AERA Open
January-March 2017, Vol. 3, No. 1, pp. 1–17
DOI: 10.1177/2332858416684764

A Temperament for Teaching? Associations Between Personality Traits and Beginning Teacher Performance and Retention

Kevin C. Bastian
University of North Carolina at Chapel Hill
David M. McCord
Western Carolina University
Julie T. Marks
University of North Carolina at Chapel Hill
Dale Carpenter
Western Carolina University

The "greening" (i.e., inexperience) of the U.S. teacher workforce puts a premium on districts and schools hiring effective and persistent beginning teachers. Given the limitations of characteristics currently available at the time of hiring (e.g., academic ability, preparation type), we built off previous research in economics and psychology to investigate associations between personality traits and first-year teachers' performance and retention in North Carolina public schools. Using the 5-factor model of personality, we find that conscientiousness is significantly associated with higher value-added estimates, higher evaluation ratings, and higher retention rates. Additionally, general self-efficacy, a subdomain of conscientiousness, is significantly associated with teacher value added and evaluation ratings. These conscientiousness results are consistent with a rich body of evidence connecting conscientiousness-related measures to employee performance and retention across professions, and they suggest that districts and schools should consider using personality trait measures, along with other valid indicators, as a way to improve teacher hiring decisions.

Keywords: *personality traits, teacher effectiveness, teacher retention*

TWENTY-FIVE YEARS ago, the modal teacher experience category in the United States was 15 years. Today, first-year teachers are the largest experience category, and approximately 25% of the teacher workforce has <5 years of experience (Ingersoll & Merrill, 2012). This "greening" of the teacher workforce is concerning for three reasons. First, novice teachers are typically less effective than their more experienced peers (Chingos & Peterson, 2011; Papay & Kraft, in press). Relative to first-year teachers, for example, North Carolina elementary grade teachers with 3 to 5 years of experience add 9% and 6% of a *SD* to student achievement on end-of-grade exams in mathematics and reading (Clotfelter, Ladd, & Vigdor, 2007). Second, novice teachers are significantly more likely to exit the profession. This attrition entails financial costs for districts and has adverse effects on school stability and student achievement (Alliance for Excellent Education, 2014; Ronfeldt, Loeb, & Wyckoff, 2013). Finally, novice teachers are disproportionately employed in high-need schools and classrooms, where students need the highest-quality resources to enhance their achievement (Bastian, Henry, & Thompson, 2013; Clotfelter et al., 2005).

Given the performance and attrition of early career teachers, this greening of the teacher workforce puts a premium on districts and schools making high-quality teacher hiring decisions. To make these decisions, districts and schools need to know the characteristics of teachers that significantly predict performance and retention. While research evidence indicates that characteristics typically available at the time of hiring—such as academic ability measures (e.g., SAT scores, grade point averages, licensure exam scores) or teacher preparation type—are associated with teacher effectiveness and retention, these characteristics explain only a small portion of the variance in these teacher outcomes (Borman & Dowling, 2008; Goldhaber, Liddle, & Theobald, 2013; Greenwald, Hedges, & Laine, 1996; Kane, Rockoff, & Staiger, 2008; Rockoff, Jacob, Kane, & Staiger, 2011; Wayne & Youngs, 2003). Districts and schools should consider such

measures in hiring decisions, but they need additional predictors to consistently hire teachers who will be effective and remain in teaching.

Toward this end, seminal research in psychology and economics highlights the importance of personality traits and personal qualities (Duckworth & Yeager, 2015) to individuals' academic, workplace, and livelihood outcomes (Ozer & Benet-Martinez, 2006). For example, measures of self-discipline, not IQ scores, more accurately predict middle school students' grade point averages (Duckworth & Seligman, 2005). Similarly, measures of childhood self-control significantly predict physical health, substance dependence, personal finances, and criminal outcomes for adults (Moffitt et al., 2011). Building from this body of work, nascent research in education indicates that teachers' personality traits are significantly associated with their performance and retention outcomes (Klassen & Tze, 2014; Robertson-Kraft & Duckworth, 2014; Rockoff et al., 2011). Personality traits, paired with other characteristics available at hiring, may present a way to predict more of the variation in teacher performance and enable the hiring of more effective and persistent teachers (Rockoff et al., 2011). This is particularly relevant since personality traits are not simply redundant with academic traits; rather, they predict unique sources of variation in outcomes of interest (Lounsbury, Sundstrom, Loveland, & Gibson, 2003).

Therefore, to further this emerging research agenda, we assess the associations between personality traits and the performance and retention outcomes of first-year teachers in North Carolina public schools (NCPS). Specifically, we ask the following research questions:

Research Question 1: Are the personality traits of first-year teachers associated with their value-added estimates?

Research Question 2: Are the personality traits of first-year teachers associated with their evaluation ratings?

Research Question 3: Are the personality traits of first-year teachers associated with their retention in NCPS?

Importantly, our analyses make two key additions to previous research connecting personality traits and personal qualities to teacher outcomes. First, compared with previous studies, our sample is larger and includes teachers with diverse forms of preparation and varying levels of academic ability (Dobbie, 2011; Duckworth, Quinn, & Seligman, 2009; Robertson-Kraft & Duckworth, 2014; Rockoff et al., 2011). This means that our results may be generalizable to a larger body of teachers. Second, while previous analyses have focused on a small number of personality trait measures—not all of which have empirical data on their validity and reliability—we employ a valid, reliable, and empirically based personality trait framework. With a more generalizable sample and a set of valid and reliable measures, we make an incremental yet important contribution: providing estimates of the associations between personality traits and the performance and retention of beginning teachers.

In the following sections we first summarize theoretical frameworks on teaching quality, describe our personality trait framework, and review research evidence connecting personality traits to workplace and teaching outcomes. Next, we detail our research sample, measures, and data analysis. Finally, we present our findings and close with a discussion of the limitations and implications of our results.

Background

Theoretical Frameworks for Teaching Quality

Over many decades, research has coalesced around several theories to explain teaching quality. Initial research (prior to the 1950s) used surveys to discern the characteristics and traits of teachers (including their dispositions and personality) that predicted student achievement gains (Cochran-Smith & Fries, 2005). While generating long lists of teacher traits correlated with effectiveness, this early research was criticized for its lack of rigor (Barr et al., 1952) and lack of results—much was still unknown about what predicts teaching success (Barr, 1953). From this dissatisfaction, education researchers proposed a more scientifically rigorous research agenda—observing classrooms in experimental and quasi-experimental designs—to identify teaching behaviors associated with student outcomes. Known as "process-product"—since a teacher process was hypothesized to lead to a student product—this research (carried out in the 1960s through the 1980s) sought to pinpoint distinct teaching actions that could be transported into teacher education and professional development. Essentially, effective teaching was a series of circumscribed teacher behaviors (Brophy & Good, 1986; Gage, 1964, 1978).

While noting benefits of process-product research, critics argued that observing teacher behaviors was not enough; instead, to understand effective teaching, it was important to know what teachers knew and how they made teaching decisions (Shulman, 1986). Thus, a next theory of teaching quality (prevalent in the 1980s and 1990s) focused on teachers' cognition, content knowledge, and pedagogical content knowledge. Acknowledging that teacher tasks matter, these ethnographic approaches asserted that it is how teachers implement those tasks, with which students, and at which times that is key to teaching quality (Grimmett & Mackinnon, 1992). Finally, with the rise of school and teacher accountability regimes linked to student test scores, a recent research agenda has defined teacher effectiveness according to the teacher characteristics and credentials significantly associated with teacher value-added estimates (Clotfelter et al., 2007, 2010).

Although the present study is most easily tied to trait-based theories of teacher effectiveness—for example,

effective teachers are outgoing, agreeable, conscientious—we contend that the relationship is more complex. Teachers with high levels of a given personality trait may be more effective; however, it is not the trait alone that explains teacher performance. Rather, it is how that trait influences teacher behavior, decision making, and learning that matters. Therefore, this study can connect to a range of theories on teacher effectiveness.

