school. No specific mentoring activities were strongly associated with beginning elementary or secondary school teachers' intention to remain in the field. Rasch analysis was used to develop an overall measure of the quantity and quantity of mentoring activities (*Did not receive; Received, not helpful; Received, somewhat helpful; Received, very helpful*; separation: 2.48, reliability: 0.86). From these responses, teachers were grouped according to the level of mentoring support received: *average, weak,* and *strong.* Of teachers assigned a mentor, only around one-fifth of beginning high school teachers and one-quarter of beginning elementary school teachers received strong levels of mentoring, while on average, beginning teachers reported receiving weak or average levels of mentoring. The authors concluded that there should be an emphasis on how mentors are selected and trained, and found that teacher collaboration and principal support are particularly beneficial to new teachers. While this study was strong overall, particularly with its use of covariates, the generalizability of its findings on a national level can be called into question due to its focus on Chicago Public School teachers. Furthermore, the cross-sectional design prohibited examination of actual turnover behavior.

A number of studies have made use of large-scale, nationally representative datasets. As with the current study, these studies utilized data from the National Center for Education Statistics (NCES). Smith and Ingersoll (2004) utilized the 1999-2000 Schools and Staffing Survey (SASS) and its longitudinal follow-up component, the 2000-2001 Teacher Follow-up Survey (TFS), to collect self-report data from 3,235 first-year

teachers. A major strength of studies using datasets such as the SASS or NTPS is the boost to the external validity that comes with a nationally-representative sample. The referenced study sought to provide descriptive information on how many first-year teachers were participating in various types of induction activities, the helpfulness of a teacher's mentor, and to explore the link between mentoring/induction and beginning teacher turnover. Multinomial logistic regression was used to examine the impact of receiving induction supports with the likelihood of teachers moving schools, remaining in the same school, or leaving the field following the completion of their first year of teaching (represented by a relative risk ratio) after controlling for teacher- and school-level covariates. This method accounted for the nested structure of the data (i.e., teachers clustered within schools). Four sets of predictors were examined on the aforementioned outcomes: teacher-level and school-level characteristics, participation in mentoring activities, participation in group induction activities, and the administration of extra resources for beginning teachers. The study included many of the same teacher- and school-level covariates as Kapadia and colleagues (2007), with the addition of teacher age, full/part-time teaching status, subject taught, and income for teacher-level factors, and urbanicity, sector (i.e., public non-charter, public charter, private), and percentage of students eligible for free or reduced priced lunch for school-level factors. Three sets of measures were used to assess participation in induction and mentoring programs; the mentoring-specific measure collected data on whether the teacher was paired with a mentor and whether the mentor taught the same subject as the mentee, while the other measures assessed group induction activities and reduced workload and/or additional resources. Approximately 70% of beginning public school teachers reported working with

a mentor. Among mentees, around 70% were matched with a mentor who taught the same subject, and around 90% reported that their mentor was helpful. Results from the logistic regression revealed that a subject area match between mentor and mentee significantly reduced the risk of leaving the field at the 90% level of confidence (*rrr* = 0.704, *p* = 0.084). Mentorship was not significantly associated with the likelihood of changing schools. When the collective impact of all mentoring and induction variables were assessed simultaneously, it was shown that teachers rarely receive one single type of support; rather, they often are given a host of mentoring and induction supports. No individual mentoring or induction support coefficient was statistically significant, rendering it difficult to isolate the unique effect of each component. When joint effects of induction and mentoring components were analyzed, it revealed that the more induction supports a teacher received, the greater the reduction in risk for turnover. However, as the number of supports increased, the number of teachers receiving these supports dropped. Results indicated that having a mentor in the same field as the mentee was a salient factor in reducing turnover risk for beginning teachers. Furthermore, teachers who were part of combined mentoring/induction programs were less likely to change schools or leave the field than teachers who did not participate in these types of programs. Overall, this study provides support for the effectiveness of mentoring and induction programs in reducing beginning teacher turnover.

A recent study by Ronfeldt and McQueen (2017) analyzed national NCES data – specifically, the SASS, BTLS, and TFS – to estimate the likelihood of teachers moving schools in their second year based on the induction supports (including mentoring) they received as first-years. Of their sample of 2,340 full-time and part-time public school

teachers, 79% reported having a mentor. Using multilevel logistic regression, with teachers nested in schools, the authors found that having a mentor as a first-year teacher significantly reduced the odds of subsequent turnover (specifically, migrating schools) in the following year by 41% - 51%. The authors point out that while having a mentor seems to positively impact beginning teachers, it is likely that certain components and quantities of mentoring are more helpful than others. Therefore, they suggest a more granular study of teacher mentoring and its component parts.

Ingersoll and Strong (2011) conducted a critical review of beginning teacher induction and mentoring research extending back through the mid-1980s. Candidates for inclusion in the review – 500 studies in total – were located through a search of educational databases using search terms including 'beginning teacher,' 'induction,' 'mentoring programs,' and 'teacher mentors.' Of this initial set, 150 were excluded because they did not meet criteria for empirical research focused on beginning teacher induction and mentoring. Finally, a total of 15 studies were selected based on the following inclusion criteria: (a) the study was evaluative rather than descriptive; (b) the aim of the study was to evaluate the effects of induction on at least one outcome (i.e., particular induction components, activities, or programs); (c) outcome variables represented effects of induction programs on beginning teachers or students (e.g., studies were excluded if the outcome(s) of interest concerned mentors rather than mentees). Additionally, studies were required to contain thorough information on data sources, methods, sample sizes, and outcomes. The bulk of studies that were ultimately selected compared teachers based on their degree of participation in induction and mentoring programs, rather than comparing participants to non-participants. Within the

final 15 studies, there are three general clusters of study outcomes: (a) occupational commitment, job satisfaction, turnover, and retention; (b) teaching methods and practices; and (c) student achievement. Compared to more common educational research topics, such as school size, there is relatively little research on early career teacher induction and mentoring (Ingersoll & Strong, 2011). Thus, while Ingersoll and Strong's (2011) review does not exclusively concern mentoring programs, the comprehensive nature of their review makes it critical to describe their key findings in this literature review. Overall, findings from the 15 studies suggest that induction programs – especially the teacher mentoring component – are beneficial to beginning teachers (Ingersoll & Strong, 2011). While each study was noted to have limitations, Ingersoll and Strong (2011) were able to present several unified conclusions. First, induction programs – particularly mentoring programs – appear to benefit early-career teachers' occupational health. In almost every study, beginning teachers participating in induction/mentoring programs reported elevated occupational commitment, job satisfaction, and retention. Second, participation in induction/mentoring programs seems to positively impact beginning teachers' classroom practices (e.g., improved performance related to lesson plan development, classroom control, interacting with students, and keeping students on task). Finally, in almost every study, mentoring was associated with increased student achievement (e.g., higher achievement test scores). One study produced contradictory findings: a randomized controlled trial conducted by Glazerman and colleagues (2010). While the study found significant differences between treatment and control groups on student achievement, there were no significant group differences on first-year teacher classroom practices or retention over a new teacher's first several years. The lack of findings with

regard to classroom practices and retention stand in stark contrast to the findings of the other studies in the review. Ingersoll and Strong (2011) offer several possible explanations for the contradictory, mixed findings. First, differences in the length of programs could impact results: Glazerman et al. (2010) found that differences in student achievement only appeared after two years of induction/mentoring participation, yet sampled teachers were only observed during the Spring semester of their first year of teaching. The other studies largely conducted multiple observations of teachers over a longer period of time, with the largest study observing teachers after their second year. As previously mentioned, these studies all found significant positive effects of induction and mentoring on classroom practices. Therefore, it could be the case that more than one year of induction is required to create meaningful change in a teacher's professional development. Another possible explanation for the different results could relate to the inherent limitations in the external validity of randomized controlled trials. In the Glazerman et al. (2010) study, the sample was limited to teachers in large, urban public school districts, and thus may not adequately represent the full national teacher workforce.

A thorough examination of the empirical research base on beginning teacher mentoring unearths a series of recurring flaws and limitations. First, several studies failed to collect information on the specific components and characteristics of mentoring and induction programs, which are elements that could impact apparent variation in outcomes. Another common – and critical – limitation was the failure to include covariates that could contribute to variation in the outcomes of interest. Strong studies include a host of teacher-level and school-level factors, in addition to controlling for organizational/working

conditions (e.g., degree of student discipline issues). In addition to the issue of potential confounding, non-randomized designs which did not include covariates also failed to address potential selection bias. Kee (2012) offers a helpful discussion of methods for addressing the issue of selection bias, including the use of multiple teacher-level demographic and background covariates. This approach can help strengthen any causal inferences that are made (Kee, 2012). According to Elze et al. (2017), covariate adjustment performs at a level on par with propensity score matching in terms of creating reliable, robust estimates. Small sample size was an issue with several studies. Additionally, a number of observation-only studies were included; while these have the opportunity to provide rich, unique information about teacher practices, potential threats to reliability and validity are inherent to such designs. A major issue in the research is the frequent lack of a comparison group. Specifically, many studies did not compare teachers who did and did not receive mentoring/induction, but rather examined teachers who received different types or amounts of supports. These studies, then, could not provide findings on the effects of mentoring itself, but rather explored whether a certain program format led to different, enhanced effects on outcomes of interest compared to another program format. With regards to internal validity, cross-sectional designs with an interest in teacher turnover generally used self-report measures of occupational commitment (i.e., how likely a teacher feels they are to stay in the same field and/or school). A potential limitation of these types of items is the uncertainty with regards to how strongly these ratings correlate with actual turnover behavior (Ingersoll & Strong, 2011). Finally, many studies were limited by their use of local samples and/or non-representative sampling,

which could lead to biased results. Conversely, studies that utilized large-scale, nationally representative datasets benefitted from their complex, multi-stage sampling designs.

Overall, this review reveals limitations and research gaps in the existing teacher mentoring research. In terms of the content of mentoring programs, much of the research does not help forward the understanding of what particular components are most helpful to beginning teachers. For example, with regard to mentoring programs, what particular elements (e.g., one-on-one meetings, help with lesson plan development) seem to benefit teachers the most? Additionally, while the quantity of induction/mentoring programs appears to be important, questions remain regarding how variations in the intensity of these programs impact study outcomes. For example, there may be differences in teacher effectiveness depending on how much contact a teacher has with their mentor. Finally, there is the important issue of context: how much do the effects of induction/mentoring vary across school settings? It is possible, for instance, that what makes for an effective program in a rural setting may differ from the critical components in an urban setting. Additionally, it may be that induction and mentoring programs are particularly beneficial for teachers with alternative routes to certification, but are less helpful for teachers with more thorough preparation. Furthermore, there is a need for nationally-representative studies that provide a more recent snapshot of teacher mentoring in U.S. public schools. With regards to outcome variables, there is a healthy amount of research that examines the impact of mentoring on beginning teachers' job satisfaction, occupational commitment, and retention. However, very few studies examine the effect of mentoring programs on beginning teachers' risk for occupational stress (particularly from a transactional framework). In the next section, I will describe the phenomenon of

occupational stress risk, and will summarize research findings that are relevant to the current study.

Teacher Stress

Teaching is generally considered to be among the most stressful occupations, with studies consistently demonstrating higher levels of self-reported workplace stress among teachers than in any other occupation (Dicke et al., 2014). Early-career teachers are thought to be particularly vulnerable to the many demands of teaching; in a recent study of beginning U.S. public school teachers in full-time positions, Fitchett et al (2018) found that approximately one-quarter (24.4%) of first-year teachers were at risk for stress. Teacher workplace stress has become a popular focus within educational research, as it correlates with a host of critical occupational health indicators including job satisfaction, burnout, occupational commitment, workplace performance, and turnover rates (Chang, 2009; Fitchett, McCarthy, et al., 2019; Gilbert, Adesope, & Schroeder, 2014; Gray & Brauen, 2013; Jennings, Frank, Snowberg, Coccia, & Greenberg, 2013; Skaalvik & Skaalvik, 2011).

There is no consistently used, unanimously accepted definition of stress, and many studies have exclusively focused on the role of objective, external workplace factors related to stress (e.g., class size, salary) (French, 1993; Loeb, Darling-Hammond, & Luczak, 2005; Zellars, Hochwarter, Perrewé, Hoffman, & Ford, 2004). While these studies have undeniably offered contributions to the field, they fail to consider the more subjective, perceptual elements involved in the stress process. These particular elements, which are well-represented within the transactional model of stress, will be discussed next.

Transactional Model of Teacher Stress

The transactional model of stress (Cox, 1978; Lazarus & Folkman, 1984) provides a framework for examining how teachers appraise their work environment, and for identifying when these appraisals are likely to render the teacher at risk for stress. The transactional model is one of the most dominant stress frameworks across disciplines. A major asset of the transactional model is that it privileges teachers' subjective perceptions of their workplace environment in assessing their vulnerability to stress, thus capturing the subjective, cognitive component that purely objective models (as previously mentioned) lack. The basic assumption underlying the transactional model is that psychological stress results from the appraised imbalance of perceived demands vis-à-vis resources (Cox, 1978; Lazarus & Folkman, 1984). Thus, teachers who appraise their available workplace resources as insufficient compared to workplace demands are theoretically at greater risk for stress than teachers who either: a) perceive their workplace demands and resources as being approximately balanced; or b) do perceive an overall imbalance, but view their resources as sufficient compared to workplace demands (McCarthy, Lambert, Lineback, Fitchett, & Baddouh, 2016). While the transactional model has inspired many studies of teacher stress, researchers generally do not adequately operationalize the model's fundamental principle that psychological stress results from the perceived imbalance between available demands and resources. As will be described in the following section, Lambert and colleagues (2009) created a useful approach to effectively operationalize the transactional model: the CARD.