Five-Factor Model of Personality: Theory and Measurement

In the foundational discipline of psychology, the last two to three decades have seen remarkable progress in the development of a broad, empirically based, and comprehensive model of personality traits that captures a large proportion of individual differences in personality. The five-factor model of personality (FFM) has become the new paradigm for personality research (Marsh et al., 2010; McCrae, 2011). This model posits five broad trait domains, often called the "Big Five," which are normally distributed in the population and together account for approximately 75% of the variance in human personality. These Big Five personality domains are *extraversion* (outgoing and energetic vs. solitary and reserved), *agreeableness* (friendly and compassionate vs. analytical and detached), *conscientiousness* (hardworking and organized vs. careless and unreliable), *neuroticism* (sensitive and nervous vs. calm and secure), and *openness to experience* (curious and imaginative vs. conventional and cautious). Hundreds of empirical studies have been conducted demonstrating that the Big Five personality traits are (a) reliably observed across raters and methods (e.g., McCrae et al., 2004), (b) quite stable across the life span in adulthood (Caspi, 2000; Terracciano, Costa, & McCrae, 2006), (c) substantially heritable (Bouchard & Loehlin, 2001), (d) evidenced across a wide range of cultures (Schmitt et al., 2007), and (e) important influences on many aspects of life (Ozer & Benet-Martinez, 2006). The fully developed FFM, as offered by Costa and McCrae (1995), includes six narrower facets under each domain, resulting in the framework shown in Table A1.

A number of self-report measurement instruments have been developed to assess the domains of the FFM. The first to include the fully faceted model was the NEO-Personality Inventory–Revised (NEO-PI-R; Costa & McCrae, 1992). Because the NEO-PI-R is copyright restricted and cannot be modified without permission, researchers have in many cases turned to public domain proxy scales available through Goldberg's (1999) International Personality Item Pool—the website of which offers a 300-item proxy for the NEO-PI-R (10 items per facet). This instrument exhibits excellent psychometric properties and very high correlations with the parent scale, but it is even longer than the NEO-PI-R itself and is unwieldy in many research designs. Johnson (2014) has collected >300,000 cases using this long form and has developed a much shorter form, known as the M5-120 (120 items), using the best 4 items for each of the 30 personality trait facets. With the M5-120, test takers respond to items on a 5-point Likert-type scale to indicate how accurately statements apply to their personality—very inaccurate, inaccurate, neither inaccurate nor accurate, accurate, and very accurate. We selected the M5-120 for use in the current study given its manageable length, excellent reliability in multiple large samples—alpha values range from 0.83 to 0.90 for Big Five domains and from 0.63 to 0.88 across all 30 facets (Johnson, 2014)—and ability to be administered via an online survey platform. The M5-120 has now been used in many thesis projects, presentations, and published studies, including those targeting teacher characteristics (Cooper, Carpenter, Reiner, & McCord, 2014).

Impacts of Personality Traits in the Workplace and Teaching

Given the prevalence of the FFM, an extensive literature base has evolved to estimate the associations between personality traits and work-related outcomes, including motivation, working environment preferences, job performance, and job retention. Regarding job performance—closely tied to our focus on teacher value added and evaluation ratings—initial and follow-up meta-analyses identify significant results for personality traits (Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Hurtz & Donovan, 2000; Judge, Heller, & Mount, 2002). Overall, conscientiousness is one of the best predictors of job performance; the combination of conscientiousness and emotional stability (low neuroticism) predicts job performance in a range of specific occupations. For instance, conscientiousness, extraversion, agreeableness, openness to experience, and emotional stability are valid predictors of performance in customer service jobs; in skilled and semiskilled positions without a significant interpersonal component, conscientiousness is often the only consistent predictor from the FFM (Hurtz & Donovan, 2000). With regard to job retention, initial analyses showed that conscientiousness is negatively associated with a variety of organizational withdrawal behaviors, such as absenteeism and quitting (Barrick & Mount, 1991). Additional analyses suggest that conscientiousness and emotional stability are positively associated with job retention (Barrick & Mount, 1996; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990).

Turning to teaching, efforts to measure the personality traits and personal qualities of prospective teachers are not new. A number of school districts employ scripted questionnaires or interview protocols, such as the Haberman Star Teacher Pre-Screener or the Teacher Perceiver Interview, to identify prospective teachers' personality traits and personal qualities and make hiring decisions. With the increasing

prevalence of student achievement and teacher evaluation data, a nascent research agenda is now able to investigate whether personality traits are associated with teacher performance and retention. Here, initial studies have returned promising findings. For example, Rockoff and colleagues show that a 1-*SD* increase in a latent noncognitive construct—extraversion, conscientiousness, personal efficacy, general efficacy, and the Haberman Star Index total score—led to a significant improvement of 0.033 *SD* in student math achievement and 0.272 *SD* in teachers' subjective evaluations (Rockoff et al., 2011). In work with Teach for America corps members, (a) Dobbie (2011) finds that a 1-*SD* increase in corps members' leadership and perseverance predicts achievement gains in math of 0.054 and 0.040 *SD*, respectively; (b) Duckworth and colleagues (2009) show that a 1-*SD* increase in measures of corps members' self-reported grittiness and life satisfaction is associated with 23% and 36% increased probability, respectively, of corps members reporting significant student achievement gains; and (c) Robertson-Kraft and Duckworth (2014) find that a measure of grit taken from corps members' resumes significantly predicts teacher performance and retention. A meta-analysis by Klassen and Tze (2014) indicated that personality traits, particularly self-efficacy (a facet of the conscientiousness domain), exhibit significant associations with evaluated teacher performance and student achievement. Last, regarding job satisfaction, Cooper and colleagues (2014) find that extraversion, agreeableness, and neuroticism (low) are significant predictors for public school teachers.

While this evidence suggests that personality traits and personal qualities help to explain what makes teachers effective and persistent, our analyses make two important additions to the existing research connecting personality traits to teacher outcomes. First, the sample in previous analyses has been small (Klassen & Tze, 2014) and often focused on teachers from Teach for America, a highly selective alternative entry program. For example, studies by Duckworth and colleagues (2009), Dobbie (2011), and Robertson-Kraft and Duckworth (2014) all focus on Teach for America corps members. Duckworth and colleagues and Dobbie included sample sizes of 390 and 384 teachers, respectively; Robertson-Kraft and Duckworth had two studies that included 154 teachers and 307 teachers. Rockoff and colleagues (2011) estimate models with >1,000 New York City mathematics teachers, but they have complete noncognitive data for only 333 teachers. Likewise, in the meta-analysis by Klassen and Tze (2014), nearly 80% of the reviewed studies included <200 teachers. Our sample includes 1,790 first-year teachers with diverse forms of teacher preparation and academic ability. Second, previous analyses have focused on only a small number of personality trait or personal quality measures—for example, grittiness, leadership, perseverance, organizational ability, and motivational ability—and not all of these measures have empirical data on their validity and reliability. We employ the FFM: a valid, reliable, and empirically based framework.

Given previous studies on the FFM and employment outcomes, we hypothesize that the conscientiousness domain and its self-efficacy facet will be significantly associated with the value-added estimates and evaluation ratings of beginning teachers. Although the literature is less robust, we also hypothesize that conscientiousness will be significantly associated with retention outcomes for beginning teachers. Other personality traits may be significantly associated with well-aligned teacher outcomes (e.g., intellect and teachers' evaluation ratings on the content knowledge standard, cautiousness and teacher retention). We consider this an exploratory study to identify those significant results and test our main hypothesis.

Data

Research Sample

To assess the relationships between personality traits and beginning teacher outcomes, we used an online survey platform to distribute the M5-120 personality trait assessment to all first-year teachers in NCPS in the 2013–2014 school year. We focus on beginning teachers, rather than a sample of more experienced teachers, given the greening of the teacher workforce and our interest in the potential use of personality traits to improve district and school hiring practices. Ideally, we would have surveyed teachers prior to their beginning teaching; however, data identifying first-year teachers and their contact information were not available until well past the start of the school year. To identify and contact these first-year teachers, we used January 2014 salary data and teacher email addresses provided by the North Carolina Department of Public Instruction. We opened the M5-120 administration in early March 2014 and closed the assessment in early April 2014. This timeline was similar to research by Rockoff and colleagues (2011) investigating the associations between cognitive and noncognitive characteristics and beginning teacher outcomes in New York City. In total, we sent the M5-120 to 6,421 first-year teachers, and 1,790—a response rate of 27.88%—completed all of the personality trait items. Of these 1,790 first-year teacher respondents, 740 worked in elementary schools, 347 in middle schools, 587 in high schools, and 116 in schools with other grade configurations (K–8, K–12, etc.).

While 1,790 respondents form a large sample for analyses investigating the relationships between personality traits and teacher outcomes, we want to make inferences that are valid beyond our sample of respondents. Toward this end, there are two concerns. First, our analyses include only those who were hired to teach in NCPS and not the full set of individuals who applied for teaching positions. It is possible that this introduces selection bias if, for example, districts hire teachers from the applicant pool with higher levels of a given personality trait.

TABLE 1
Characteristics of M5-120 Respondents and Nonrespondents

	M5-120	
Characteristics	Respondents	Nonrespondents
Female	80.65*	77.93
Minority	19.09+	21.02
Age	28.21**	26.93
Teacher preparation		
In-state prepared	49.97**	55.86
Out-of-state prepared	26.91+	24.49
Alternative entry	23.11**	19.65
Select teaching licenses		
Elementary Grades	39.42**	44.67
Mathematics	10.78	11.26
Science	9.70	8.90
English	12.38	11.31
Social studies/history	11.92	11.40
Special education	13.29**	8.83
Standardized value-added estimate	−0.334	−0.375
Evaluation standard		
Leadership	3.06	3.04
Classroom Environment	3.14	3.11
Content Knowledge	3.05*	3.02
Facilitating Student Learning	3.04	3.02
Reflecting on Practice	3.08**	3.04
Retention rate in North Carolina public schools	87.59	86.03
School size	787.99	785.08
Total per-pupil expenditures	8,500.55	8,410.05
Average teacher salary supplements	3,451.74**	3,604.29
Short-term suspension rate (per 100 students)	18.71*	16.89
Violent act rate (per 1,000 students)	8.98*	7.87
Percentage free and reduced-price lunch	64.44	64.64
Percentage minority	56.69*	58.17
Performance composite	51.31	51.31
Teacher count	1790	4631

+$p < .10$. *$p < .05$. **$p < .01$.

Second, our analyses focus only on respondents to the M5-120 among those hired. Therefore, Table 1 displays descriptive data—teacher demographics, teacher preparation route, select teaching license areas, teacher outcomes, and school characteristics—on the extent to which respondents to the M5-120 differed from their nonrespondent peers. In comparison to nonrespondents, M5-120 completers were more likely to be female and white and were slightly older. Regarding teacher preparation, a lower percentage of respondents were prepared at in-state teacher education programs; conversely, a higher percentage of respondents were prepared out of state or entered the profession alternatively. A lower percentage of respondents held elementary grades teaching licenses, and a higher percentage of respondents had special education licenses. Importantly, the value-added estimates and retention rates of respondents were comparable to those of nonrespondents. Regarding teacher evaluation ratings, respondents had significantly higher ratings on two standards (Standard 3: content knowledge and Standard 5: reflecting on practice) and comparable ratings on the remaining three standards. Finally, respondents worked in schools with fewer minority students and with higher rates of short-term suspension and violent acts. The percentage of free and reduced-price lunch students and the percentage of state assessments passed (performance composite) were similar across groups. Overall, while these observable differences between respondents and nonrespondents should not seriously compromise the generalizability of results, we caution that findings still need to be interpreted carefully.

Outcome Measures

We include three teacher outcome measures in this study: teacher value-added scores, teacher evaluation ratings, and teacher retention. The full sample includes all 1,790 first-year teachers with personality trait data. However, as detailed in the following, the analysis samples differ among the teacher outcome measures, on the basis of the available outcome data and the missingness in the model covariates.

Teacher Value Added. To assess whether personality traits are associated with beginning teacher value added, we use teachers' Education Value-Added Assessment System (EVAAS) scores estimated by the SAS Institute—the official measure of value added used for teacher evaluation in NCPS. For NCPS, there are two types of EVAAS models:

- *Multivariate response model*—a random effects model that estimates teacher value added to student achievement on the state's end-of-grade exams (Grades 3–8) for mathematics and reading
- *Univariate response model*—a hybrid random and fixed effects model that estimates teacher value added to student achievement on the state's end-of-course exams (Algebra 1, biology, and English 2), fifth- and eighth-grade science exams, and all other secondary grade courses with final exams (e.g., chemistry, geometry, U.S. history; Wright, White, Sanders, & Rivers, 2010)

For our sample, the complete list of subject areas in which teachers have EVAAS estimates is as follows: Algebra 1, Algebra 2, biology, chemistry, civics, environmental science, English 1, English 2, English 4, geometry, math,

physical science, physics, reading, science, social studies, U.S. history, and world history. The multivariate response model accounts for students clustering within teachers and for students and their peers clustering within different teachers in different years. Essentially, the teacher effect in a multivariate response model is a random effect adjusted for the impact of the past and future teachers that a student has. The univariate response model accounts for the clustering of students within teachers and incorporates 2 years of prior student test scores but no other student, classroom, or school characteristics. Here, the teacher effect is captured by a teacher-level random effect (Rose, Henry, & Lauen, 2012).

To ease interpretability of the EVAAS estimates, which are expressed in either normal curve equivalency units (multivariate response model) or scale score points (univariate response model), we standardized the EVAAS scores, within test (e.g., fourth-grade mathematics, seventh-grade reading, U.S. history) across all NCPS teachers with value-added estimates. This allows us to interpret coefficients as the association between a particular personality trait and a percentage of a *SD* in teacher effectiveness (effect size). For analyses, we made these standardized EVAAS estimates the outcome variable and ran models using all available value-added estimates across school levels. Overall, the full sample for these analyses includes 1,230 value-added estimates for 857 first-year teachers—teachers can have value-added estimates in multiple subjects (e.g., fourth-grade math and reading)—with personality trait data and EVAAS scores in the 2013–2014 school year.

Teacher Evaluation Ratings. The outcome variable for this analysis comes from the North Carolina Educator Evaluation System (NCEES), an evaluation rubric in place across NCPS in which school administrators rate teachers across five professional teaching standards: Teachers demonstrate leadership (Standard 1); teachers establish a respectful environment for a diverse group of students (Standard 2); teachers know the content that they teach (Standard 3); teachers facilitate learning for their students (Standard 4); and teachers reflect on their practice (Standard 5). To evaluate teachers, school administrators use formal classroom observations (at least three are required during the school year) and paper-based evidences to document key indicators of teaching and to provide teachers with summative ratings of *not demonstrated* (Level 1), *developing* (Level 2), *proficient* (Level 3), *accomplished* (Level 4), or *distinguished* (Level 5) on each of the five NCEES standards. Importantly, this evaluation framework captures a range of teaching practices and has multiple rating levels to better distinguish teaching practice. Furthermore, work by Henry and Guthrie (2015) indicates that these evaluation ratings are significantly correlated with teacher value added—concurrent correlations of approximately 0.20 between value-added estimates and each evaluation standard. For these analyses the dependent variable is a 1–5 ordinal value, and the sample includes the 1,579 first-year teachers who completed the M5-120 and were evaluated by a school administrator in the 2013–2014 school year.

Teacher Retention. We assess whether the personality traits of first-year teachers are significantly associated with their return to NCPS in the 2014–2015 school year. To track teacher retention, we used September 2014 salary data provided by the North Carolina Department of Public Instruction and created a dichotomous variable, where a value of 1 indicates returning to teach in NCPS and a value of 0 indicates exiting teaching in NCPS. Overall, of the 1,779 teachers in this analysis, 1,558 (87.57%) returned to NCPS in the 2014–2015 school year.

Focal Measures

To create focal measures for this analysis, we followed standard M5-120 scoring procedures and normed the responses of first-year teachers, by gender and by four age categories (10–20, 21–40, 41–60, and 61–95), against the full population of individuals who have completed the assessment (Johnson, 2014). Over 70% of our sample was part of the female, aged 21–40, norming category; another 17% of our sample was in the male, aged 21–40, category. In total, we normed these female, aged 21–40, responses against the responses of approximately 6,500 M5-120 completers; we normed the male, aged 21–40, responses against the responses of nearly 4,000 M5-120 completers.

For our sample of beginning teachers, Table 2 displays the mean values of the Big Five personality domains and the 30 personality trait facets. In comparison to the population of individuals completing the M5-120 (whose average values on the personality trait domains and facets are equal to zero), first-year teacher respondents have higher levels of extraversion, agreeableness, and conscientiousness. Conversely, first-year teacher respondents have lower levels of neuroticism and openness to experience. These values for conscientiousness and its facets are particularly notable given the research work connecting conscientiousness-related traits to outcomes of interest in education and the workplace (Barrick & Mount, 1991; Barrick et al., 2001; Rockoff et al., 2011).