Measuring Risk for Stress Within a Transactional Framework

Measuring a teacher's stress risk in terms of perceived school/classroom demands and resources was made possible through the creation of the Classroom Appraisal of Resources and Demands (CARD, Lambert, McCarthy, O'Donnell, & Wang, 2009). The focus of the CARD is intentionally placed on risk for stress rather than the stress response itself, primarily because symptoms of stress are more situation-dependent and transient as compared to the more enduring nature of risk for stress. The Classroom Demands scale (α =.92) measures perceived demands along the following categories: problematic student behaviors, other student-related demands (e.g., poor attendance), administrative demands, and lack of instructional resources (e.g., availability of teaching supplies). Perceived resources are assessed via the Classroom Resources scale (α =.94), which surveys the subjective helpfulness of the following areas of classroom resources: availability of school support personnel (administrators and classroom aides), other adults (community volunteers), instructional support materials, and specialized instructional resources (e.g., resources for children performing below grade level). The Resources and Demands scales yield distinct information, as evidenced by low correlations across multiple studies (r = -.208 to -.080).

The CARD operationalizes the transactional model by calculating a difference score, labeled the Appraisal Index score, between a teacher's Demands scale score and Resources scale score. The Appraisal Index score – named such because it was explicitly designed to capture the overall appraisal of a teachers' perceived resources vis-à-vis demands – serves as a continuous measure of a teacher's risk for occupational stress. Higher appraisal index scores indicate increased risk for stress. A review of 18 studies

using the CARD with local samples of teachers demonstrated support for the validity of using CARD Appraisal Index scores as proxies for stress risk: teachers with the highest Appraisal Index scores (i.e., greatest imbalance of classroom demands versus resources) reported lower occupational commitment, lower job satisfaction, and more student behavior concerns than did teachers with lower scores (McCarthy, Lambert, & Reiser, 2014). In addition to the aforementioned occupational health outcomes, associations have also been demonstrated between CARD scores, burnout symptoms, and occupational mobility (McCarthy et al., 2016).

The CARD also allows for the classification of teachers into groups based on their comparative risk for stress, serving as a convenient tool for identifying those teachers who are likely to be most at risk for stress (and, by extension, at greatest risk for the other occupational health problems associated with stress risk, including burnout, turnover, and low job satisfaction). To construct the groups, cut-scores are used to distinguish the boundaries of each classification group (cut scores are at the upper and lower limit of the 95% confidence interval around zero difference), resulting in three groups that differentiate teachers according to their risk for stress: (1) teachers perceiving classroom resources as greater than demands (Resourced group), (2) teachers perceiving classroom demands as equal to resources (Balanced group), and (3) teachers perceiving classroom demands as greater than resources (Demanded group). According to the transactional model of stress, teachers who perceive an imbalance wherein they feel more “demanded” than “resourced” are theoretically expected to be at the greatest risk for stress, followed by individuals who perceive demands and resources to be approximately balanced, followed by teachers who perceive themselves as more

“resourced” than “demanded” (McCarthy, Lambert, & Reiser, 2014). These three stress-risk classification groups – Resourced, Demanded, and Balanced – have been shown to differ meaningfully from one another with regards to their risk for stress (McCarthy et al., 2016).

While the CARD was initially used to examine local teacher samples (McCarthy, Lambert, et al., 2016), researchers have since successfully applied the CARD method of operationalizing the transactional model to nationally representative teacher samples. The gains in generalizability that come with producing national over local estimates can expand the potential impact a study has on greater society, as findings from such studies can inform educational policy and decision-making on a national level. Using the nationally representative Schools and Staffing Survey, Lambert et al. (2015) identified items on the SASS that were similar to those on the CARD, and then used item responses to create a proxy variable for measures of classroom demands and resources. Teachers were reliably classified into appraisal groups based on their responses to this subset of SASS items. In subsequent studies, the CARD scoring and classification procedure has also been successfully applied to data from the Beginning Teacher Longitudinal Study (BTLS), an extension of the SASS, and from the National Teacher and Principal Survey (NTPS), a follow-up to the SASS; meaning that the transactional model of stress has now been successfully operationalized in both local and national samples of teachers. This is significant because, among transactional stress researchers, only a handful have examined teacher stress on a national scale (Lhospital & Gregory, 2009). However, there still is scant transactional research that is both nationally representative and that focuses on first-year teachers.

Research on Beginning Teachers' Risk for Stress

While some novice teachers undoubtedly thrive upon first facing the demands of teaching (e.g., accountability, bureaucracy, student contexts) many beginners perceive these demands as outweighing the resources allocated to them (e.g., professional development, mentoring, staff/administrative support, financial incentives). Many novice teachers report feeling overwhelmed by their new workplace environment, and frequently respond by entering into "survival mode" (Costigan & Crocco, 2004; Katz, 1972). Perhaps unsurprisingly, beginning teachers as a group have some of the highest rates of occupational burnout and turnover (Ingersoll, 2003). Previous transactional stress research has demonstrated that when teachers appraise their classroom demands as higher than their available resources – i.e., are at risk for occupational stress – the consequences can include occupational burnout (Chang, 2009) reduced job satisfaction, and decreased occupational commitment (Gilbert et al., 2014). Thus, knowing how a teacher is classified in terms of stress risk (Resourced, Balanced, or Demanded) can be a helpful tool for identifying teachers who might need more support in order to avoid the likely repercussions in the areas of job satisfaction, burnout, commitment, and turnover.

This research extends to early-career teachers: McCarthy and colleagues (2019), looking specifically at beginning teachers, found that being classified as Demanded in one's first year of teaching can lead to a breakdown in occupational commitment in later years. Specifically, first-year teachers in the Demanded group were more likely than their Balanced and Resourced colleagues to move schools in their second and fourth years, and by year three, they were departing the profession entirely at higher rates than non-Demanded colleagues.

Are school-based mentoring programs a potential panacea for the numerous occupational ills faced by beginning teachers? As previously mentioned, there is research indicating mentoring programs can positively impact teachers and reduce their risk for turnover (Ingersoll & Kralik, 2004; Smith & Ingersoll, 2004). Perhaps there is a relationship between receiving mentoring as a first-year teacher and risk for stress.

Research on Mentoring and Beginning Teacher Stress

While the research exploring the relationship between mentoring programs and beginning teacher stress is scant, two recent studies investigate this link. Fitchett, McCarthy, Lambert, and Boyle (2018) analyzed cross-sectional data from first-year teacher respondents to wave 1 of the NCES Beginning Teacher Longitudinal Study (BTLS) to examine the relationship between first-year teacher supports (including mentoring) and risk for occupational stress. Data from wave 1 were taken from the 2007-08 SASS, and included items assessing workplace climate, teacher attitudes, new teacher supports, and demographics of first-year teachers and their schools. The study sample (n = 1,760; sample size rounded to nearest 10 per NCES data disclosure protocol) was restricted to full-time public elementary and secondary public school teachers. Each first-year teacher's stress grouping (Resourced, Balanced, or Demanded) was analyzed in relation to the following six following new teacher supports: access to a mentor, common planning time, seminars, additional help, and supportive communication. ANOVA results from the analysis of teachers' supports across stress group classification indicated that first-year teachers classified as Demanded reported different new teacher supports – except for access to a mentor – compared to Resourced teachers. While there were descriptive differences, there were no statistically significant differences between

Resourced teachers (85.1%) and Demanded teachers (77.4%), [$\chi^2(2, 1760) = 9.86$, $p > .05$] in terms of their access to a mentor. The authors had hypothesized that Demanded teachers would report different – specifically, lower – levels of support than Resourced teachers; they made sense of these findings by pointing out that a large portion of U.S. teachers have access to mentors, which may account for the lack of detectable differences. The authors suggest, however, that differences in stress risk may be unearthed when mentoring programs are examined on a more granular level; specifically, they suggest looking at the different components of mentoring rather than simply whether or not a teacher had a mentor.

A second study which included results specific to occupational stress risk and teacher mentoring was conducted by McCarthy, Fitchett, Lambert, and Boyle (2019). This study was a longitudinal companion study to Fitchett et al. (2018), and used BTLS data with the overarching purpose of determining whether first-year teachers classified as Demanded were more likely to move schools or leave the field (i.e., exhibit turnover behavior) at any point over the next four years (data was collected over five waves between 2007 and 2012) than were teachers classified as Balanced or Resourced. A component of this study tested whether the association between Demanded group membership and turnover status remained after controlling for several teacher-level demographic covariates, one of which was having a mentor. In addition, all analysis held constant the following individual-level teacher and school characteristics: teachers' self-reported ethnicity, gender, certification type, status as a special education teacher, perceived quality of teacher training programs, and aspects of the school setting (i.e. urban vs. suburban, rural vs. suburban, Title I, and school percentage of students

qualifying for the National School Lunch Program). The sample consisted of full-time first-year teacher respondents the BTLS (n_{Wave1} = 1,760). The percentage of teachers who reported having a mentor remained consistent across all five waves, ranging from 81.6% – 82.8%. Logistic regression models were applied to examine whether Demanded group membership was associated with turnover status, as measured by membership in the 'total leavers' or 'stayers' group for each wave ('total leavers' included teachers who moved schools or left the field over the course of the survey, while 'stayers' included teachers who remained in the same school or returned to teaching over the course of the survey). The outcome of interest was membership in the Demanded group, with a comparison category consisting of teachers classified as either Balanced or Resourced. For waves II and V, membership in the Demanded group was significantly associated with being in the 'total leavers' group, prior to adding teacher-level demographic covariates to the model. For wave II, the adjusted Wald statistic for the overall model was statistically significant ($F(1,88) = 11.338$, $p < .001$), with an odds ratio indicating that teachers classified as Demanded were 2.342 times more likely to be in the 'total leavers' group than teachers classified as either Resourced or Balanced. For wave V, the Bonferroni adjusted statistic was statistically significant ($F(1,88) = 6.786$ $p < .0125$), with an odds ratio of 2.132. Next, the following teacher-level demographic covariates were added to the models: male gender, non-white ethnicity, being a special education teacher, alternative certification, having a mentor, and rating of teacher preparation experience. None of these covariates, including having a mentor, were significantly related to turnover status in any of the waves, suggesting that risk for stress is associated with turnover regardless of teacher and school characteristics. Again, however, a significant limitation

of this study is the single item on mentoring, which does not capture the varied content or characteristics of mentoring programs.

The Current Study

Purpose

The primary purpose of the current study was to empirically evaluate the effects of school-based mentoring programs on first year teachers' risk for occupational stress, while addressing some of the common limitations found in research relating to teacher mentoring and occupational stress. In particular, the study aimed to address the following limitations that are widespread in the research base: (a) small sample size; (b) use of local samples rather than nationally-representative samples (c) dearth of research on beginning teachers, particularly those in their first year; (d) lack of a comparison group (i.e., mentored vs not mentored); (e) absence of, or insufficient type/number of control variables; (f) tendency to evaluate stress from frameworks other than the transactional model. I addressed these gaps through the use of nationally representative, cross-sectional data collected from new teachers, using an empirically supported methodology (i.e., the CARD classification and scoring system) for measuring teacher' s classroom appraisals and operationalizing risk for stress. The data source was the 2015-16 National Teacher and Principal Survey (NTPS), which measures teachers' perceptions of their working conditions and provides demographic data on teachers and schools.

A secondary purpose of the current study was to investigate the relationship between teacher stress and several indicators of occupational health, namely Autonomy, Workplace Fatigue, and Job Satisfaction. Previous research has found a connection between teachers' risk for stress and several indicators of occupational health, including burnout, job satisfaction, classroom autonomy, and professional commitment (Lambert et al., 2015; McCarthy et al., 2009; Veldman et al., 2016). What has not been less thoroughly

investigated is the link between occupational health variables and beginning teacher stress risk, particularly from a transactional perspective. A recent study by Fitchett and colleagues (2018) shed some light onto this research gap, finding that first-year teachers classified at risk for stress (according to the CARD scoring system) reported greater symptoms of burnout (*d* = 1.48) and less classroom control (*d* = .62) than teachers with low enough Appraisal Index scores to be classified as not at risk for stress. Seeing as this study analyzed the 2007-08 Schools and Staffing Survey, it is important to take a second look at occupational health in relation to first-year teacher stress using the more recent 2015-16 NTPS. For the present study, I investigated the link between Appraisal group classification and three key occupational health variables: Workplace Fatigue, Job Satisfaction, and Autonomy.

Research Questions:

The current study used the 2015-16 NTPS dataset to answer the following research questions (RQs):

Research Question 1: Are first-year teachers assigned a mentor less likely to be at risk for stress (i.e., report lower Appraisal Index scores) compared to teachers who are not assigned a mentor, controlling for study covariates?

Research Question 2: Are teachers assigned a mentor less likely to be classified as Demanded (i.e., most at risk for stress) and more likely to be classified as Balanced or Resourced (i.e., least at risk for stress) compared to teachers not assigned a mentor?

Research Question 3: Are first-year teachers assigned a mentor more likely than non-mentored teachers to be classified as Demanded, after controlling for study covariates?

Research Question 4: When assigned mentors, are Demanded teachers reporting different mentoring experiences compared to Resourced and Balanced teachers?

Research Question 5: Are first-year teachers classified as Demanded more likely to report lower Autonomy, greater Workplace Fatigue, and lower Job Satisfaction compared to Resourced and Balanced teachers?

Method

Data and Sample

National Teacher and Principal Survey (NTPS)

The data source for the current study was the 2015-16 National Teacher and Principal Survey (NTPS). The NTPS is a nationally representative survey of public teachers, schools, and administrators from grades K-12 in the 50 U.S. states and the District of Columbia. Conducted by the National Center for Education Statistics (NCES), the NTPS is the largest and most comprehensive survey of the U.S. teacher workforce, and includes cross-sectional survey data on teacher perspectives and working conditions, as well as teacher-level demographics and school-level contextual information. The NTPS produces national estimates for public elementary and secondary teachers, including national estimates according to subject matter taught and by full-time or part-time status (Taie & Goldring, 2017). The NTPS is a redesigned version of the Schools and Staffing Survey (SASS), which was administered every four years and spanned from the 1987-88 school year through 2011-12. While the NTPS has a different sample and differs structurally from the SASS, the content of the 2015-16 NTPS is largely similar to that of the 2011-12 SASS, with many items remaining the same (Taie & Goldring, 2017). As such, there is a high level of continuity between the SASS and NTPS, encouraging the development of studies that replicate and/or extend existing SASS research.