In the present study, we perform two main analyses. First, we estimate models in which we include all Big Five personality trait domains—extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience—into a single regression. These models assess whether any of the Big Five personality trait domains are significantly associated with teacher value added, teacher evaluation ratings, or teacher retention. Second, we estimate a series of models in which we include the six facets for each Big Five domain—for example, for conscientiousness, we run a model controlling for self-efficacy, orderliness, dutifulness, achievement

TABLE 2
First-Year Teachers' Personality Trait Values

Big Five domains									
Extraversion		Agreeableness		Conscientiousness		Neuroticism		Openness to experience	
0.453 (0.753)		0.568 (0.775)		0.763 (0.795)		−0.590 (0.834)		−0.386 (0.909)	
Personality facets									
Friendliness	0.523 (0.749)	Trust	0.458 (0.773)	Self-efficacy	0.339 (0.809)	Anxiety	−0.172 (0.931)	Imagination	−0.680 (0.922)
Gregariousness	0.258 (0.918)	Morality	0.255 (0.742)	Orderliness	0.498 (0.854)	Anger	−0.706 (0.787)	Artistic interests	−0.023 (0.917)
Assertiveness	0.321 (0.745)	Altruism	0.503 (0.738)	Dutifulness	0.593 (0.728)	Depression	−0.753 (0.753)	Emotionality	−0.185 (0.938)
Activity level	0.570 (0.772)	Cooperation	0.493 (0.682)	Achievement seeking	0.570 (0.659)	Self-consciousness	−0.213 (0.822)	Adventurousness	−0.188 (0.892)
Excitement seeking	−0.414 (0.767)	Modesty	0.166 (0.823)	Self-discipline	0.638 (0.796)	Impulsiveness	−0.489 (0.856)	Intellect	−0.158 (0.890)
Cheerfulness	0.563 (0.689)	Sympathy	0.375 (0.868)	Cautiousness	0.535 (0.818)	Vulnerability	−0.148 (0.883)	Liberalism	−0.209 (0.988)

Note. Values are presented as *M* (*SD*) for the Big Five personality domains and the 30 personality trait facets for our sample of first-year teachers.

seeking, self-discipline, and cautiousness. We note that the use of broad personality domains, as well as the narrower facets hierarchically subordinate to each domain, poses some challenges with regard to statistical analysis. However, in personality trait research, information is valued at the domain and facet levels, with studies addressing the incremental validity of personality facets over domains alone (Griffin & Hesketh, 2004; Paunonen & Ashton, 2001; Thomas et al., 2013). Therefore, we follow established analysis procedures in psychology (Bagby, Costa, Widiger, Ryder, & Marshall, 2005; Ross, Lutz, & Bailley, 2002) and conduct separate regression analyses for personality domains and facets.

Covariates

To better isolate the associations between personality traits and beginning teacher outcomes, we include a set of school characteristics in analyses. These characteristics include school size, total per-pupil expenditures, average teacher salary supplements, short-term suspension rates, violent acts rates, the percentage of racial and ethnic minority students, and the percentage of students qualifying for subsidized school meals. These data come from the North Carolina Department of Public Instruction school report card and school expenditures files. We control for these variables because school size, resources, orderliness, and demographics have been associated with where teachers work and with measures of teacher performance and retention (Borman & Dowling, 2008; Clotfelter et al., 2005; Hanushek, Kain, & Rivkin, 2004; Lankford, Loeb, & Wyckoff, 2002; Whitehurst, Chingos, & Lindquist, 2014). Specifically, we control for school characteristics in teacher value-added analyses because school context may influence teacher effectiveness and the EVAAS model explicitly excludes measures of context (Wright et al., 2010). Since our value-added analyses combine EVAAS estimates across subject areas, we also insert a set of subject area indicators into those models. We control for school characteristics in teacher evaluation analyses given recent work showing that teachers' ratings are significantly correlated with measures of school and classroom context—even with the random assignment of teachers (Steinberg & Garrett, 2016; Whitehurst et al., 2014). Finally, we control for school context in teacher retention analyses given the connections between school characteristics (e.g., school resources, orderliness, and demographics) and teacher attrition in prior research and meta-analyses (Borman & Dowling, 2008).

Data Analysis

After administering the M5-120 to first-year teachers in March and April 2014, we took the following steps: (a) norming the beginning teacher survey responses in the summer of 2014, (b) receiving 2013–2014 teacher performance data (value added and evaluation ratings) and September 2014 pay data (for teacher retention analyses) in late 2014 and early 2015, (c) cleaning and processing these administrative

data in spring 2015, and (d) performing quantitative analyses. In what follows, we describe our value-added, evaluation rating, and retention analyses.

Teacher Value Added

To assess whether personality traits predict first-year teachers' value added to student achievement, we specified an ordinary least squares regression model. Our analyses cluster standard errors at the school level and control for a set of school covariates and a set of subject area indicators. This approach takes advantage of the full amount of variation in the EVAAS data and makes teacher effectiveness comparisons across our full sample of first-year teachers with value-added estimates (McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004). Controlling for model covariates, coefficients indicate how a 1-*SD* increase in a personality trait—one of the Big Five domains or facets—predicts first-year teachers' value-added estimates. Our preferred analyses combine across all school levels; in Appendix B, we display results from separate models in elementary, middle, and high schools.

Teacher Evaluation Ratings

Since value-added estimates do not provide information on specific aspects of teaching quality (e.g., content knowledge or classroom management) and are available only for teachers in tested grades/subject areas, we estimated the associations between personality traits and first-year teachers' evaluation ratings (Goldring et al., 2015). Here, we specified an ordered logistic regression model for each of the five NCEES standards, where the outcome variable is a first-year teacher's 1–5 evaluation score. In these models, we control for a set of school covariates and cluster standard errors at the school level. Coefficients from these models indicate how a 1-*SD* increase in a personality trait—one of the Big Five domains or facets—predicts the odds of earning higher evaluation ratings. In our ordered logistic analyses, we tested the proportional odds assumption—that the relationship between each pair of outcome groups is the same (Brant, 1990)—and found that the assumption is not violated in our sample.

Teacher Retention

Since teacher attrition has adverse effects on student achievement, is financially costly for districts and schools, and may necessitate the hiring of novice teachers to fill open positions, we assessed the relationship between personality traits and teacher retention in NCPS (Alliance for Excellent Education, 2014; Ronfeldt et al., 2013). For these analyses, we specified a logistic regression model with a set of school covariates and standard errors clustered at the school level. Coefficients from these models indicate how a 1-*SD* increase in a personality trait—one of the Big Five domains or facets—predicts the odds of returning to teach in NCPS in the 2014–2015 school year.

Results

In the present study, we report results as being statistically significant at three p value thresholds: $p < .10$, .05, and .01. Despite the increased probability of type I errors, this choice reflects our view of this research as exploratory—testing whether established associations between personality traits and job performance hold for the teaching profession. Likewise, we acknowledge that given our relatively large sample, we may detect results that are statistically but not practically significant. Therefore, we stress that our interpretation of results focuses on consistent patterns across teacher outcomes and on findings consistent with established theory/previous empirical results. In the Results and Discussion sections, we provide examples that illustrate the practical significance of our findings.

Teacher Value Added

For models across all school levels, the top panel of Table 3 shows that conscientiousness is positively associated with teacher value-added estimates, while agreeableness is negatively associated with teacher value-added estimates. On average, a 1-*SD* increase in conscientiousness is associated with an 8%-of-a-*SD* increase in teacher effectiveness; a 1-*SD* increase in agreeableness is associated with an 11%-of-a-*SD* decrease in teacher effectiveness. At the facet level, two conscientiousness facets have significant results—general self-efficacy positively predicts teacher value added, while dutifulness is negatively associated with teacher effectiveness. Two additional facets—modesty (agreeableness) and impulsiveness (neuroticism)—are also negatively associated with teacher value-added estimates.

To supplement these overall results, Tables B1 and B2 display separate value-added estimates for teachers in elementary, middle, and high schools. Like the main results, in these analyses conscientiousness is positively associated with teacher value added in elementary and high schools, while agreeableness is negatively associated with value added in elementary schools. At the facet level, several value-added results are of note: (a) Cheerfulness, cooperation, and self-efficacy positively predict teacher value added in elementary schools; (b) no facets are positively associated with teacher value added in middle schools; and (c) cautiousness and intellect positively predict value added in high schools.