NTPS Missing Data Procedures. All NTPS data is fully imputed. The NCES used two imputation approaches to handle missing data in the NTPS dataset: the donor respondent and central tendencies methods. In the event that a donor case could not be matched, the mean or mode from groups of similar cases was used. All imputed data was

then assessed for consistency with the questionnaire; in the event of an inconsistency, Census Bureau analysts determined an appropriate value (Taie & Goldring, 2017).

NTPS Sampling. NTPS employs a complex, multistage sampling design which includes a systematic, probability proportionate to size (PPS) sampling strategy (Raj, 1964). The primary sampling unit for NTPS is the school. The NTPS universe of schools is restricted to the 50 U.S. states plus the District of Columbia, with the following school types excluded from the sample: Department of Defense overseas schools, overseas U.S. territory schools, and schools that do not offer classroom instruction in grades 1-12 or ungraded equivalents. Some schools were purposely oversampled based on characteristics including urbanicity (i.e., rural, urban, town, or suburban area; please reference https://nces.ed/gov/surveys/ruraled/definitions.asp for school locale definitions), charter school status, and grade level, and sample sizes were inflated for schools in the six states containing the fewest schools. This resulted in a sample of 7,130 traditional public schools (i.e., not public charter or typical charter). Teachers were first randomly sampled from lists that were either provided by sampled schools, obtained from school websites, or supplied from a vendor. Next, teachers were stratified by subject area and then selected systematically with equal probability from each stratum, with teachers sampled in an ongoing process as the roster information became available to the NCES. A maximum of 20 teachers were sampled from each school, with an average of 6 to 8 sampled per school. The final sample included 31,540 full-time public school teachers (Riddles, Wallace, Rizzo, & Marker, 2017). In August of 2015, questionnaires were sent to school administrators for distribution to teachers; repeated and ongoing follow-up operations were then conducted from October 2015 through May 2016. A total of 24,570

full-time public school teachers completed the survey (Riddles et al., 2017). The resulting teacher-level data is available through the 2015-16 NTPS restricted-use Public School Teacher data file, and is capable of producing national estimates of teachers by subject matter taught and by full- or part-time teaching status.

Sample for the Current Study

Participants in the current study were all full-time traditional public school teachers who were in their first year of teaching (specifically, who entered the workforce in either 2014 or 2015) and responded to the 2015-16 NCES National Teacher and Principal Survey (NTPS). Teachers who were not full-time public school teachers were excluded because including less common teacher types (e.g., part-time, itinerant, charter, public charter) teachers might bias estimates. Teachers meeting the aforementioned inclusion criteria were included in the sample regardless of whether or not they reported having a mentor, as this permitted group comparisons (i.e., teachers who do receive mentoring versus those who do not, thus addressing a common limitation of mentoring research). With the application of replicate weights and final sampling weights, the NTPS provides a nationally-representative sample of first-year traditional public school teachers for the 2015-16 school year. Normalized final weights were calculated and used to report percentages and frequencies, except where otherwise noted. Demographic information for the full ($n_{unweighted}$ = 1,250) and mentored ($n_{unweighted}$ = 1,000) samples is provided in Table 1.

Table 1

Demographic Characteristics of the Sample

Demographic Variable	Weighted %		Unweighted %[a]	
	Full Sample (*N* = 143,530)	Mentored Sample (*n* = 118,255)	Full Sample (*N*=1,250)	Mentored Sample (*n*=1,000)
Gender				
Male	23.6%	23.8%	30%	30%
Ethnicity				
Hispanic/Latino	12.2%	11.3%	10%	10%
Race				
White	87.3%	87.1%	90%	90%
African American/Black	10.2%	10.4%	10%	10%
Asian	3.8%	3.3%	0%	0%
Native Hawaiian/Pacific Islander	0.6%	0.7%	0%	0%
American Indian/Alaska Native	1.8%	1.8%	0%	0%
Year began Teaching				
2014	26.2%	24.7%	30.0%	30%
2015	73.8%	75.3%	70.0%	70%
Area of Focus				
Alternatively Certified	26.9%	26.2%	30%	30%
Special Education	14.1%	15.2%	10%	10%
ESL/Bilingual	2.5%	2.3%	0%	0%
Urbanicity of School				
City	34.3%	32.8%	30%	30%
Suburb	34.8%	35.2%	30%	30%
Town	12.0%	12.5%	20%	20%
Rural	18.9%	19.5%	20%	20%
Percent IEP (mean)	16.9%	17.1%	20%	20%
Percent LEP/ELL (mean)	11.7%	11.0%	10%	10%
School Level				
Secondary	32.6%	32.9%	40%	40%
Elementary	67.4%	67.1%	60%	60%

Note. Racial categories may exceed 100% due to participants selecting more than one identifier.

[a]Unweighted percentages and sample sizes rounded to the nearest 10 per NCES disclosure procedures.

Strategies Used to Address the Complex Sampling Design of the NTPS. Due to the complex sampling design of the NTPS, adjustments must be made in order to produce nationally-representative estimates and ensure the accuracy of hypothesis tests. To aid in this process, the NCES provides final sampling weights and replicate weights for each teacher in the sample; these weights were applied to all analyses.

Sampling weights adjusted the samples to reflect the true demographic composition of the U.S. national teacher workforce. The NCES provides a final sampling weight for each teacher (variable name: TFNLWGT), which adjusts for differential rates of non-response among subgroups and accounts for the unequal probability of selection.

Replicate weights were used to calculate appropriate standard errors of estimates. As previously noted, the NTPS employs a complex sampling design; however, statistical packages (including Stata, used in the current study) automatically assume simple random sampling has been used. If uncorrected, estimates of sampling errors will often be underestimated. To address this issue, NTPS provides a set of 200 teacher replicate weights (variable names: TREPWT1 - TREPWT200) that are associated with each respondent's final sampling weight. NTPS replicate weights use the jackknife resampling technique for variance estimation. This procedure involves repeatedly dropping a small portion of the sample to create a number of subsamples (replicates). Using revised weights for each replicate sample, weighted estimates of the statistic of interest (the jackknife replicates) are then computed. Finally, the variability between these replicate statistics is used to obtain an unbiased estimate of the variance of the full sample statistic.

The Stata software package was selected for use in the current analyses due to its ability to incorporate weights in the calculation of variance and parameter estimates.

As recommended by the NCES, I accounted for the complex survey design by using Stata's "svyset" command, which requires entering the main sampling weight (TFNLWGT), the full set of replicate weights (TREPWT1 - TREPWT200), and the method of variance estimation (JACKKNIFE).

Procedures

Operationalizing Risk for Stress

The NTPS includes workplace climate items that are similar to CARD items assessing particular classroom demands and resources, making this questionnaire ideal for studying teacher stress within the framework of Lazarus and Folkman's (1984) Transactional Model of stress. Mirroring the procedures used in recent early-career teacher SASS research (e.g., Fitchett et al., 2018; Lambert, Fitchett, McCarthy, & Boyle, 2016; McCarthy et al., 2019, 2016) I used teacher responses to NTPS items assessing workplace appraisals (see Appendix C for the item list) to create a continuous measure of stress risk (Appraisal Index score) and to group teachers according to their respective stress-vulnerability (Resourced, Balanced, and Demanded stress groups).

Using items from the NTPS, I applied the CARD scoring system (Lambert et al., 2015) to appraise teachers' levels of perceived classroom Demands and Resources, as well as classify teachers according to stress risk. First, to measure teachers' risk for occupational stress, I created a continuous variable from NTPS items assessing teachers' perceptions of workplace demands and resources. Guided by the initial CARD scoring strategy (Lambert et al., 2015), I followed the procedures used in a similar study of first-year teacher stress (Fitchett et al., 2018) The Appraisal Index score measures risk for stress using a difference score (Demands minus Resources), thereby operationalizing a

key tenet of transactional theory, which is that risk for stress is a function of the appraised imbalance of perceived workplace demands vis-à-vis resources (Cox, 1978; Lazarus & Folkman, 1984).

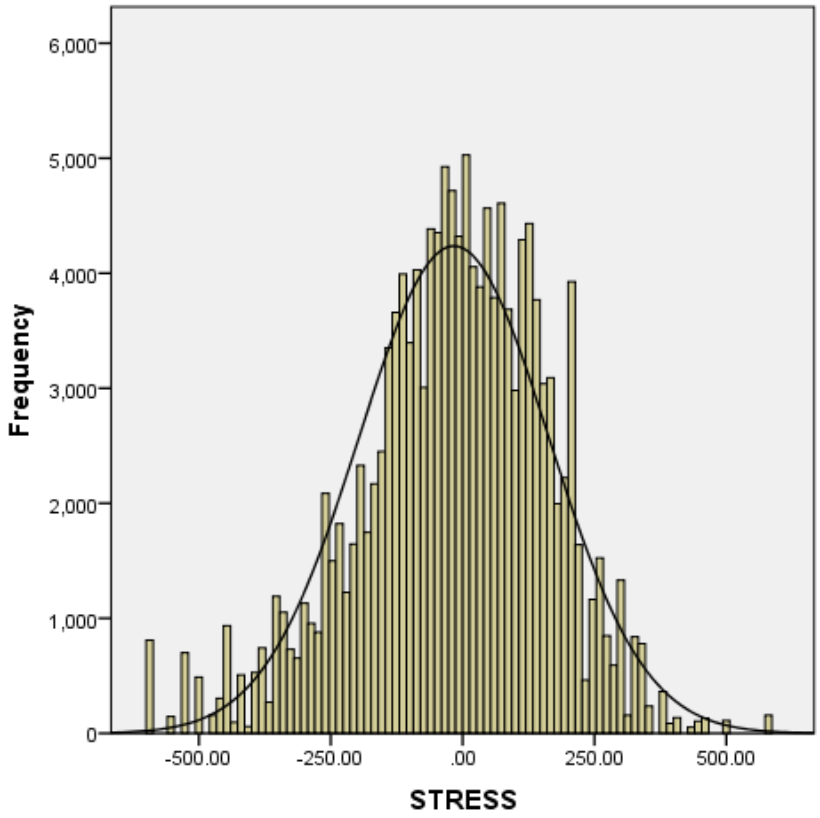
The first step in creating the stress risk variable (i.e., Appraisal Index) was constructing a Demands scale and a Resources scale using pre-identified items from the NTPS that are thematically, theoretically, and conceptually similar to CARD items. The Resources scale included self-reported Likert-type items on teacher perceptions of principal leadership, access to teaching resources, adherence to school policies, and communication among school staff. The Demands scale included self-reported Likert-type items measuring school- and classroom-level rates of student misbehavior and teacher absences. Appendix C provides a list of NTPS items selected for use in constructing the Resources and Demands scales. Next, the Rasch rating scale model, a case of one parameter item response theory (IRT), was applied with the WINSTEPS software package to create interval scale scores for each teacher from both the Demands and Resources scales (Lambert et al., 2017). Demands and Resources scores were scaled to have a mean of 500 and a standard deviation of 100. Because work environments for elementary and secondary teachers are hypothesized to differ with regard to organizational structure and workplace culture, I followed the lead of earlier studies (e.g., Fitchett et al., 2018) and constructed these interval scale scores separately for the two teacher grade levels. When these steps were followed with a sample of full-time public school teachers who responded to the 2015-16 NTPS (i.e., not limited to first-year teachers), the resulting Classroom Demands ($\alpha_{elementary} = .876$, $\alpha_{secondary} = .898$) and

Resources ($\alpha_{elementary}$ = .852, $\alpha_{secondary}$= .853) scales yielded scores with adequate internal consistency reliability (Fitchett et al., 2020).

At this stage, I calculated the Appraisal Index score for each teacher by calculating the difference between the Demands and Resources scale scores. An Appraisal Index score of zero represents no difference (e.g., demands and resources are perceived as approximately balanced, such that D=R); positive scores indicate that perceived demands outweigh resources (D>R), while negative scores indicate resources outweigh demands (D<R). Figure 1 displays the distribution of Appraisal Index scores for the full sample.

Figure 1

Distribution of Appraisal Index Scores in the Full Sample



[a]Unweighted percentages and sample sizes rounded to the nearest 10 per NCES disclosure procedures.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Teacher and Principal Survey (NTPS), "Public School Teacher Data File," 2015-16.

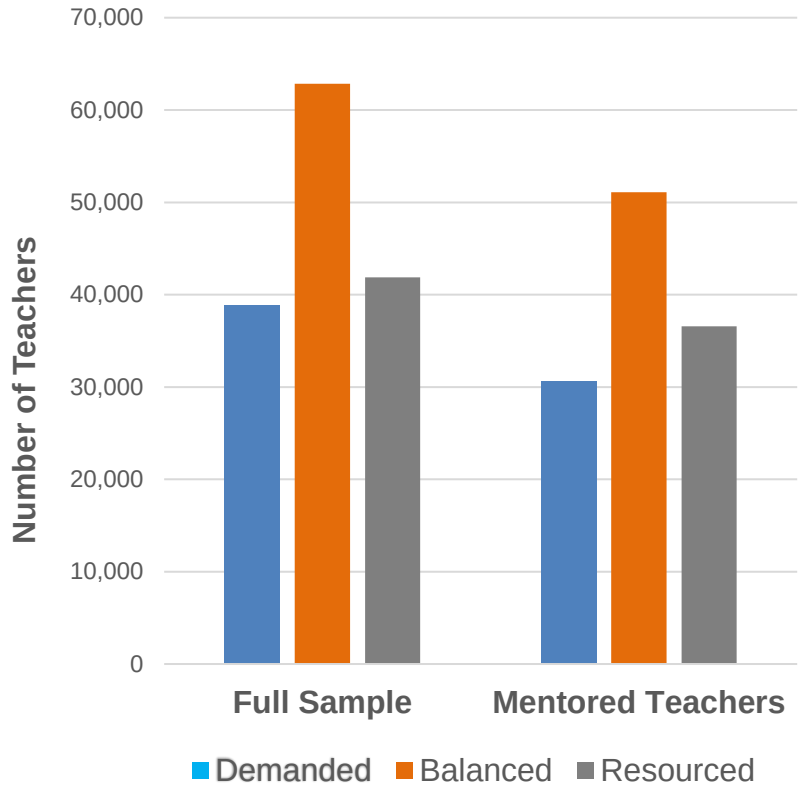
Next, teachers were classified into the Resourced, Balanced, and Demanded subgroups. The primary utility of the stress-risk classification groups is their straightforward interpretation; unlike the scale scores, which can be difficult to translate from the abstract into something more tangible, the implication of being in a particular classification group is relatively simple to comprehend. For example, it is relatively clear that a teacher classified as Demanded is at greater risk for stress than a teacher classified as Resourced. However, the practical meaning of a particular Appraisal Index score is less clear-cut, and therefore potentially less useful to policy makers and school administrators.