Teacher Evaluation Ratings

Table 4 shows that conscientiousness is significantly associated with higher teacher evaluation ratings across all

TABLE 3
Personality Traits and Teacher Value Added (All School Levels)

Big Five model				
Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness to experience
−0.007	−0.114**	0.083+	0.014	0.010
(0.051)	(0.042)	(0.045)	(0.046)	(0.039)

Personality facet models									
Friendliness	−0.063	Trust	−0.051	Self-efficacy	0.094+	Anxiety	0.014	Imagination	−0.008
	(0.057)		(0.039)		(0.049)		(0.049)		(0.035)
Gregariousness	−0.068	Morality	0.003	Orderliness	−0.060	Anger	0.004	Artistic interests	−0.046
	(0.044)		(0.051)		(0.040)		(0.045)		(0.039)
Assertiveness	0.059	Altruism	−0.033	Dutifulness	−0.101+	Depression	0.033	Emotionality	0.008
	(0.047)		(0.057)		(0.052)		(0.051)		(0.033)
Activity level	0.008	Cooperation	0.057	Achievement seeking	0.052	Self-consciousness	−0.012	Adventurousness	−0.021
	(0.043)		(0.046)		(0.065)		(0.042)		(0.037)
Excitement seeking	0.039	Modesty	−0.107**	Self-discipline	0.042	Impulsiveness	−0.069+	Intellect	0.036
	(0.042)		(0.039)		(0.057)		(0.036)		(0.044)
Cheerfulness	0.076	Sympathy	−0.019	Cautiousness	0.033	Vulnerability	0.018	Liberalism	−0.006
	(0.049)		(0.043)		(0.040)		(0.053)		(0.030)

Note. Coefficients are expressed as a *SD* in teacher effectiveness with *SE*s in parentheses. Cases per Big Five trait, n = 1,230.
+$p < .10$. **$p < .01$.

TABLE 4
Big Five Personality Domains and Teacher Evaluation Ratings

	Leadership	Classroom environment	Content knowledge	Facilitating learning	Reflecting on practice
Extraversion	1.073 (.473)	0.987 (.878)	0.948 (.562)	0.963 (.687)	0.851+ (.087)
Agreeableness	0.933 (.420)	0.951 (.536)	0.874 (.113)	0.918 (.318)	1.040 (.662)
Conscientiousness	1.264** (.006)	1.292** (.002)	1.284** (.004)	1.262** (.006)	1.183+ (.063)
Neuroticism	0.982 (.834)	0.911 (.243)	0.988 (.893)	0.942 (.484)	0.900 (.228)
Openness	0.910 (.234)	1.024 (.727)	1.165* (.033)	1.045 (.525)	1.018 (.814)

Note. Cells report odds ratios, with p values in parentheses. Cases per rating category, n = 1,579.
+$p < .10$. *$p < .05$. **$p < .01$.

five NCEES standards—leadership, classroom environment, content knowledge, facilitating student learning, and reflecting on practice. These results are similar to the relationships between conscientiousness and subjective evaluation ratings found by Rockoff and colleagues (2011) in New York City. For facilitating student learning—a standard focused on planning appropriate instruction, using a variety of instructional methods, developing students' critical thinking skills, and assessing student learning in multiple ways—a first-year teacher with a conscientiousness value 2 *SD* below the population mean had an 18% predicted probability for rating at *developing* (Level 2) and an 8% predicted probability for rating at *accomplished* (Level 4). Conversely, a first-year teacher with a conscientiousness value 2 *SD* above the population mean had an 8% predicted probability for rating at *developing* and an 18% predicted probability for rating at *accomplished*. Additionally, openness to experience predicts significantly higher evaluation ratings for content knowledge, while extraversion is associated with significantly lower evaluation ratings for reflecting on practice.

At the individual facet level, several results are of note in Table 5. Assertiveness, an extraversion facet, predicts higher evaluation ratings for four of five teaching standards. For conscientiousness, three facets are significantly associated with higher evaluation ratings—general self-efficacy for all five standards, cautiousness for four standards, and achievement seeking for two standards. These positive self-efficacy results mirror those for teacher value added and suggest that general self-efficacy may be a facet-level predictor of teacher performance. Finally, for openness to experience, imagination predicts lower evaluation ratings for four

TABLE 5
Personality Trait Facets and Teacher Evaluation Ratings

Personality facets	Leadership	Classroom environment	Content knowledge	Facilitating learning	Reflecting on practice
Extraversion					
Friendliness	1.173 (.154)	0.995 (.963)	1.162 (.157)	1.077 (.471)	1.072 (.553)
Gregariousness	0.908 (.301)	0.944 (.502)	0.908 (.298)	0.932 (.429)	0.902 (.279)
Assertiveness	1.305** (.003)	1.179+ (.059)	1.206* (.043)	1.225* (.016)	1.150 (.142)
Activity level	1.001 (.986)	1.021 (.801)	0.971 (.747)	1.003 (.968)	1.000 (.997)
Excitement seeking	0.935 (.430)	0.939 (.438)	0.961 (.644)	0.968 (.684)	0.856+ (.079)
Cheerfulness	0.939 (.547)	1.182+ (.081)	0.929 (.472)	0.969 (.746)	1.014 (.885)
Agreeableness					
Trust	1.096 (.267)	1.171* (.045)	1.102 (.241)	1.098 (.268)	1.070 (.420)
Morality	0.922 (.406)	0.889 (.192)	0.897 (.266)	0.926 (.418)	1.000 (.998)
Altruism	0.978 (.826)	1.037 (.698)	1.102 (.333)	1.004 (.963)	1.099 (.360)
Cooperation	1.118 (.313)	1.219+ (.058)	0.966 (.771)	1.019 (.864)	1.027 (.809)
Modesty	0.889 (.144)	0.918 (.225)	0.906 (.228)	0.965 (.637)	1.022 (.776)
Sympathy	1.035 (.676)	0.965 (.629)	1.056 (.520)	1.050 (.528)	0.966 (.656)
Conscientiousness					
Self-efficacy	1.251* (.012)	1.307** (.001)	1.339** (.002)	1.227* (.025)	1.213* (.044)
Orderliness	0.961 (.637)	0.983 (.828)	0.872 (.107)	0.882 (.113)	0.970 (.734)
Dutifulness	0.878 (.186)	0.988 (.900)	0.899 (.262)	0.952 (.582)	0.935 (.515)
Achievement seeking	1.291* (.029)	1.031 (.785)	1.141 (.254)	1.208+ (.086)	1.058 (.625)
Self-discipline	0.973 (.816)	0.985 (.890)	0.995 (.965)	0.963 (.735)	0.952 (.686)
Cautiousness	1.116 (.175)	1.175* (.041)	1.164+ (.054)	1.214* (.013)	1.197* (.042)
Neuroticism					
Anxiety	1.159 (.145)	1.046 (.612)	1.024 (.816)	1.055 (.552)	1.053 (.595)
Anger	0.957 (.635)	0.837* (.046)	0.930 (.468)	0.945 (.503)	0.891 (.247)
Depression	0.878 (.187)	0.945 (.573)	1.119 (.265)	0.895 (.252)	0.992 (.941)
Self-consciousness	1.006 (.933)	1.025 (.745)	0.970 (.705)	1.002 (.973)	0.982 (.811)
Impulsiveness	0.993 (.934)	0.955 (.495)	0.934 (.349)	0.988 (.866)	0.903 (.188)
Vulnerability	0.857 (.165)	0.938 (.499)	0.915 (.424)	0.943 (.577)	0.986 (.903)
Openness to experience					
Imagination	0.845* (.012)	0.861* (.023)	0.924 (.225)	0.868* (.030)	0.859* (.033)
Artistic interests	0.952 (.501)	1.076 (.298)	0.995 (.948)	0.966 (.652)	0.934 (.387)
Emotionality	0.992 (.903)	1.019 (.750)	1.011 (.854)	1.017 (.792)	1.015 (.814)
Adventurousness	1.005 (.946)	0.922 (.257)	1.018 (.813)	0.937 (.368)	0.949 (.518)
Intellect	1.051 (.522)	1.081 (.307)	1.144 (.119)	1.195* (.025)	1.190+ (.053)
Liberalism	1.023 (.732)	1.066 (.284)	1.094 (.150)	1.076 (.236)	1.091 (.209)

Note. Cells report odds ratios, with *p* values in parentheses. Cases per rating category, $n = 1{,}579$.
$^{+}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

teaching standards, and intellect predicts higher evaluation ratings for two standards.

Teacher Retention

The top panel of Table 6 indicates that conscientiousness is significantly associated with first-year teachers returning to NCPS in the 2014–2015 academic year. To ease interpretation of results, we converted the odds ratios to predicted retention probabilities. First-year teachers with a conscientiousness value of zero (the population mean for all M5-120 completers) had an 86% predicted probability of returning to NCPS. Those with a conscientiousness value 2 *SD* below the mean had an 80% predicted probability of retention, while those with a value 2 *SD* above the mean had a 91% predicted probability of retention. At the level of individual personality trait facets, cautiousness—a conscientiousness facet that was significantly associated with four evaluation standards—positively predicted returning to NCPS; conversely, two openness to experience facets—adventurousness and imagination (which

TABLE 6
Personality Traits and Teacher Retention in North Carolina Public Schools

Big Five model									
Extraversion		Agreeableness		Conscientiousness		Neuroticism		Openness to experience	
1.076 (.530)		0.971 (.768)		1.249* (.036)		1.104 (.371)		0.976 (.799)	
Personality facet models									
Friendliness	1.053 (.736)	Trust	1.086 (.401)	Self-efficacy	1.151 (.260)	Anxiety	0.985 (.909)	Imagination	0.866+ (.079)
Gregariousness	0.879 (.284)	Morality	0.989 (.932)	Orderliness	1.022 (.826)	Anger	1.083 (.437)	Artistic Interests	1.099 (.325)
Assertiveness	0.996 (.997)	Altruism	1.193 (.180)	Dutifulness	0.931 (.537)	Depression	0.985 (.901)	Emotionality	1.030 (.736)
Activity level	1.151 (.189)	Cooperation	0.915 (.507)	Achievement seeking	1.238 (.103)	Self-consciousness	1.126 (.246)	Adventurousness	0.830* (.032)
Excitement seeking	0.902 (.329)	Modesty	0.953 (.608)	Self-discipline	0.840 (.208)	Impulsiveness	0.919 (.317)	Intellect	1.104 (.303)
Cheerfulness	1.239 (.105)	Sympathy	0.949 (.595)	Cautiousness	1.215* (.048)	Vulnerability	0.911 (.480)	Liberalism	1.033 (.673)

Note. Cells report odds ratios, with *p* values in parentheses. Cases per Big Five trait, $n = 1{,}779$.
$^{+}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

was negatively associated with four evaluation standards)—were associated with lower rates of retention in NCPS. Conceptually, these facet results make sense: Teachers whose self-reported personality traits suggest that they make considered and deliberate decisions were more likely to return, while those who prefer novelty and variety were more likely to exit.

Discussion

Our major finding is that conscientiousness, one of the Big Five personality traits, is significantly associated with higher teacher value-added estimates, higher teacher evaluation ratings across all five professional teaching standards, and higher rates of teacher retention in NCPS. These results are statistically and policy significant. For example, a 1-*SD* increase in conscientiousness is associated with an 8%-of-a-*SD* increase in teacher value added. This is more than one third of the average difference in value added between first- and second-year teachers and nearly one fourth of the average difference between first-year and midcareer (11–15 years of experience) teachers. Likewise, a 1-*SD* increase in conscientiousness is associated with a 2.5–percentage point increase in teacher retention. With 86.5% of first-year teachers returning to NCPS in 2014–2015, this conscientiousness result represents a 20% increase in retention.

At the level of personality trait facets, general self-efficacy (a subdomain of conscientiousness) significantly predicts our measures of teacher performance: value added and evaluation ratings. So why might conscientiousness matter? Conscientiousness incorporates traits related to dependability, such as being thorough, responsible, confirming, careful, and planful, as well as a set of characteristics related to motivation, including drive and persistence (Barrick & Mount, 1991). Given the structure of schools and the demands of teaching, the dependability and motivation-related traits should both be associated with job performance and retention for teachers.

These results add to the research literature on the FFM and theoretical conceptions of teaching quality. Empirically, our conscientiousness and self-efficacy findings are consistent with a rich body of research connecting the FFM to job performance and retention (Barrick & Mount, 1991; Barrick et al., 2001; Hough et al., 1990). As with many other professions, teaching outcomes are positively associated with levels of employee conscientiousness. Our unique contributions to the FFM-employment research base include an exclusive focus on public school teachers, a relatively large sample size, and the availability of multiple policy-relevant outcome measures—including two measures of teaching performance. Theoretically, our results most closely connect to initial conceptions of teacher quality that focused on individual traits as predictors of teacher performance. Despite this connection, we believe that our results make a more nuanced contribution to theories of teacher effectiveness. Conscientiousness alone does not explain teacher performance. Rather, it is likely that more

conscientious teachers engage in certain behaviors and make certain teaching decisions, and that is what contributes to their success. Future research can extend this work by examining whether there are differences in teaching practices and decision making for teachers with different levels of conscientiousness. Additionally, given recent studies connecting teachers to students' noncognitive outcomes—for example, happiness and behavior in class (Blazar & Kraft, in press), growth mind-set (Kraft & Grace, 2016), and engagement in school (Jackson, 2012)—future work needs to assess whether teacher conscientiousness is associated with a broader range of student outcomes. To the extent that teacher conscientiousness predicts student cognitive and noncognitive outcomes, this solidifies the use of conscientiousness (or other personality traits with predictive validity) as a potential measure for consequential decision making.

From a policy and practice perspective, our results, coupled with other studies on personality traits and teacher outcomes, suggest that districts and schools may want to start using personality trait measures as a way to improve their hiring decisions. Given the stability of the FFM across cultures (Schmitt et al., 2007), these results may apply equally well to hiring decisions in the United States and abroad. Toward this end, we note several concerns and guidelines for integrating personality trait data into hiring. First, regarding the ethical implications of considering personality traits in hiring decisions, research shows that this practice is fairly common in the public sector: >60% of public sector managers report using personality measures in their jobs, and >20% use them for hiring and promotion (Cooper, Knotts, Johnson, & McCord, 2012). Certain reform-minded school districts, such as Washington, DC, public schools, are already considering personality trait measures and dispositions in teacher hiring (Jacob, Rockoff, Taylor, Lindy, & Rosen, 2015). Second, the effectiveness of personality traits in hiring decisions depends on efficient, timely, and information-rich hiring practices by districts and schools. Recent research shows that many new teachers have limited interactions with school/district personnel during the hiring process, that the process is information poor, and that hiring often occurs after the start of the school year (Liu & Johnson, 2006). Thus, overlaying personality trait measures on top of the hiring process may have a limited impact without more systemic hiring changes.

Third, when personality trait measures are integrated into the hiring process, such measures should be one of a number of valid indicators that inform hiring decisions. For example, recent research shows that rubric-based hiring systems focused on a range of key teaching competencies benefit districts' hiring decisions (Goldhaber, Grout, & Huntington-Klein, 2014). Fourth, consequential hiring decisions are also improved when indicators significantly predict multiple outcomes of interest—capturing a more comprehensive measure of teacher quality (Bill and Melinda Gates Foundation, 2013). In the case of our findings, conscientiousness significantly predicted value added, evaluation ratings, and retention. Conversely, agreeableness and its facet of modesty were negatively associated with teacher value added but did not significantly predict evaluation ratings or retention. Consistency in results provides more compelling evidence that a measure should be included in consequential decision making. Fifth, while the FFM explicitly captures many personality traits, we recognize that there are other noncognitive and soft skill measures (e.g., motivation for teaching, grit, task persistence) that are significantly associated with teacher outcomes. Districts and schools integrating personality trait measures into their hiring decisions may want to explore a range of valid and reliable measures.

Finally, we acknowledge a potential limitation of this work: First-year teachers answered these personality trait items after acquiring teaching experience and in a low-stakes setting. Teacher responses and their subsequent relationships with outcomes of interest may differ before teachers acquire teaching experience and in high-stakes environments. While we would have preferred a survey administration prior to the start of teachers' careers—before the successes and challenges of classroom teaching and to better approximate information available at the time of hire—we note two points that solidify our findings: First, M5-120 survey items are general and do not ask about specific teaching attitudes; second, previous studies affirm the stability of personality trait measures across the life span of adults (Caspi, 2000; Terracciano et al., 2006). Likewise, the associations between personality traits and teacher outcomes may change over time as teachers gain experience—for example, conscientiousness may be more or less predictive of teacher performance. Future research that includes veteran teachers and/or follows cohorts of teachers over multiple years would provide a longer-term perspective (and more reliable and stable measures of teacher performance) that benefits considerations of personality traits in the hiring process.

Taken together, this study makes an important contribution by linking the FFM to teaching outcomes and showing that personality traits are positively associated with policy-relevant teacher outcomes. We encourage continued research that focuses on different samples, examines the mechanisms of personality trait findings, and evaluates whether personality traits, when used in consequential decision making, continue to predict teacher outcomes.