To group teachers according to their relative stress risk, a 95% confidence interval was formed around no difference between the Demands and Resources scale scores (McCarthy et al., 2016). Teachers who provided difference scores greater than the upper limit of this interval were classified in the Demanded group (Demands higher than Resources; most at risk for stress), those who provide difference scores below the lower limit of the confidence interval were placed in the Resourced group (Resources higher than Demands; least at risk for stress), and those with difference scores within the interval (Resources approximately equal to Demands) were classified in the Balanced group. Figure 2 shows the number of teachers in each Appraisal group within the full sample and within the mentored sub-sample. In the full sample, 27% of teachers were classified as Demanded, 43.8% as Balanced, and 29.2% as Resourced. In the mentored sample, 25.9% of teachers were classified in the Demands group, 43.2% in the Balanced group, and 30.9% were classified as Resourced. As can be seen, first-year teachers are similarly

distributed across Appraisal groups in the full and mentored samples, with the majority in the Balanced group, followed by Resourced and then Demanded.

Figure 2

Appraisal Group Classification by Subsample



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Teacher and Principal Survey (NTPS), "Public School Teacher Data File," 2015-16.

Analysis

In the first stage of the data analysis, I used SPSS to produce descriptive data on mentor status, occupational health indicators (Workplace Fatigue, Job Satisfaction and Autonomy), risk for occupational stress, and demographic covariates (see Appendix F for a list of study covariates). Specifically, I controlled for year began teaching, alternative certification, gender, race, ethnicity, special education area of focus, English as a second

language (ESL)/bilingual area of focus, urbanicity (city, suburban, town, rural), school level (elementary vs secondary), percent of the teacher's students with an Individualized Education Plan (IEP), and percent of the teacher's students designated as Limited English Proficient (LEP)/ English Language Learner (ELL)]. The covariates were included in the analysis, which might account for differences in the outcome variables as well as any apparent relationships among teacher stress risk, mentoring (both mentor status and types of supports received), and occupational health. There are a host of studies identifying a number of teacher and workplace characteristics related to the prediction of the outcome variables. For example, Kapadia et al. (2007) found that teacher background characteristics and type of teacher preparation correlated with beginning teachers' classroom experiences. Novice male teachers, for example, were more likely to report a positive workplace experience than female teachers (Kapadia et al., 2007). Research also suggests notable differences in teacher workplace experiences across race and ethnicity (Achinstein, Ogawa, & Freitas, 2010; Griffin & Tackie, 2016), including risk for stress (Fitchett, Dillard, McCarthy, Lambert, & Mosley, 2020). Teachers in the sample were classified by NCES as first-year teachers new to the field, but there was some variation in which year the teacher reported entering the field Because the NCES sampled beginning teachers at two different time points, year began teaching is included to control for any potential variation in outcomes that could arise from beginning teaching in 2014 versus 2015. Several important school-level factors that have known links to teachers' workplace experience are urbanicity and socio-economic status (SES) of the teacher's school, as well as the percent of students with an IEP or identified as LEP. Urbanicity classifies a school's location as being rural, town, suburban, or urban. Teachers in poor,

urban schools are more likely to encounter additional stressors (e.g., heavier workloads, lower funding, and increased student discipline problems) than teachers who teach in rural, town, or suburban locales (Thompson, 2017; Ryan et al., 2017). Therefore, urbanicity was also included as a workplace-related covariate. Lambert and colleagues (2018) found that classroom percentages of students with IEPs and with LEP were positively associated with the probability of teachers' Demands group membership.

To provide an overall picture of how first-year teachers appraised their workplace during the 2015-16 school year, I calculated estimates of the proportion of teachers classified as Demanded, Balanced, or Resourced.

Prior to running any descriptive statistics, I used SPSS to determine whether multilevel modeling would be appropriate, I used the "AGGREGATE" and "N_BREAK" commands to determine how many teachers were clustered within each school. Seeing as no school cluster contained more than four teachers (M =1.37, SD = 0.638), I followed the guidance of McCarthy et al. (2017) and determined that a single-level model was appropriate for all analyses.

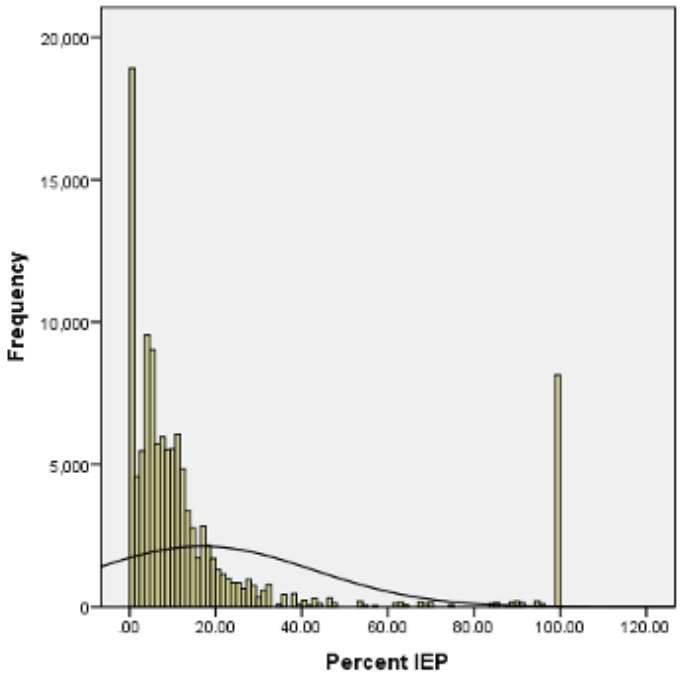
Next, I re-coded categorical independent variables into dummy variables, and set valid skips as "system missing" for continuous covariates (%IEP and %LEP). Please reference the variable key (Appendix B) for a full list of the original NTPS items and corresponding recoded study variables. The normally-distributed Autonomy variable was then created by first summing the scaled scores on School Influence, Pedagogical Influence, and Classroom Control, and then transforming into a score ("AUTO_SS") scaled to have a mean of 500 and a standard deviation of 100. The Cronbach's alpha for the thirteen Autonomy items was 0.857, indicating a high level of shared covariance and

thus supported combining the School Influence, Pedagogical Influence, and Classroom Control items into one scale.

For the next step, I applied sampling weights to convey an accurate picture of the population. Beginning with the scale covariates – percent of the teacher's students with an IEP ("IEP") and percent of the teacher's students designated as LEP/ELL ("LEP") – I calculated covariate means, distributions, and standard deviations. Figure 3 presents descriptive information for %IEP (figure 3a) and %LEP/ELL (figure 3b).

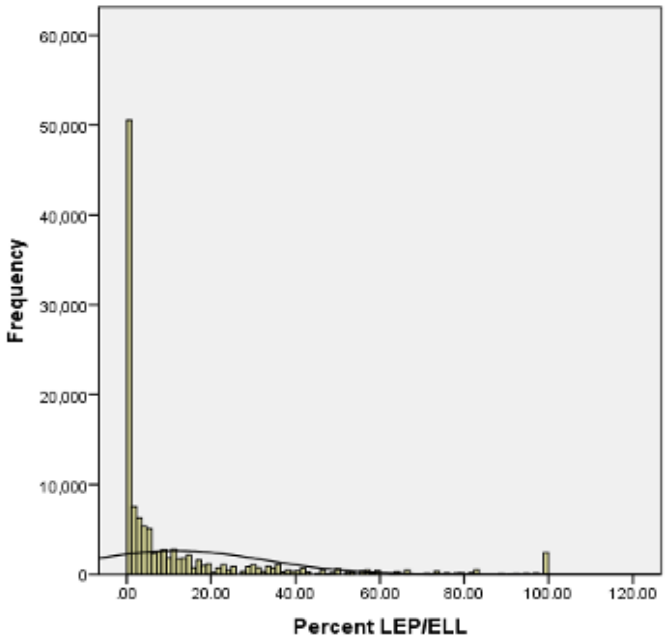
Figure 3

3a. Weighted Distribution of % of Teacher's Students with an IEP



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Teacher and Principal Survey (NTPS), "Public School Teacher Data File," 2015-16.

3b. Distribution of % of Teacher's Students who are ELL/LEP



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Teacher and Principal Survey (NTPS), "Public School Teacher Data File," 2015-16.

I then transitioned to using Stata to address the main research questions. As previously noted, mentoring programs can vary greatly in their content and characteristics. Thus, I used mentoring-specific items from the 2015-16 NTPS (see Appendix D for the list of mentoring support-related NTPS items) to provide information throughout my analyses on mentoring prevalence and on types of mentoring supports received. Appendix E contains all NTPS items used to create the Autonomy, Workplace Fatigue, and Job Satisfaction variables.

To address Research Question 1, I conducted a multiple regression of mentor status on the Appraisal Index with a set of control variables. Covariate adjustment was used to correct for selection bias, covariate imbalance, and potential confounding; this method has been shown to produce reliable and efficient estimates that are as robust as

those produced by propensity score matching (Elze et al., 2017). Following previous research on teacher stress (e.g., Fitchett et al., 2020; Fitchett et al., 2018; McCarthy et al., 2019), I included the following control variables for teacher characteristics: year began teaching, alternative certification, gender, race, ethnicity, special education area of focus, ESL/bilingual area of focus, urbanicity (city, suburban, town, rural), school level (elementary vs secondary), % of the teacher's students with an IEP, and % of the teacher's students designated as LEP/ELL].

For Research Question 2, I used a chi-square test of independence to compare the mentor status of teachers in each of the Appraisal group classifications. Sampling weights and replicate weights were applied for all analyses completed in Stata, and the "svy" command was used to indicate the survey design of the data. The weighted estimate of first-year teachers who participated in a mentoring program was 82.4%.

For Research Question 3, I conducted a binomial logistic regression to explore the link between mentoring programs and first-year teachers' risk for occupational stress. Specifically, I determined whether the likelihood of being classified as Demanded (versus Resourced or Balanced) is related to mentoring participation, while controlling for teacher characteristics. To properly carry out the logistic regression, I collapsed the dependent variable, appraisal group classification, into two categories: Demanded (teachers classified as Demanded), and Resourced/Balanced (teachers classified as either Resourced or Balanced). Postestimation contrasts and pairwise comparisons were run on Urbanicity due to it being a multi-category variable.

A subset of the population, first-year teachers who reported working with a mentor (n = 1,000) was analyzed for Research Question 4. Four separate chi-squares, one for

each type of mentoring support -- help with paperwork or record keeping, lesson demonstration, help with preparing lessons that address student learning standards, and assistance with developing student assessment tools -- were run to calculate the weighted percent of mentored teachers, by Appraisal group classification, who received each type of support.

The final step of the analysis, addressing RQ5, required three regression models, one for each occupational health DV (Autonomy, Workplace Fatigue, and Job Satisfaction). I included the full set of the aforementioned covariates. Post-estimation contrasts and pairwise comparisons were run for both multi-category IVs (Appraisal group and Urbanicity).

Results

Research Question 1: First-year teachers assigned a mentor are less likely to be at risk for stress (i.e., report lower Appraisal Index scores) compared to teachers who are not assigned a mentor, controlling for study covariates.

A multiple regression was conducted to investigate RQ1, with the following variables regressed onto the Appraisal Index: mentor status, year began teaching, alternative certification, sex, race, ethnicity, special education status, ESL status, urbanicity, school level, percent of the teacher's students with an IEP, and percent of the teacher's students who were ELLs/LEPs. Results are presented in Table 2. Together, these variables statistically significantly predicted Appraisal Index score [$F(14, 190) = 7.10$, $p < .001$, $R^2 = 0.105$]. As previously described, increases in Appraisal Index score (i.e., higher positive scores reflect higher risk for stress) theoretically correspond with increased risk for occupational stress. The following variables added statistically significantly to the prediction of Appraisal Index score, at $p < .05$: mentor status ($t = -2.62$, $p = .01$), ethnicity ($t = -3.21$, $p < .01$), special education ($t = -2.79$, $p < .01$), urbanicity [$F(3, 200) = 10.29$, $p < .001$), and LEP/ELL percentage ($t = 4.59$, $p < .001$). These findings indicate that having a mentor, identifying as Hispanic, and teaching special education are associated with lower risk for stress, while higher LEP/ELL percentages are associated with higher risk for stress. The standard error of the mentoring coefficient, 14.322, shows that having a mentor was associated with over 1/3 of a standard deviation decrease in the Appraisal Index Score. Pairwise comparisons of means on urbanicity (URBANS12) using the Bonferroni adjustment indicated that three comparisons were significant: (a) suburban teachers ($M = -93.535$*)* had significantly ($p < .001$) lower mean Appraisal Index

scores than teachers in city schools (*M* = -12.182), with a 95% confidence interval of the difference between means from -125.004 to -37.703; (b) rural teachers (*M* = -79.587) had significantly (*p* = .001) lower mean Appraisal Index scores than teachers in city schools (*M* = -12.182), with a 95% confidence interval in the difference in means from -115.185 to -19.625; and (c) teachers in town schools (*M* = -34.401) had significantly (*p* < .01) higher mean Appraisal Index scores than suburban teachers (*M* = -93.535), with a 95% confidence interval in difference between means of 15.409 to 102.859.