Appendix A

TABLE A1

Five-Factor Model of Personality

Factors: Facets	Low	High
Extraversion	Reserved, sober, aloof, retiring	Sociable, active, talkative, optimistic
E1: Friendliness	Formal, reserved, distant	Like people, close attachments
E2: Gregariousness	Loners	Enjoy company of others
E3: Assertiveness	Keep in background	Dominant, forceful, ascendant
E4: Activity level	Leisurely, relaxed in tempo	Rapid tempo, vigorous, busy
E5: Excitement seeking	Low need for thrills	Crave excitement and stimulation
E6: Cheerfulness	Less exuberant	Cheerful and optimistic
Agreeableness	Cynical, rude, uncooperative	Trusting, helpful, good-natured
A1: Trust	Skeptical, suspicious	Believe others are honest
A2: Morality	Willing to manipulate, lie	Frank, sincere, genuine
A3: Altruism	Self-centered	Generous, concerned for others
A4: Cooperation	Aggressive, competitive	Defers to others, cooperates
A5: Modesty	Arrogant, conceited	Humble, self-effacing
A6: Sympathy	Hard-hearted, realistic	Tender-minded
Conscientiousness	Careless, lazy, unreliable	Organized, reliable, hardworking
C1: Self-efficacy	Low opinion of abilities, inept	Feel well prepared, competent
C2: Orderliness	Disorganized	Neat, tidy, organized
C3: Dutifulness	Casual conscience and morality	Strictly ethical and principled
C4: Achievement striving	Lackadaisical, not driven	High aspirations and drive
C5: Self-discipline	Tend to procrastinate, quitters	Self-motivated to get job done
C6: Cautiousness	Hasty, snap decisions	Cautious and deliberate
Neuroticism	Calm, relaxed, unemotional	Worrying, nervous, emotional
N1: Anxiety	Calm, relaxed	Fearful, apprehensive, worrying
N2: Anger	Easygoing, slow to anger	Ready to experience anger
N3: Depression	Rarely experience depression	Prone to guilt, sadness, dejection
N4: Self-consciousness	Undisturbed by awkward situations	Sensitive to ridicule
N5: Impulsiveness	High tolerance for frustration	Desires are irresistible
N6: Vulnerability	Good coping	Unable to cope with stress
Openness to experience	Conventional, unartistic	Curious, broad interests
O1: Imagination	Prosaic, keep mind on task at hand	Vivid imagination, active fantasy life
O2: Artistic interests	Uninterested in art and beauty	Deep appreciation for art, poetry
O3: Emotionality	Blunted affect, low value for feelings	Experience deep, intense feelings
O4: Adventurousness	Prefer routine	Prefer novelty and variety
O5: Intellect	Narrow focus, low curiosity	Enjoy philosophical arguments
O6: Liberalism	Accept authority, tradition, conservative	Ready to reexamine values

Note. Descriptions of traits adapted from Costa and McCrae (1992).

Appendix B: Teacher Value-Added Results by School Level

TABLE B1

Big Five Personality Domains and Teacher Value Added

Domains	Elementary schools	Middle schools	High schools
Extraversion	−0.049 (0.079)	−0.058 (0.098)	−0.032 (0.073)
Agreeableness	−0.162* (0.075)	−0.023 (0.081)	−0.094 (0.059)
Conscientiousness	0.184* (0.079)	−0.123 (0.094)	0.131* (0.058)
Neuroticism	0.030 (0.078)	−0.095 (0.098)	0.021 (0.064)
Openness	0.099 (0.065)	−0.064 (0.076)	−0.027 (0.059)
Cases, *n*	478	288	461

Note. Coefficients are expressed as *SD*s in teacher effectiveness, with *SE*s in parentheses.
*$p < .05$.

TABLE B2
Personality Trait Facets and Teacher Value Added

Personality facets	Elementary schools	Middle schools	High schools
Extraversion			
Friendliness	−0.109 (0.094)	−0.079 (0.116)	0.010 (0.089)
Gregariousness	−0.143^{+} (0.080)	−0.009 (0.089)	−0.089 (0.065)
Assertiveness	0.066 (0.082)	−0.088 (0.093)	0.043 (0.070)
Activity level	0.120 (0.073)	−0.015 (0.083)	−0.053 (0.070)
Excitement seeking	0.055 (0.073)	0.063 (0.083)	−0.007 (0.063)
Cheerfulness	0.139^{+} (0.081)	0.051 (0.106)	0.094 (0.077)
Agreeableness			
Trust	−0.124^{+} (0.066)	0.032 (0.077)	−0.027 (0.064)
Morality	−0.060 (0.072)	−0.064 (0.113)	0.087 (0.084)
Altruism	−0.079 (0.095)	−0.015 (0.109)	−0.023 (0.086)
Cooperation	0.139^{+} (0.074)	0.092 (0.085)	−0.004 (0.085)
Modesty	−0.164* (0.070)	−0.037 (0.083)	−0.124* (0.055)
Sympathy	0.103 (0.074)	−0.121 (0.082)	−0.031 (0.065)
Conscientiousness			
Self-efficacy	0.146^{+} (0.081)	0.054 (0.087)	0.075 (0.084)
Orderliness	−0.024 (0.067)	0.020 (0.091)	−0.135* (0.062)
Dutifulness	−0.062 (0.075)	−0.194^{+} (0.102)	−0.152 (0.095)
Achievement seeking	0.076 (0.101)	0.041 (0.118)	0.100 (0.104)
Self-discipline	−0.023 (0.090)	−0.083 (0.116)	0.142 (0.093)
Cautiousness	0.036 (0.071)	−0.015 (0.080)	0.128* (0.058)
Neuroticism			
Anxiety	−0.039 (0.084)	−0.046 (0.108)	0.068 (0.071)
Anger	0.010 (0.081)	−0.094 (0.077)	0.059 (0.075)
Depression	0.050 (0.084)	0.137 (0.130)	−0.011 (0.071)
Self-consciousness	−0.005 (0.075)	−0.013 (0.082)	0.013 (0.066)
Impulsiveness	−0.038 (0.067)	−0.042 (0.085)	−0.108* (0.049)
Vulnerability	0.029 (0.097)	0.039 (0.094)	−0.032 (0.076)
Openness to experience			
Imagination	0.058 (0.066)	−0.026 (0.063)	−0.051 (0.051)
Artistic interests	−0.027 (0.065)	−0.080 (0.096)	−0.044 (0.055)
Emotionality	−0.009 (0.058)	0.013 (0.078)	0.027 (0.038)
Adventurousness	0.046 (0.061)	−0.023 (0.077)	−0.127* (0.055)
Intellect	−0.030 (0.064)	−0.031 (0.086)	0.175* (0.070)
Liberalism	0.034 (0.050)	0.029 (0.052)	−0.068 (0.045)
Cases, *n*	478	288	461

Note. Coefficients are expressed as *SD*s in teacher effectiveness, with *SE*s in parentheses.
$^{+}p < .10$. $^{*}p < .05$.

Acknowledgments

We thank Alisa Chapman with the University of North Carolina General Administration for her continued promotion and support of the Teacher Quality Research Initiative and the deans of the schools and colleges of education at the 15 University of North Carolina system institutions engaged in teacher education for their comments and feedback throughout the research process.

Funding

We acknowledge partial funding for this project from the Teacher Quality Research Initiative.

References

Alliance for Excellent Education. (2014). *On the path to equity: Improving the effectiveness of beginning teachers*. Retrieved from http://all4ed.org/wp-content/uploads/2014/07/PathToEquity.pdf

Bagby, R. M., Costa, P. T., Widiger, T. A., Ryder, A. G., & Marshall, M. (2005). *DSM-IV* personality disorders and the five-factor model of personality: A multi-method examination of domain-and facet-level predictions. *European Journal of Personality*, *19*(4), 307–324.

Barr, A. S. (1953). The measurement of teacher characteristics and prediction of teaching efficiency. *Review of Educational Research*, *22*(3), 169–174.

Barr, A. S., Bechdolt, B. V., Coxe, W. W., Gage, N. L., Orleans, J. S., Remmers, H. H., & Ryans, D. G. (1952). Supplement: Report of the Committee on the Criteria of Teacher Effectiveness. *Review of Educational Research*, *22*(3), 238–263.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1–26.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, *81*, 261–272.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *Personality and Performance, 9*(1–2), 9–30.

Bastian, K. C., Henry, G. T., & Thompson, C. L. (2013). Incorporating access to more effective teachers into assessments of educational resource equity. *Education Finance and Policy*, *8*(4), 560–580.