Table 2

Predicting Appraisal Index score from Mentor Status and Covariates

Predictor	*b*	SE	*t*	*p*	95% CI
Has mentor	**-37.462****	14.322	-2.616	.010	[-65.705, -9.219]
Began teaching in 2014	-15.175	12.328	-1.231	.220	[-39.485, 9.134]
Alternatively certified	14.922	15.242	0.979	.329	[-15.134, 44.978]
Male	15.334	11.884	1.290	.198	[-8.101, 38.769]
White	12.987	19.684	0.660	.510	[-25.829, 51.803]
Hispanic	**-65.372****	20.390	-3.206	.002	[-105.581, -25.163]
Special education	**-89.806****	32.225	-2.787	.006	[-153.352, -26.259]
ESL	36.871	64.194	0.574	.566	[-89.717, 163.460]
Urbanicity					
Suburban	**-81.353*****	16.380	-4.967	.000	[-113.653, -49.053]
Town	-22.219	16.181	-1.373	.171	[-54.127, 9.688]
Rural	**-67.405*****	17.929	-3.759	.000	[-102.761, -32.049]
Elementary	15.747	12.928	1.218	.225	[-9.746, 41.240]
% IEP	0.445	0.339	1.312	.191	[-0.224, 1.114]
% LEP/ELL	**1.471*****	0.312	4.588	.000	[0.839, 2.103]

Note. * p<.05, **p<.01,*** p<.001.

Research Question 2: First-year teachers assigned a mentor are less likely to be classified as Demanded (i.e., most at risk for stress) and more likely to be classified as Balanced or Resourced (i.e., least at risk for stress) compared to teachers without a mentor.

For RQ2, a chi-square test of independence was used to compare mentor status – i.e., whether or not a teacher was assigned a mentor – of first-year teachers between

stress group classifications. Results are displayed in Table 3. There was a significant relationship between not having a mentor and stress group classification [χ^2 (2, N = 1,250) = 9.337, p < .05]. As anticipated, the proportion of Demanded teachers without a mentor (0.211) was greater than the proportion of Balanced (0.187) and Resourced (0.127) teachers. In other words, first-year teachers classified in the Demands group were approximately 8% less likely than Resourced or Balanced teachers to have an assigned mentor.

Table 3

Proportion of Teachers in Each Appraisal Group, by Mentor Status

Mentor Status	Appraisal Group		
	Resourced	Balanced	Demanded
Mentor	.873	.813	.789
No Mentor	.127	.188	.211

Mentor Status	Appraisal Group		
	Resourced	Balanced	Demanded
Mentor	.873	.813	.789
No Mentor	.127	.188	.211

Mentor Status	Appraisal Group		
	Resourced	Balanced	Demanded
Mentor	.873	.813	.789
No Mentor	.127	.188	.211

Research Question 3: Mentor status was not a significant predictor of self-reporting into the Demanded category. Ethnicity, special education status, urbanicity, and percent ELL/LEP were all significant predictors of being classified as Demanded, after controlling for study covariates.

For RQ3, a binomial logistic regression was run to investigate the effects of mentor status and study covariates (year began teaching, alternative certification, sex, race, ethnicity, area of specialization [special education, ESL], urbanicity [city, town, rural], school level taught [elementary, secondary] IEP percentage, and LEP/ELL percentage) on the likelihood of first-year teachers being classified as Demanded versus either Resourced or Balanced. The omnibus test indicated that the logistic regression model was significant [$F(14, 190) = 4.48$, $p < .001$]. Table 4 reports the results of the binomial regression. Four independent variables were significant predictors of belonging to the Demands group at the $\alpha \leq .05$ level: Hispanic ($t = -1.986$, $p = .048$), special education ($t = -2.726$, $p = .007$), urbanicity [$F(3, 200) = 8.36$, $p < .001$], and LEP/ELL percentage ($t = 2.648$, $p = .009$). The odds ratio of Hispanic (0.582) indicates that a teacher who identifies as Hispanic is 1.72 times less likely to be in the Demands group than a teacher who does not identify as Hispanic. Looking at a teacher's area of specialization, we see that a special education teacher is 4.00 times less likely than a non-special education teacher to be in the Demanded group (OR = 0.250). Post-estimation comparisons between each level of urbanicity revealed one significant comparison between suburban vs city teachers, with suburban teachers being 2.67 times less likely than teachers in city schools to be classified as Demanded (OR = 0.374, $t = -4.98$, $p < .001$). Finally, for every one unit

increase in the percentage of LEP/ELL learners in a teacher's classroom, the odds of this teacher belonging to the Demands group increases by a factor of 1.011.

Table 4

Predicting Demands Group Membership

Predictor	Odds Ratio (SE)	*t*	*p*	95% CI for Odds Ratio
Has mentor	0.819	-0.937	.350	[0.538, 1.247]
	(0.174)			
Began 2014	0.843	-1.018	.310	[0.605, 1.174]
	(0.142)			
Alternatively Certified	1.319	1.521	.130	[0.921, 1.887]
	(0.240)			
Male	0.948	-0.322	.748	[0.684, 1.314]
	(.157)			
White	1.534	1.779	.077	[0.955, 2.464]
	(0.369)			
Hispanic	**0.582***	-1.986	.048	[0.341, 0.996]
	(0.158)			
Special Education	**0.247****	-2.726	.007	[0.090, 0.679]
	(0.127)			
ESL	2.863	1.374	.171	[0.633, 12.945]
	(2.191)			
Urbanicity				
Suburban vs City	**0.374*****	-4.98	.000	[0.221, 0.633]
	(0.074)			

Table 4 (Continued)

Town vs City	0.538 (0.127)	-2.64	.054	[0.287, 1.007]
Rural vs City	0.588 (0.132)	-2.36	.116	[0.323, 1.071]
Town vs Suburb	1.437 (0.342)	1.52	.778	[0.762, 2.712]
Rural vs Suburb	1.573 (0.344)	2.07	.239	[0.878, 2.818]
Rural vs Town	1.094 (0.272)	0.36	1.000	[0.565, 2.120]
IEP percentage	1.005 (0.005)	1.014	.312	[0.995, 1.014]
LEP/ELL percentage	**1.011**** (0.004)	2.648	.009	[1.003, 1.019]

Note. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Research Question 4: When assigned mentors, Demanded teachers are reporting different mentoring experiences compared to Resourced and Balanced teachers.

For RQ4, four separate chi-square tests of independence were run to investigate the relationship between appraisal group classification (Resourced, Balanced, Demanded) and the types of supports mentors provided to their assigned teachers (help with paperwork or record keeping, lesson demonstration, assistance with preparing lessons that met student learning standards, help with developing student assessment

tools). The relation between appraisal group classification and mentoring support was significant in all four cases. Results are displayed in Table 5.

For the first chi-square test, a statistically significant relationship was found between stress group classification and whether a teacher received their mentor's support with carrying out paperwork or record keeping duties [χ^2 (2, N = 1,000) = 30.175, p < .001]. Among teachers with an assigned mentor, the proportion of teachers in the Demands group who received a mentor's help with paperwork or record keeping (0.530) was less than the proportion of teachers in both the Balanced (0.615) and Resourced (0.749) groups who received this support.

Similarly, a statistically significant relationship was found between stress group classification and whether a teacher received lesson development support from their mentor [χ^2 (2, N = 1,000) = 32.911, p < .001]. Among teachers with an assigned mentor, the proportion of teachers in the Demands group whose mentors demonstrated lessons for them (0.426) was less than the proportion of teachers in both the Balanced (0.531) and Resourced (0.665) groups who received this support.

The third chi-square test revealed a statistically significant relationship between stress group classification and whether a teacher received lesson preparation support from their mentor [χ^2 (2, N = 1,000) = 31.343, p < .001]. Among teachers with an assigned mentor, the proportion of teachers in the Demands group whose mentors helped them prepare lessons that addressed student learning standards (0.504) was less than the proportion of teachers in both the Balanced (0.640) and Resourced (0.731) groups who reported receiving this form of mentoring support.

Finally, a statistically significant relationship was revealed between stress group classification and whether a teacher's mentor assisted them with the development of student assessment tools [χ^2 (2, N = 1,000) = 45.901, p < .001]. Among teachers with an assigned mentor, the proportion of teachers in the Demands group whose mentors helped them develop student assessment tools (0.515) is less than the proportion of teachers in both the Balanced (0.668) and Resourced (0.784) groups who reported this type of mentoring.

In each set of analyses, the Resourced group emerged as the stress classification group with the greatest proportion of teachers receiving the mentoring support, followed by Balanced and Demanded teachers. This aligns with our study hypothesis. Also in line with expectations, across all types of mentoring supports, there were fewer observations for the Demanded group versus the Resourced or Balanced groups.

Table 5

Results of Chi-square Tests for Mentoring Support Status by Appraisal Group

Mentoring Support	Appraisal Group		
	Demanded	Balanced	Resourced
Paperwork			
Yes	.530	.615	.749
No	.470	.385	.251
Lesson Demonstration			
Yes	.426	.531	.665
No	.574	.469	.335
Student Learning Standards			
Yes	.504	.640	.731
No	.496	.360	.269
Student Assessment			
Yes	.515	.668	.784
No	.485	.332	.216

Research Question 5: First-year teachers classified in the Demands group are more likely to report lower Autonomy, greater Workplace Fatigue, and lower Job Satisfaction compared to Resourced & Balanced teachers, controlling for study covariates.

Three separate linear regression analyses were conducted to evaluate RQ5. Stress group classification was regressed onto each of the following occupational health variables, controlling for study covariates: Autonomy, Workplace Fatigue, and Job Satisfaction. Results are presented in Table 6. Together, the covariates and stress group

classification variable statistically significantly predicted Autonomy [$F(16, 180) = 13.93$, $p < .001$, $R^2 = 0.205$]. The following variables statistically significantly predicted Autonomy, at an alpha of $p < .05$: appraisal group [$F(2, 200) = 85.98$, $p < .001$], mentor status ($t = -2.933$, $p < .01$), alternative certification ($t = -2.391$, $p < .05$), sex ($t = 2.569$, $p < .05$), school level ($t = -2.933$, $p < .01$). For stress group classification, pairwise comparisons of the means using the Bonferroni adjustment indicated all three comparisons were significant: (a) teachers in the Balanced group ($M = 0.072$) reported significantly ($p < .001$) lower Autonomy scores than teachers in the Resourced group ($M = 0.632$), with a 95% confidence interval of the difference between means from -0.721 to -0.398; (b) teachers in the Demands group ($M = -0.415$) reported significantly ($p < .001$) lower Autonomy scores than teachers in the Resourced group ($M = 0.632$), with a 95% confidence interval of the difference between means from -1.25 to -0.845; (c) teachers in the Demands group ($M = -0.415$) reported significantly ($p < .001$) lower Autonomy scores than Balanced ($M = 0.072$) teachers, with a 95% confidence interval of the difference between means from -0.691 to -0.282. As anticipated, the greatest discrepancy in means was seen between the Resourced and Demanded groups.

Results from the second regression showed the covariates and stress group classification variable statistically significantly predicted Workplace Fatigue [$F(16, 180) = 19.03$, $p < .001$, $R^2 = 0.249$]. Appraisal group [$F(2, 200) = 140.63$, $p < .001$] and race ($t = -2.262$, $p < .05$) statistically significantly predicted Workplace Fatigue. Bonferroni-adjusted pairwise comparisons for appraisal group classification were all significant: (a) teachers in the Balanced group ($M = 500.521$) reported significantly ($p < .001$) greater Workplace Fatigue scores than teachers in the Resourced group ($M = 437.707$), with a

95% confidence interval of the difference between means from 47.627 to 78.001; (b) teachers in the Demands group (M = 569.433) reported significantly (p < .001) greater Workplace Fatigue scores than teachers in the Resourced group (M = 437.707), with a 95% confidence interval of the difference between means from 112.316 to 151.136; (c) teachers in the Demands group (M = 569.433) reported significantly (p < .001) greater Workplace Fatigue scores than Balanced (M = 500.521) teachers, with a 95% confidence interval of the difference between means from 49.613 to 88.212. Again, the greatest discrepancy in means was seen between the Resourced and Demands groups.

Finally, the covariates and stress group classification variable statistically significantly predicted teachers' scores on Job Satisfaction [F(16, 180) = 65.31, p < .001, R^2 = 0.444]. As seen with Workplace Fatigue, appraisal group classification [F(2, 200) = 400.13, p < .001] and race (t = 4.333, p < .001) were the only variables to significantly predict Job Satisfaction. Pairwise comparisons of the Job Satisfaction means between appraisal groups using the Bonferroni adjustment indicated all three comparisons were significant: (a) teachers in the Balanced group (M = 487.600) reported significantly (p < .001) lower Job Satisfaction scores than teachers in the Resourced group (M = 567.196), with a 95% confidence interval of the difference between means from -95.472 to -63.722; (b) teachers in the Demands group (M = 387.315) reported significantly (p < .001) lower Job Satisfaction scores than teachers in the Resourced group (M = 567.196), with a 95% confidence interval of the difference between means from -195.197 to -164.567; (c) teachers in the Demands group (M = 387.315) reported significantly (p < .001) lower Job Satisfaction scores than Balanced (M =487.600) teachers, with a 95% confidence interval of the difference between means from -117.130 to -83.440. Again, as with Autonomy and

Workplace Fatigue, the greatest discrepancy in Job Satisfaction means was seen between the Resourced and Demands groups.

Table 6

Teacher Autonomy, Workplace Fatigue, and Job Satisfaction by Appraisal Group

		Resourced Group (R)[a]	Balanced Group (B)[b]	Demands Group (D)[c]	Contrasts	*t*	95% CI
Autonomy	Weighted mean	.632	.072	-.415			
	(SE)	(.116)	(.116)	(.119)	B vs R	**-8.36*****	[-0.721, -0.398]
	Unweighted mean	.517	-.023	-0.51	D vs R	**-12.56*****	[-1.248, -0.845]
	(SE)	(.049)	(.037)		D vs B	**-5.74*****	[-0.691, -0.282]
Workplace Fatigue	Weighted mean	437.707	500.521	569.433			
	(SE)	(11.335)	(11.344)	(12.599)			
	Unweighted mean				B vs R	**9.99*****	[47.627, 78.001]
	(SE)				D vs R	**16.39*****	[112.316, 151.1356]
					D vs B	**8.62*****	[49.613, 88.212]
Job Satisfaction	Weighted mean	567.196	487.600	387.315			
	(SE)	(9.54)	(9.46)	(9.93)			
	Unweighted mean				B vs R	**-12.11*****	[-95.472, -63.722]
	(SE)				D vs R	**-28.36*****	[-195.197, -164.567]
					D vs B	**-14.37*****	[-117.130, -83.440]

Note. Each of the regression analyses includes a host of covariates. Unweighted values are rounded to the nearest tens, as per NCES disclosure protocol.

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

[a] Resourced weighted *n* = 41,895; unweighted *n* = 360

[b] Balanced weighted *n* = 62,847; unweighted *n* = 550

[c] Demands weighted *n* = 38,788; unweighted *n* = 330

Discussion

Overall, results cautiously suggest that formal mentoring programs play an important role in beginning teacher wellness. The following discussion highlights the implications of key study findings.

Mentor Status and First-Year Teacher Risk for Stress

One general question this study aimed to address was whether mentoring programs impact beginning teachers' risk for stress. The results suggest that a first-year teacher's mentor status is indeed associated with their stress risk, as measured continuously using the Appraisal Index. Here, having a mentor is associated with a decreased risk for stress. In the language of appraisals, this means teachers with a mentor appraise their workplace as overall more resourced than demanded.

First-year teachers often find themselves grappling with a host of challenges; in other words, the sources of demands for beginning teachers are numerous. The so-called "painful beginning" (Huberman, 1989) that is a teacher's first year often leaves beginning teachers uncertain about their roles, questioning their abilities, and overwhelmed by responsibilities (DePaul, 2000; Grissom, Kern, & Rodriguez, 2015; Weasmer & Woods, 1998). Formal mentoring programs were developed to aid beginning teachers in the areas of role modeling, psychosocial support, and career support, all in a pointed effort to make the first year of teaching more manageable (Chun et al., 2012). Therefore, one could reasonably expect that teachers with a mentor might perceive themselves to be more resourced and less demanded than their un-mentored peers. Indeed, this is what RQ1 and RQ2 suggest. Whether stress risk is operationalized as an Appraisal Index score or an Appraisal group classification, having a mentor was associated with lower risk for

stress. Interestingly, Fitchett et al. (2018) did not find significant differences in stress risk based on mentoring status. Fitchett et al. (2018) was similarly cross-sectional, and analyzed a precursor to the NTPS, the 2007-2008 SASS. It is unclear why this study, which also used a chi-square test, diverged from the current study's findings. It is possible that there is something inherently different about the current sample of teachers, or perhaps the quality of mentoring programs has improved over time. Additionally, there are slight differences in the questionnaire items used to derive the Demands and Resources items in the two studies, which may relate to the contrasting results. However, the Demands and Resources scales, which always are constructed using the same CARD scoring methodology, consistently capture the same common categories of resources and demands across studies, and thus this is a less compelling explanation than the previous ones offered.

In sum, findings from RQ1 and RQ2 indicate a potentially positive impact of mentoring programs on first-year teachers' stress risk (i.e., mentored teachers had lower Appraisal scores and were more likely to be in the Resourced group). This suggests that mentoring may in fact equip beginning teachers with valuable resources, offsetting at least some aspects of the many demands they face during their first year in the field.

We know that beginning teachers whose workplace appraisals place them at greatest risk for stress are at the highest risk to eventually leave the field or change schools (McCarthy et al., 2019). Turnover often has a devastating impact on teachers, students, and entire school organizations (Ronfeldt et al., 2013). Therefore, formal mentoring programs may be a vital tool for resourcing early-career teachers, and consequently decreasing the likelihood of subsequent turnover.

Exploring the relationship between mentor status and stress risk was approached differently for RQ3, where a binary logistic regression was used, with stress risk operationalized in the form of Appraisal group classifications. Curiously, teachers with a mentor were *not* significantly more likely to be classified as Demanded (versus Balanced or Resourced) than teachers without a mentor when the logistic regression was run. This distinction may be explained by inherent differences in how the analyses were carried out. Compared to RQ2, where stress risk was also measured categorically, RQ3 included a host of covariates in the analysis. This suggests that there is meaningful explanatory power in some of the covariates; in other words, mentor status alone is not driving variation in Demands group membership. Comparing RQ1 to RQ3, we see that the dependent variables (Appraisal Index score and mentor status) are different. When stress risk was assessed using a continuous scale, mentorship emerges as a robust predictor; specifically, having a mentor was associated with over 1/3 of a standard deviation decrease in the Appraisal Score (with lower Appraisal Scores indicating a reduced risk for stress). However, when measured categorically in the logistic regression, mentor status was no longer a significant predictor of stress risk. It is possible that mentorship is important to some extent in ameliorating teachers' stress risk, but whether or not a teacher actually meets the threshold for being classified as Demanded might depend on additional factors. Additional covariates to be considered might include subject area match with one's mentor, percent of students receiving free or reduced lunch (a proxy for school SES), school size, racial/ethnic demographic match between a teacher and their students, income, and age. When considering the potential impact of additional variables, it is important recall that in both

the full and mentored samples, less than 30% of teachers were classified as Demanded (27% in the full sample, and 25.9% in the mentored sample). With such a narrow range of observations falling into the Demanded category, there is only so much variability possible within this small group. Thus, a hefty proportion of the overall variation in Appraisal Group classification is attributable to the factors related to being Resourced or Balanced, making it highly challenging to "move the needle," so-to-speak, into the Demanded category. In conclusion, while mentorship may help reduce teachers' risk for stress to some degree, it may not be sufficient on its own to tip the scales toward a teacher being classified as Demanded.

Mentoring Supports and First-Year Teacher Risk for Stress

One of the most notable strengths of this study was its inclusion of mentoring variables beyond mentor status. Specifically, RQ4 used four separate chi-squares to calculate the weighted percent of mentored teachers, by Appraisal group classification, who received each type of support.

In the current study, mentoring supports – help with paperwork or record keeping, lesson demonstration, assistance with preparing lessons that met student learning standards, help with developing student assessment tools – were associated with Appraisal categories, with the Demands group reporting less access to each type of support. Given the wide variety of formal mentoring programs, policymakers have a vested interest in identifying which specific components of these programs are worth keeping. All four of the support types examined in the current study differ among Resourced, Balanced, and Demanded teachers, the Resourced group consistently emerged as the stress classification group with the greatest proportion of teachers

receiving that mentoring support. Thus, these types of supports may help new teachers cope with the many demands inherent to the first year of teaching.

New teachers are said to often struggle with the gaps between academic theories of teaching and what end up being the practical realities of the job (Gaede, 1978). Thus, receiving practical support from a mentor may play a role in a new teacher appraising their classroom as more resourced than demanded. For example, student learning standards are often discussed in terms of the significant pressure felt by teachers expected to meet such standards (Youngs et al., 2012). Teachers struggling to teach in ways that align with formal learning standards have been shown to suffer in terms of instructional quality, commitment to the field, and overall effectiveness. Thus, receiving a mentor's support with preparing lessons that explicitly address student learning standards likely helps teachers perceive their workplace as more resourced overall. Similarly, a beginning teacher who is able to watch their mentor demonstrate a lesson, or receive help with developing student assessment tools, is theoretically armed with practical know-how that they may not have come into the field with. Receiving a mentor's help with paperwork or record keeping could help offset some of the hefty demands placed on a new teacher's time, given the numerous tasks and responsibilities placed upon them (DePaul, 2000; Grissom, Kern, & Rodriguez, 2015; Weasmer & Woods, 1998).

There are, of course, additional components to many of the mentoring programs currently in place in U.S. schools, such as classroom observation, identification of professional learning goals, conducting formal mentee assessments, and co-teaching a lesson with the mentee. The current study was limited to evaluating the specific types of supports assessed in the NTPS questionnaire. Thus, future research might explore the

usefulness of other types of supports, as well as exploring other measures of efficacy beyond stress-risk classification.

First-Year Teacher Risk for Stress and Occupational Health

Indicators of occupational health, such as burnout (Chang, 2009) reduced job satisfaction, and decreased occupational commitment (Gilbert et al., 2014) have been linked to teachers' risk for occupational stress (Fitchett et al., 2018; Lambert et al., 2015; Veldman et al., 2016). Thus, RQ5 results from the three regression models were in line with previous findings; here, stress group classification significantly predicted Autonomy, Workplace Fatigue, and Job Satisfaction. Consistent patterns emerged across each occupational health variable, with the greatest discrepancy in means emerging between the Resourced and Demanded groups.

These findings bolster existing research suggesting that the transactional model and corresponding CARD classification system are useful tools for examining not only teacher stress, but also in identifying potential occupational health concerns. Given study findings, school administrators armed with the knowledge that a given first-year teacher has been classified as Demanded, and thus at greatest risk for stress, could reasonably expect this teacher to also feel less autonomous in the workplace, experience increased workplace fatigue, and feel less satisfied with their job than their Resourced and Balanced colleagues. The tone set in the beginning phase of a teacher's career is critical in determining their likelihood of remaining in the field (Berman, 2004; Doney, 2013). Berman (2004) notes that "Talented teachers will not last long in a culture that undermines or is neutral to their needs and interests, leaves them isolated, or fails to promote their growth" (p. 118). Thus, school administrators taking a proactive approach to addressing

turnover could selectively funnel resources, including mentoring supports, to first-year teachers classified as Demanded.

Limitations

The current study was a quantitative, cross-sectional investigation of first-year teachers' risk for stress across occupational health indicators, mentor status, and types of mentoring supports received. Results from the study cautiously suggest associations between stress risk and both occupational health and mentoring. While the study is unable to make causal inferences or provide information on the long-term impact of mentoring on teacher stress and occupational health, the findings nevertheless provide a useful platform for future studies that are designed to provide this information.

All findings are the result of self-reported questionnaire data, which can suffer from bias; it is unclear how much an individual's self-report would correlate with their actual behaviors and experiences. However, the nature of transactional stress research is inherently subjective, seeing as it depends on teachers' perceptions of demands and resources. Furthermore, all data was de-identified, likely reducing the chance of teachers' responses being swayed by social desirability bias. Therefore, self-reported data is not seen as a major design flaw in the current study.

Perhaps the most significant limitation of the study was that all data related to mentoring, stress, and occupational health was inherently limited by the questions asked on the NTPS teacher questionnaire. The undeniable benefit of using the NTPS is that it provides a large-scale, nationally representative dataset. However, a major tradeoff is that a researcher is bound by the NCES's wording and item selection; neither of which pay any regard to a researcher's unique research questions. For example, none of the items on the NTPS were specifically designed to measure risk for stress; instead, proxy variables needed to be created to capture this information. Similarly, occupational health

indicators were approximated based on NTPS items that fit the general categories theoretically captured by the Workplace Fatigue, Job Satisfaction, and Autonomy constructs. Thus, it is possible that these proxies do not accurately represent the intended constructs. However, previous research (e.g., McCarthy et al., 2016; McCarthy et al., 2019) suggests that NCES datasets (NTPS, SASS, BTLS) all contain comparable items which, when used to create stress risk and occupational health variables, perform as expected. There were several mentoring-related items on the NTPS that were not included in the current study, but which may have the potential to add useful information. For example, Smith and Ingersoll (2004) found that subject area match between a mentor and mentee was associated with a significant reduction in risk for leaving the field of teaching. It is possible that a similar relationship could exist between subject match and risk for stress and/or occupational health. Additionally, there is an NTPS item written to capture how effective a teacher found their mentoring experience to be. This could be an interesting covariate to include in future studies, given that not every mentor is necessarily a “good” or “helpful” mentor. The quality, amount, and types of support provided by mentors has been described as a “continuum,” (SREB, 2018), ranging from no support to a rich, consistent, person-focused mentoring experience.

Finally, alternative study designs may offer future studies some advantages over the current design. For one, qualitative studies may provide an additional level of depth, as teachers would be capable of providing more specific and nuanced information on what aspects of their mentoring relationship they found most and least helpful, what their mentoring program consisted of beyond help with paperwork, lesson planning, and assessment. Furthermore, based on the findings of this study, I would be most interested

in interviewing teachers landing on both extremes of the appraisal index distribution (i.e., teachers deemed at most risk for stress and at least risk for stress based on their Appraisal Index scores). Rich material could be gathered on the quality and nature of the mentoring relationship, such as whether rapport was developed, how collaborative was the approach to problem solving, et cetera. What additional factors outside of being mentored might have been involved in the more Demanded teachers experiencing their workplace as highly demanding?

Secondly, it could be interesting to explore the long-term effectiveness of mentoring, something that cannot be captured in a cross-sectional study. After all, the purpose of mentoring is arguably to produce lasting benefits for teachers, not just to assist educators in their first year of teaching. Therefore, future studies are recommended to incorporate longitudinal analyses. The SASS, the previous iteration of the NTPS, contained a longitudinal study arm – the Beginning Teacher Longitudinal Study (BTLS) -- allowing for the tracking of first-year teachers over the first five years of their careers. McCarthy and colleagues (2019) used the BTLS to longitudinally examine stress vulnerability and turnover among teachers who started their careers during the 2007-08 schoolyear. While the NTPS does not currently offer a version of the BTLS, any future additions of the sort would provide fertile research opportunities related to the impact of mentoring over the beginning years of a teacher's career.