Bill and Melinda Gates Foundation. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study*. Retrieved from http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf.

Blazar, D., & Kraft, M. A. (in press). Teacher and teaching effects on students' attitudes and behaviors. *Educational Evaluation and Policy Analysis*.

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research*, *78*(3), 367–409.

Bouchard, T. J., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, *31*, 243–273.

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, *46*(4), 1171–1178.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York, NY: MacMillan.

Caspi, A. (2000). The child is father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology*, *78*, 158–172.

Chingos, M. M., & Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review*, *30*(3), 449–465.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review*, *24*(4), 377–392.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, *26*(6), 673–682.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources*, *45*(3), 655–681.

Cochran-Smith, M., & Fries, K. (2005). Researching teacher education in changing times: Politics and paradigms. In M. Cochran-Smith, & K. Zeichner (Eds.), *Studying teacher education* (pp. 69–109). Mahwah, NJ: Erlbaum.

Cooper, C. A., Carpenter, C. D., Reiner, A., & McCord, D. M. (2014). Personality and job satisfaction: Evidence from a sample of street level bureaucrats. *International Journal of Public Administration*, *37*(3), 155–162.

Cooper, C. A., Knotts, H. G., Johnson, A. J., & McCord, D. M. (2012). Taking personality seriously: The Five Factor Model and public management. *American Review of Public Administration*, *43*, 397–415.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R™) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, *64*, 21–50.

Dobbie, W. (2011). *Teacher characteristics and student achievement: Evidence from Teach for America*. Retrieved from http://blogs.edweek.org/edweek/teacherbeat/teachercharacteristics-july2011.pdf

Duckworth, A. L., Quinn, P. D., & Seligman, M. E. (2009). Positive predictors of teacher effectiveness. *Journal of Positive Psychology*, *4*(6), 540–547.

Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939–944.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*(4), 237–251.

Gage, N. L. (1964). *Psychological theory and empirical research for teacher education*. Address presented to the Associated Organizations for Teacher Education, Chicago, IL.

Gage, N. L. (1978). *The scientific basis of the art of teaching*. New York, NY: Teachers College Press.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg, Netherlands: Tilburg University Press.

Goldhaber, D., Grout, C., & Huntington-Klein. (2014). *Screen twice, cut once: Assessing the predictive validity of teacher selection tools*. Seattle, WA: Center for Education Data and Research.

Goldhaber, D., Liddle, S., & Theobald, R. (2013). The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review*, *34*(2), 29–44.

Goldring, E., Grissom, J. A., Rubin, M., Neumerski, C. M., Cannata, M., Drake, T., & Schuermann, P. (2015). Make room value added: Principals' human capital decisions and the emergence of teacher observation data. *Educational Researcher*, *44*(2), 96–104.

Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research*, *66*(3), 361–396.

Griffin, B., & Hesketh, B. (2004). Why openness to experience is not a good predictor of job performance. *International Journal of Selection and Assessment*, *12*(3), 243–251.

Grimmett, P., & Mackinnon, A. (1992). Craft knowledge and the education of teachers. In G. Grant (Ed.), *Review of research in education* (Vol. 18, pp. 385–456). Washington, DC: American Educational Research Association.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources*, *39*(2), 326–354.

Henry, G. T., & Guthrie, J. E. (2015). Using multiple measures of developmental teacher evaluation. In J. Grissom & P. Youngs (Eds.), *Improving teacher evaluation systems: Making the most of multiple measures* (pp. 143–155). New York, NY: Teachers College Press.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effects of response distortion on those validities. *Journal of Applied Psychology*, *75*, 581–595.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, *85*, 869–879.

Ingersoll, R., & Merrill, E. (2012). *Seven trends: The transformation of the teaching force* (Working Paper No. WP-01). Philadelphia, PA: Consortium for Policy Research in Education.

Jackson, C. K. (2012). *Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina* (Working Paper No. 18624). Cambridge, MA: National Bureau of Economic Research.

Jacob, B., Rockoff, J., Taylor, E., Lindy, B., & Rosen, R. (2015). *Teacher applicant hiring and teacher performance: Evidence from DC public schools*. Retrieved from https://www.newyorkfed.org/medialibrary/media/research/education_seminar_series/jrtlr_teach_dc_23_feb_2015.pdf

Johnson, J. A. (2014). Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality*, *51*, 78–89.

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, *87*, 530–541.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, *27*, 615–631.

Klassen, R. M., & Tze, V. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, *12*, 59–76.

Kraft, M. A., & Grace, S. (2016). *Teaching for tomorrow's economy? Teacher effects complex cognitive skills and social-emotional competencies*. Providence, RI: Brown University.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, *24*(1), 37–62.

Liu, E., & Johnson, S. M. (2006). New teachers' experiences of hiring: Late, rushed, and information-poor. *Educational Administration Quarterly*, *42*(3), 324–360.

Lounsbury, J. W., Sundstrom, E., Loveland, J. M., & Gibson, L. W. (2003). Intelligence, "Big Five" personality traits, and work drive as predictors of course grade. *Personality and Individual Differences*, *35*(6), 1231–1239.

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., & Trautwein, U. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*(3), 471–491.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, *29*(1), 67–101.

McCrae, R. R. (2011). Personality theories for the 21st century. *Teaching of Psychology*, *38*(3), 209–214.

McCrae, R. R., Costa, P. T., Jr., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., . . . Urbanek, T. (2004). Consensual validation of personality traits across cultures. *Journal of Research in Personality*, *38*, 179–201.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2693–2698.

Ozer, D. J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401–421.

Papay, J. P., & Kraft, M. A. (in press). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics*.

Paunonen, S., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, *81*(3), 524–539.

Robertson-Kraft, C., & Duckworth, A. L. (2014). True grit: Trait-level perseverance and passion for long-term goals predicts effectiveness and retention among novice teachers. *Teachers College Record*, *116*(3), 1–27.

Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, *6*(1), 43–74.

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

Rose, R., Henry, G. T., & Lauen, D. L. (2012). *Comparing value added models for estimating teacher effectiveness: Technical briefing*. Retrieved from https://publicpolicy.unc.edu/files/2015/07/Comparing-Value-Added-Models-for-Estimating-Teacher-Effectiveness1.pdf

Ross, S. R., Lutz, C. J., & Bailley, S. E. (2002). Positive and negative symptoms of schizotypy and the five-factor model: A domain and facet level analysis. *Journal of Personality Assessment*, *79*(1), 53–72.

Schmitt, D. P., Allik, J., McCrae, R. R., Benet-Martinez, V., Alcalay, L., Ault, L., . . . Zupancic, A. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, *38*, 173–212.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(4), 4–14.

Steinberg, M. P., & Garrett, R. (2016). Classroom composition and measured teacher performance: What do teacher observation scores really measure? *Educational Evaluation and Policy Analysis*, *38*(2), 293–317.

Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin*, *32*, 999–1009.

Thomas, K. M., Yalch, M. M., Krueger, R. F., Wright, A. G., Markon, K. E., & Hopwood, C. J. (2013). The convergent structure of *DSM-5* personality trait facets and five-factor model trait domains. *Assessment*, *20*(3), 308–311.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89–122.

Whitehurst, G. J., Chingos, M. M., & Lindquist, K. M. (2014). *Evaluating teachers with classroom observations: Lessons learned in four districts*. Retrieved from the Brookings Institute website: http://www.brookings.edu/~/media/research/files/reports/2014/05/13-teacher-evaluation/evaluating-teachers-with-classroom-observations.pdf

Wright, S. P., White, J. T., Sanders, W. L., & Rivers, J. C. (2010). *SAS EVAAS statistical models*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.437.6615&rep=rep1&type=pdf

Authors

KEVIN C. BASTIAN is a research associate and the director of the Teacher Quality Research Initiative at the Education Policy Initiative at Carolina at the University of North Carolina at Chapel Hill; kbastian@email.unc.edu. His research focuses on educator preparation and effectiveness.

DAVID M. MCCORD is a professor in the department of psychology and in the College of Education and Allied Professions at Western Carolina University; mccord@wcu.edu. His research focuses on the five-factor model of personality and relationships between personality traits and characteristics and behaviors.

JULIE T. MARKS is the director for the Education Policy Initiative at Carolina at the University of North Carolina at Chapel Hill; jtmarks@email.unc.edu. Her research focuses on educational evaluation and school turnaround.

DALE CARPENTER is the dean of the College of Education and Allied Professions at Western Carolina University; carpenter@wcu.edu. His research focuses on teacher preparaton, personality traits, and special education.