Conclusion

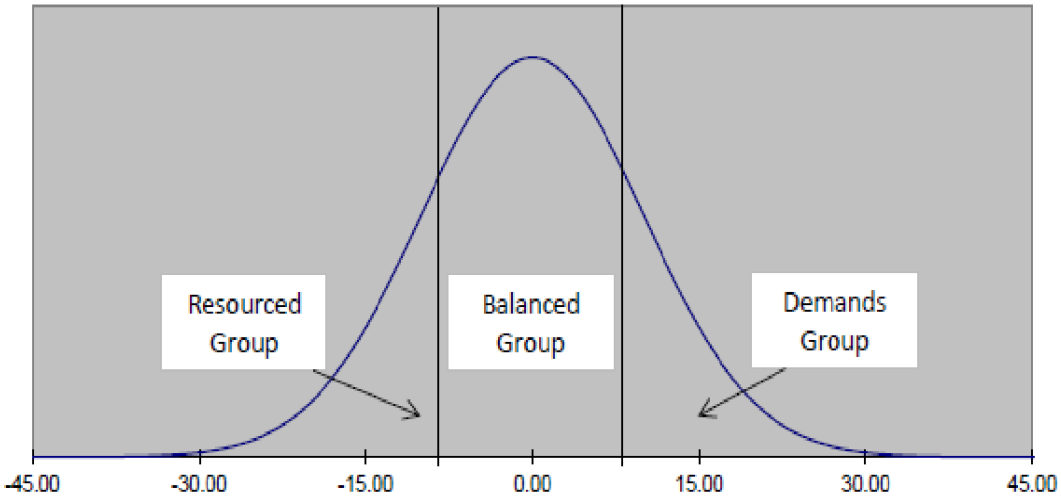
In line with existing research, results from the current study suggest that mentoring programs may serve as an effective tool for addressing the massive burden typically placed upon early-career teachers. Researchers have found mentoring programs have general positive impacts on beginning teachers and decrease turnover among this population (Ingersoll & Kralik, 2004; Smith & Ingersoll, 2004). What was less clear was whether mentoring programs might be able to act specifically on perceived stress among first-year teachers. Indeed, I found that not only are mentored first-year teachers more likely to report lower Appraisal Index scores (i.e., lower risk for stress) than their un-mentored colleagues, but teachers in the Demands group (the highest risk stress group) were less likely to have a mentor than their fellow Resourced or Balanced teachers. However, this pattern of results did not hold up when the binomial logistic regression was run; whether or not a teacher has a mentor did *not* significantly predict their odds of being classified as Demanded. As previously explained, methodological differences may be driving these conflicting results. While not entirely consistent, together these findings tentatively suggest that there may be something advantageous about mentoring as a whole when it comes to teachers' risk for stress. Seeing as stress risk has been linked to a host of critical teacher outcomes (e.g., Lambert et al., 2012), it may be worthwhile to continue focusing time and resources into developing teacher mentoring programs.

As previously mentioned, there tends to be quite a large degree of variation in the characteristics of mentoring programs. Thus, in addition to looking at mentoring more generally, I was also inspired to investigate whether and how particular mentoring supports related to teacher stress risk. Similar patterns were found across types of

mentoring supports, with the Resourced group consistently having the greatest number of teachers receiving each type of mentoring support, followed by Balanced and trailed by Demanded teachers. Distributions of stress groups across support types were markedly similar, suggesting no one form of mentoring support is superior to another. In conclusion, there seems to be a link between receiving mentoring supports -- regardless of the type -- and whether a teacher is classified as at risk for stress. While these results are by no means conclusive, this could mean that a strong mentoring program would include multiple aspects of support. Perhaps it is simply spending one-on-one time with a mentor that is the critical factor at play when it comes to stress risk, or it could be that each type of support offers unique benefits. Future studies could benefit from delving more deeply into how particular profiles of mentoring -- e.g., different combinations of mentoring supports -- relate to differences in stress risk. Based on these future findings, more nuanced recommendations could be made to policy makers and school administrators in terms of how to construct a maximally effective mentoring program.

Appendix A

CARD Classification: Distribution of Appraisal Index Scores



Appendix B

Variable Key

Original NTPS Variables			Corresponding Derived Variables		
Name	***Label***	***Values***	***Name***	***Label***	***Values***
T0100	Main assignment	1=regular full-time teacher	**n/a**	n/a	n/a
T0108	Yr began teaching	[scale]	**BEGAN_2014_DUMMY**	Began teaching in 2014 dummy	1=2014; 0=2015
T0400	Alt cert	1=yes; 2=no	**ALT_CERT_DUMMY**	Alternative certification dummy	1=Alt Cert; 0=Not Alt Cert
T0294	Gender	1=Male; 2=Female	**MALE_DUMMY**	Male dummy	1=male; 0=female
T0928	Hispanic	1=yes; 2=no	**HISPANIC_DUMMY**	Hispanic dummy	1=Hispanic; 0=Non-Hispanic
T0929	Race-White	1=White; -8=valid skip	**WHITE_DUMMY**	White dummy	1=White; 0=Non-White
T0930	Race-Black or African American	1=Black/AA; -8=valid skip	**n/a**	n/a	n/a
T0931	Race-Asian	1=Asian; -8=valid skip (Valid skip otherwise)	**n/a**	n/a	n/a
T0932	Race-Native Hawaiian or Other Pacific Islander	1=Hawaiian/Pac Island; -8=valid skip (Valid skip otherwise)	**n/a**	n/a	n/a
T0933	Race-American Indian or Alaska Native	1=Amer Ind/Alaska Native; -8=valid skip (Valid skip otherwise)	**n/a**	n/a	n/a
MNASGN	General field of main teaching assignment	SPED=2; ESL=5	**SPED_DUMMY** **ESL_DUMMY**	Special Ed Dummy ESL/Bilingual dummy	1=SPED; 0=not special-ed 1=ESL/Bilingual; 0=not ESL/bilingual
IEP_T	% of teacher's students with an IEP	(scale)	**IEP**	% IEP with system missing	(scale)
LEP_T	% of a teacher's students who were ELLs or of LEP	(scale)	**LEP**	% LEP/ELL with system missing	(scale)
URBANS12	Collapsed school locale code (urbanicity)	1=City; 2=Suburb; 3=Town; 4=Rural	**URBANS12**	n/a (stata dummy coded; city is reference group)	n/a
T1523	Assigned mentor/master	1=yes; 2=no	**NO_MENTOR_DUMMY**	No Mentor Dummy	1=No mentor; 0=Has mentor
T1526	Mentor -- Paperwork	1=yes; 2=no	**MENTORSUP_PAPERWORK_DUMMY**	Mentoring supports -- paperwork dummy	1=Yes; 0=No
T1527	Mentor -- Lessons	1=yes; 2=no	**MENTORSUP_LESSONS_DUMMY**	Mentoring supports -- lessons dummy	1=Yes; 0=No
T1528	Mentor -- Learning standards	1=yes; 2=no	**MENTORSUP_STANDARDS_DUMMY**	Mentoring supports -- learning standards dummy	1=Yes; 0=No
T1529	Mentor -- Student assessment	1=yes; 2=no	**MENTORSUP_ASSESSMENT_DUMMY**	Mentoring supports-- student assessment dummy	1=Yes; 0=No

Constructed Variables			Corresponding Derived Variables		
Name	***Label***	***Values***	***Name***	***Label***	***Values***
SCHLEV2	[2-level elem/sec]	1= Elementary;	**ELEMENTARY_DUMMY**	Two-level elementary dummy	1= Elementary;0=Secondary
WF_SS	Workplace fatigue (scaled score)	(scale)	n/a	n/a	n/a
JS_SS	Job satisfaction (scaled score)	(scale)	n/a	n/a	n/a
AUTO_SS	Autonomy (scaled score)	(scale)	n/a	n/a	n/a
RES_SS	Resources (scaled score)	(scale)	n/a	n/a	n/a
DEM_SS	Demands (scaled score)	(scale)	n/a	n/a	n/a
STRESS	Appraisal Index Score	(scale)	n/a	n/a	n/a
STRGRP	Stress Group	1=Resourced; 2=Balanced; 3=Demanded	**DEMANDS_DUMMY**	Demands group dummy	1=Demands Group; 0=Else (Resourced or Balanced)

Appendix C

Table C1

List of CARD-matched NTPS Items

Demands	(14 items)	
NTPS Item #	NTPS Variable #	Content of Survey Item
7-3 c.	T1715	Student misbehavior
7-3 p.	T1728	Class cutting and tardiness
7-4 a.	T1731	Student tardiness
7-4 b.	T1732	Student absenteeism
7-4 c.	T1733[Sec]	Student class cutting
7-4 d.	T1734	Teacher absenteeism
7-4 e.	T1735[Sec]	Students dropping out
7-4 f.	T1736	Student apathy
7-4 g.	T1737	Lack of parental involvement
7-4 h.	T1738	Poverty
7-4 i.	T1739	Students unprepared to learn
7-4 j.	T1740	Poor student health
Resources	(12 items)	
NTPS Item #	NTPS Variable #	Content of Survey Item
7-3 a.	T1713	Supportive admin
7-3 d.	T1716	Supportive parents
7-3 e.	T1717	Necessary materials
7-3 g.	T1719	Principal enforces rules
7-3 h.	T1720	Teachers enforce rules
7-3 i.	T1721	Colleagues share beliefs
7-3 j.	T1722	Principal communicates vision
7-3 k.	T1723	Cooperation among staff
7-3 l.	T1724	Staff recognition
7-3 o.	T1727	Support for students with special needs

Note: [Sec] denotes items used in the Secondary Demands scale only.

Table C2

Wording of NTPS Items Contributing to Stress Group Classification

Resources Scale	Demands Scale
Resources (recoded strongly disagree 1 to strongly agree 4)	*Demands (recoded strongly disagree 1 to strongly agree 4)*
	The level of student misbehavior in this school interferes with my teaching
The school administration's behavior toward the staff is supportive and encouraging	
I receive a greater deal of support from parents for the work I do	The amount of student tardiness and class cutting in this school interferes with my teaching
Necessary materials such as textbooks, supplies, and copy machines are available as needed by the staff	*Demands (serious problem 1 to not a problem 4)*
My principal enforces school rules for student conduct and backs me up when I need it	Student tardiness
Rules for student behavior are consistently enforced by teachers in this school, even for students who are not in their classes	Student absenteeism
Most of my colleagues share my beliefs and values about what the central mission of the school should be	Teacher absenteeism
The principal knows what kind of school he/she wants and has communicated it to the staff	Students cutting class [Sec]
There is a great deal of cooperative effort among the staff members	Students dropping out [Sec]
In this school, staff members are recognized for a job well done	Student apathy
I am given the support I need to teach students with special needs	Lack of parental involvement
	Poverty
	Students come to school unprepared to learn
	Poor student health

Note: [Sec] denotes items used in the Secondary Demands scale only.

Appendix D

NTPS Items Contributing to Mentoring Support Variables

5-7a. In your FIRST year of teaching, were you assigned a master or mentor teacher by your school or district?
(*If you are in your first year of teaching, please answer for THIS school year)*

1. Yes
2. No → Go to section 6 on page 26

5-8. Did your assigned master or mentor teacher provide the following types of support during your FIRST year of teaching?
(*If you are in your first year of teaching, please answer for THIS school year)*

a. Helped with paperwork or record keeping
    1. Yes
    2. No
b. Demonstrated lessons
    1. Yes
    2. No
c. Helped you prepare lessons that address student learning standards
    1. Yes
    2. No
d. Helped you develop student assessment tools
    1. Yes
    2. No

Appendix E

NTPS Items Contributing to Occupational Health Variables

Autonomy	(13 items)	
NTPS Item #	NTPS Variable #	Content of Survey Item
7-1 a.	T1700	School influence - student performance standards
7-1 b.	T1701	Pedagogical influence - curriculum
7-1 c.	T1702	School influence - professional development programs
7-1 d.	T1703	School influence - evaluating teachers
7-1 e.	T1704	School influence - hiring new teachers
7-1 f.	T1705	School influence - discipline
7-1 g.	T1706	School influence - budget
7-2 a.	T1707	Pedagogical influence - selecting textbooks
7-2 b.	T1708	Pedagogical influence - selecting content
7-2 c.	T1709	Classroom control - selecting techniques
7-2 d.	T1710	Classroom control - grading students
7-2 e.	T1711	Classroom control - discipline
7-2 f.	T1712	Classroom control - homework

Workplace Fatigue	(5 items)	
NTPS Item #	NTPS Variable #	Content of Survey Item
7-5 a.	T1741	Agree - teaching not worth it
7-5 d.	T1744	Agree - leave for better pay
7-5 e.	T1745	Agree - transfer to another school
7-5 f.	T1746	Agree - less enthusiasm
7-5 g.	T1747	Agree - too tired for school

Job Satisfaction	(3 items)	
NTPS Item #	NTPS Variable #	Content of Survey Item
7-3 q.	T1729	Generally satisfied
7-5 b.	T1742	Agree- teachers satisfied
7-5 c.	T1743	Agree - school is well run

Appendix F

Study Covariates

Gender (Male)
Ethnicity (Hispanic)
Race (White)
Year began teaching (2014, 2015)
Special education teacher
ESL/bilingual teacher
Urbanicity of school (city, suburban, town, rural)
Percent of teacher's students with an IEP
Percent of teacher's students designated as LEP or ELL
School level taught (elementary, secondary)

References

Achinstein, B., Ogawa, R., Sexton, D., & Freitas, C. (2010). Retaining teachers of color: A pressing problem and a potential strategy for 'hard-to-staff' schools. *Review of Educational Research, 80*(1), 71-101.

Berman, S. (2004). Effective recruitment and induction into the school district. In Goodlad, J. & McMannon, T. (Eds.), *The teaching career* (pp. 117-139). New York: Teachers College Press.

Bradley, L., & Gordon, S. P. (1994). Comparing the ideal to the real in state-mandated teacher induction programs. *Journal of Staff Development*, *15*(3), 44–48.

Burke, P. F., Aubusson, P. J., Schuck, S. R., Buchanan, J. D., & Prescott, A. E. (2015). How do early career teachers value different types of support? A scale-adjusted latent class choice model. *Teaching and Teacher Education*, *47*, 241–253. https://doi.org/10.1016/j.tate.2015.01.005

Chang, M. L. (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. *Educational Psychology Review*, *21*(3), 193–218. https://doi.org/10.1007/s10648-009-9106-y

Chun, J. U., Sosik, J. J., & Yun, N. Y. (2012). A longitudinal study of mentor and protégé outcomes in formal mentoring relationships. *Journal of Organizational Behavior*, *33*(8), 1071–1094.

Cornu, R. L. (2005). Peer mentoring: Engaging pre-service teachers in mentoring one another. *Mentoring & Tutoring: Partnership in Learning*, *13*(3), 355–366.

Costigan, A. T., & Crocco, M. S. (2004). *Learning to teach in an age of accountability*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cox, T. (1978). *Stress*. Macmillan.

DePaul, A. (2000). *Survival guide for new teachers: How new teachers can work effectively with veteran teachers, parents, principals, and teacher educators*. Retrieved from http://eric.ed.gov/?id=ED442791

Doney, P. A. (2013). Fostering resilience: A necessary skill for teacher retention. *Journal of Science Teacher Education*, *24*(4), 645–664. https://doi.org/10.1007/s10972-012-9324-x

Elze, M. C., Gregson, J., Baber, U., Williamson, E., Sartori, S., Mehran, R., Nichols, M., Stone, G. W., & Pocock, S. J. (2017). Comparison of propensity score methods and covariate adjustment: Evaluation in 4 cardiovascular studies. *Journal of the American College of Cardiology*, *69*(3), 345–357. https://doi.org/10.1016/j.jacc.2016.10.060

Farber, B. A. (1991). Crisis in education: Stress and burnout in the American teacher. Jossey-Bass.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record*, *103*(6), 1013–1055. https://doi.org/10.1111/0161-4681.00141

Fessler, R. (1992). The teacher career cycle. In R. Fessler & J. Christensen (Eds.), *The teacher career cycle: Understanding and guiding the professional development of teachers*. Boston: Allyn and Bacon.

Fideler, E., & Haselkorn, D. (1999). *Learning the ropes: Urban teacher induction programs and practices in the United States*. Belmont, MA: Recruiting New Teachers.

Fitchett, P. G., Dillard, J., McCarthy, C. J., Lambert, R. G., & Mosley, K. (2020). Examining the intersectionality among teacher race/ethnicity, school context, and risk for occupational stress. *Education Policy Analysis Archives*, *28*, 87.

Fitchett, P. G., McCarthy, C. J., Lambert, R. G., & Boyle, L. (2018). An examination of US first-year teachers' risk for occupational stress: Associations with professional preparation and occupational health. *Teachers and Teaching*, *24*(2), 99–118. https://doi.org/10.1080/13540602.2017.1386648

Fitchett, P. G., McCarthy, C. J., Lambert, R. G., Eyal, M., Playfair, E. C., & Dillard, J. B. (2019). Examining teacher stress-vulnerability in the US secondary school context. *Educational Review*, *0*(0), 1–24. https://doi.org/10.1080/00131911.2019.1619521

French, N. K. (1993). Elementary teacher stress and class size. *Journal of Research & Development in Education*, *26*(2), 66–73.

Gaede, O. F. (1978). Reality shock: A problem among first-year teachers. *The Clearing House*, *51*(9), 405–409. Retrieved from JSTOR.

Gilbert, R. B., Adesope, O. O., & Schroeder, N. L. (2014). Efficacy beliefs, job satisfaction, stress and their influence on the occupational commitment of English-medium content teachers in the Dominican Republic. *Educational Psychology*, *34*(7), 876–899. https://doi.org/10.1080/01443410.2013.814193

Glazerman, S., Isenberg, E., Dolfin, S., Bleeker, M., Johnson, A., Grider, M., Jacobus, M., Ali, M., Duncan, A., Easton, J. Q. (2010). *Impacts of comprehensive teacher induction: Final results from a randomized controlled study*.

Goddard, J. T., & Foster, R. Y. (2001). The experiences of neophyte teachers: A critical

constructivist assessment. *Teaching and Teacher Education*, *17*(3), 349–365.

Goddard, R., O'Brien, P., & Goddard, M. (2006). Work environment predictors of beginning teacher burnout. *British Educational Research Journal*, *32*(6), 857–874.

Goldrick, L. (2009). *A teacher development continuum: The role of policy in creating a supportive pathway into the profession* (p. 14). Retrieved from New Teacher Center website: https://newteachercenter.org/wp-content/uploads/BRF_ATeacherDevelopmentContinuum_TheRoleofPolicy.pdf

Gray, L., & Brauen, M. (2013). Strategies for Longitudinal Analysis of the Career Paths of Beginning Teachers: Results from the First through Fourth Waves of the 2007-08 Beginning Teacher Longitudinal Study. Research and Development Report. NCES 2013-336. *National Center for Education Statistics*. Retrieved from http://eric.ed.gov/?id=ED544179

Grayson, J. L., & Alvarez, H. K. (2008). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education*, *24*(5), 1349–1363.

Griffin, A., & Tackie, H. (2016). *Through our eyes: Perspectives and reflections from black teachers.* Washington, DC: U.S. Department of Education.

Grissom, J. A., Kern, E. C., & Rodriguez, L. A. (2015). The "representative bureaucracy" in education educator workforce diversity, policy outputs, and outcomes for disadvantaged students. *Educational Researcher*, *44*(3), 185–192. https://doi.org/10.3102/0013189X15580102

Guin, K. (2004). Chronic teacher turnover in urban elementary schools. *Education Policy Analysis Archives*, *42*(12), 1–30.

Hargreaves, A. (2005). Educational change takes ages: Life, career and generational factors in teachers' emotional responses to educational change. *Teaching and Teacher Education*, *21*(8), 967–983.

Hebert, E., & Worthy, T. (2001). Does the first year of teaching have to be a bad one? A case study of success. *Teaching and Teacher Education*, *17*(8), 897–911.

Hobson, A. J., Ashby, P., Malderez, A., & Tomlinson, P. D. (2009). Mentoring beginning teachers: What we know and what we don't. *Teaching and Teacher Education*, *25*(1), 207–216. https://doi.org/10.1016/j.tate.2008.09.001

Hong, J. Y. (2012). Why do some beginning teachers leave the school, and others stay? Understanding teacher resilience through psychological lenses. *Teachers and Teaching*, *18*(4), 417–440.

Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education*, *21*(4), 343–356.

Huberman, M. (1989). The professional life cycle of teachers. *Teachers College Record*, *91*(1), 31–57.

Ingersoll, R. (2003). Is there really a teacher shortage? A research report. *Center for the Study of Teaching and Policy*. Retrieved from http://eric.ed.gov/?id=ED499091

Ingersoll, R., & Kralik, J. M. (2004). The impact of mentoring on teacher retention: What the research says. *GSE Publications*, 127.

Ingersoll, R., Merrill, E.; Stuckey, D., & Collins, G. (2018). Seven trends: The transformation of the teaching force – Updated October 2018. CPRE Research Reports. Retrieved from https://repository.upenn.edu/cpre_researchreports/108.

Ingersoll, R., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, *81*(2), 201–233.

Ingersoll, R., & Smith, T. M. (2004). Do teacher induction and mentoring matter? *NASSP Bulletin*, *88*(638), 28–40.

Jenlink, P. M. (2014). *Teacher identity and the struggle for recognition: Meeting the challenges of a diverse society*. R&L Education.

Jennings, P. A., Frank, J. L., Snowberg, K. E., Coccia, M. A., & Greenberg, M. T. (2013). Improving classroom learning environments by Cultivating Awareness and Resilience in Education (CARE): Results of a randomized controlled trial. *School Psychology Quarterly*, *28*(4), 374.

Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology*, *20*(2), 178–187.

Kapadia, K., Coca, V., & Easton, J. Q. (2007). *Keeping new teachers: A first look at influences of induction in the Chicago Public Schools*. Retrieved from Consortium on Chicago School Research, University of Chicago website: https://core.ac.uk/download/pdf/71348059.pdf

Katz, L. G. (1972). Developmental stages of preschool teachers. *The Elementary School Journal*, *73*(1), 50–54.

Kee, A. N. (2012). Feelings of preparedness among alternatively certified teachers. *Journal of Teacher Education*, *63*(1), 23–38. https://doi.org/10.1177/0022487111421933

Kraft, & Papay, J. P. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Effectiveness and Policy Analysis*, *36*(4), 476–500.

Lambert, R. G., Fitchett, P. G., McCarthy, C., & Boyle, L. H. (2016, July). *Identification of beginning U.S. teachers' risk for stress and occupational health*. Presented at the Annual Conference of the Society for Stress and Anxiety Research, Zagreb, Croatia.

Lambert, R. G., McCarthy, C. J., Fitchett, P. G., Lineback, S., & Reiser, J. (2015). Identification of elementary teachers' risk for stress and vocational concerns using the national Schools and Staffing Survey. *Education Policy Analysis Archives*, *23*(43). http://doi.org/10.14507/epaa.v23.1792

Lambert, R. G., McCarthy, C., O'Donnell, M., & Wang, C. (2009). Measuring elementary teacher stress and coping in the classroom: Validity evidence for the Classroom Appraisal of Resources and Demands. *Psychology in the Schools*, *46*(10), 973–988.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer Publishing Company.

Lhospital, A. S., & Gregory, A. (2009). Changes in teacher stress through participation in pre-referral intervention teams. *Psychology in the Schools*, *46*(10), 1098–1112.

Little, J. W. (1990). Chapter 6: The mentor phenomenon and the social organization of teaching. *Review of Research in Education*, *16*(1), 297–351. https://doi.org/10.3102/0091732X016001297.

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

Chicago Press.

McCarthy, C. J., Fitchett, P. G., Lambert, R. G., & Boyle, L. (2019). Stress vulnerability in the first year of teaching. *Teaching Education*, 1–20.

McCarthy, C. J., Lambert, R. G., Lineback, S., Fitchett, P., & Baddouh, P. G. (2016). Assessing teacher appraisals and stress in the classroom: Review of the classroom appraisal of resources and demands. *Educational Psychology Review*, *28*(3), 577-603.

McCarthy, C. J., Lambert, R. G., O'Donnell, M., & Melendres, L. T. (2009). The relation of elementary teachers' experience, stress, and coping resources to burnout symptoms. *The Elementary School Journal*, *109*(3), 282–300. https://doi.org/10.1086/592308

McCarthy, C. J., Lambert, R. G., & Reiser, J. (2014). Vocational concerns of elementary teachers: Stress, job satisfaction, and occupational commitment. *Journal of Employment Counseling*, *51*(2), 59–74. https://doi.org/10.1002/j.2161-1920.2014.00042.x

McCarthy, C. J., Lineback, S., Boyle, L. H., Fitchett, P. G., & Lambert, R. (2016, July). *Perceived demands and resources among early career teachers: Linking risk for stress with professional mobility*. Presented at the Annual Conference of the Society for Stress and Anxiety Research, Zagreb, Croatia.

Mitchell, H. J. (1999). Group Mentoring: Does it work? *Mentoring & Tutoring*, *7*(2), 113–120.

Prilleltensky, I., Neff, M., & Bessell, A. (2016). Teacher stress: What it is, why it's important, how it can be alleviated. *Theory Into Practice*, *55*(2), 104–111.

https://doi.org/10.1080/00405841.2016.1148986

Raj, D. (1964). The use of systematic sampling with probability proportionate to size in a large scale survey. *Journal of the American Statistical Association*, *59*(305), 251–255. https://doi.org/10.1080/01621459.1964.10480715

Riddles, M., Wallace, L., Rizzo, L., & Marker, D. (2017). *NTPS 2015-2016 analysis of response rates and field collection experience*. Rockville, MD: Westat.

Roehrig, A. D., Bohn, C. M., Turner, J. E., & Pressley, M. (2008). Mentoring beginning primary teachers for exemplary teaching practices. *Teaching and Teacher Education*, *24*(3), 684–702. https://doi.org/10.1016/j.tate.2007.02.008

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4–36.

Ronfeldt, M., & McQueen, K. (2017). Does new teacher induction really improve retention? *Journal of Teacher Education*, *68*(4), 394–410. https://doi.org/10.1177/0022487117702583

Rust, F. O. (1994). The first year of teaching: It's not what they expected. *Teaching and Teacher Education*, *10*(2), 205–217.

Ryan, S. V., von der Embse, N. P., Pendergast, L. L., Saeki, E., Segool, N., & Schwing, S. (2017). Leaving the teaching profession: The role of teacher stress and educational accountability policies on turnover intent. *Teaching and Teacher Education*. 66. 1-11. Retrieved from http://dx.doi.org/10.1016/j.tate.2017.03.016.

Singer, J. D., & Willett, J. B. (1996). Methodological issues in the design of longitudinal research: Principles and recommendations for a quantitative study of teachers' careers. *Educational Evaluation and Policy Analysis*, *18*(4), 265–283.

Skaalvik, S., & Skaalvik, E. M. (2011). Teacher job satisfaction and motivation to leave the teaching profession: Relations with school context, feeling of belonging, and emotional exhaustion. *Teaching and Teacher Education*, *27*(6), 1029–1038. https://doi.org/10.1016/j.tate.2011.04.001

Smith, T. M., & Ingersoll, R. (2004). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal*, *41*(3), 681–714.

Spooner-Lane, R. (2017). Mentoring beginning teachers in primary schools: Research review. *Professional Development in Education*, *43*(2), 253–273. https://doi.org/10.1080/19415257.2016.1148624

SREB. (2018). *Mentoring new teachers: A fresh look*. Southern Regional Education Board.

Taie, S., & Goldring, R. (2017). Characteristics of public elementary and secondary school teachers in the United States: Results from the 2015-16 National Teacher and Principal Survey. First look. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017072.

Thompson, K. (2017) *Elementary Title I teachers perception of stress, burnout and the impact on retention.* (Ed.D. Dissertations). 71. Retrieved from https://commons.cuportland.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article =1111&context=edudissertations

Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research*, *54*(2), 143–178.

Veldman, I., Admiraal, W., van Tartwijk, J., Mainhard, T., & Wubbels, T. (2016). Veteran

teachers' job satisfaction as a function of personal demands and resources in the relationships with their students. *Teachers and Teaching*, *22*(8), 913–926.

Weasmer, J., & Woods, A. M. (1998). Facilitating success for new teachers. *Principal*, *78*(2), 40–42.

Wong, H. K. (2004). Induction programs that keep new teachers teaching and improving. *NASSP Bulletin*, *88*(638), 41–58.

Youngs, P., Frank, K., Thum, Y.M. & Low, M. (2012). The Motivation of teachers to produce human capital and conform to their social contexts. *Yearbook of the National Society for the Study of Education, 111*, 248-272.

Zellars, K. L., Hochwarter, W. A., Perrewé, P. L., Hoffman, N., & Ford, E. W. (2004). Experiencing job burnout: The roles of positive and negative traits and states. *Journal of Applied Social Psychology*, *34*(5), 887–911. https://doi.org/10.1111/j.1559-1816.2004.tb02576.x