



# Do Teachers' Personality Traits Predict Their Performance? A Comprehensive Review of the Empirical Literature From 1990 to 2018

ETS RR–19-04

Harrison J. Kell

*December 2019*



Discover this journal online at  
**Wiley Online Library**  
wileyonlinelibrary.com

# ETS Research Report Series

---

## EIGNOR EXECUTIVE EDITOR

James Carlson  
*Principal Psychometrician*

## ASSOCIATE EDITORS

Beata Beigman Klebanov  
*Senior Research Scientist*

Heather Buzick  
*Senior Research Scientist*

Brent Bridgeman  
*Distinguished Presidential Appointee*

Keelan Evanini  
*Research Director*

Marna Golub-Smith  
*Principal Psychometrician*

Shelby Haberman  
*Consultant*

Priya Kannan  
*Managing Research Scientist*

Sooyeon Kim  
*Principal Psychometrician*

Anastassia Loukina  
*Research Scientist*

John Mazzeo  
*Distinguished Presidential Appointee*

Donald Powers  
*Principal Research Scientist*

Gautam Puhan  
*Principal Psychometrician*

John Sabatini  
*Managing Principal Research Scientist*

Elizabeth Stone  
*Research Scientist*

Rebecca Zwick  
*Distinguished Presidential Appointee*

## PRODUCTION EDITORS

Kim Fryer  
*Manager, Editing Services*

Ariela Katz  
*Proofreader*

Ayleen Gontz  
*Senior Editor*

---

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

## RESEARCH REPORT

# Do Teachers' Personality Traits Predict Their Performance? A Comprehensive Review of the Empirical Literature From 1990 to 2018

Harrison J. Kell

Educational Testing Service, Princeton, NJ

Interest in the relationship between teachers' personality traits and their performance is rising. It is often remarked, however, that the topic has received little research attention. This study surveys what empirical research has been conducted among U.S. K–12 teachers from 1990 to 2018. The search, conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and structured by the Big Five personality taxonomy, uncovered 13 articles. Effect sizes were statistically small, with the most consistent finding being that agreeableness was negatively related to most outcomes across most studies. Implications and future directions for research and practice are discussed.

**Keywords** Personality traits; Big Five personality traits; teacher effectiveness; teacher performance; teacher quality; teacher personality

doi:10.1002/ets2.12241

Teachers play a critical role in the modern world. Teachers can be a positive influence on a wide variety of their students' short- and long-term outcomes, including their grades, state assessment scores, health, extracurricular activities, college attendance, adult income, and retirement savings (Chamberlain, 2013; Chetty, Friedman, & Rockoff, 2014). Given teachers' pervasive impact, considerable effort has been expended trying to understand the characteristics that differentiate effective from ineffective teachers,<sup>1</sup> as teachers differ in the extent to which they have a positive effect on student outcomes (Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). Research into these characteristics serves the important purpose of helping to promote basic understanding of how such characteristics are related to effective teaching practices—but it also informs practical needs, such as identifying variables that might help eventually improve teacher hiring policies.

Unfortunately, the characteristics usually investigated (e.g., certification status, standardized test scores) account for only a small amount of the difference between effective and ineffective teachers, and it is often not feasible to hire based on candidates' experience, which is the most powerful predictor of teacher performance (Rockoff, Jacob, Kane, & Staiger, 2011). Consequently, some have begun to call for the investigation of “nontraditional” predictors of teachers' effectiveness, with the hope that heretofore underresearched variables may be capable of further explaining differences in teachers' performance (Bastian, 2013; Rockoff et al., 2011). If such characteristics can be reliably identified and they are both reasonably stable and present prior to entry into the hiring process, they could be viable for consideration during that hiring process.

Personality traits constitute one domain whose promise for predicting teacher performance is currently being investigated. Some in the United States (e.g., Kennedy, 2012; Rose, English, & Finney, 2014; Thornton, Peltier, & Hill, 2005) advocate not only for the study of the relationship between teachers' performance and their personality traits but also suggest that personality traits potentially be given explicit consideration during the teacher selection and hiring process. Despite this recent interest for exploring—and capitalizing on—the relationship between teachers' personality traits and their effectiveness, it is often remarked that current knowledge about that relationship is limited (e.g., Robertson-Kraft & Duckworth, 2014). The purpose of this report is to fill this gap in the literature by comprehensively surveying what is known about the relationship between teachers' personality traits and their performance. We first review how personality traits are conceptualized in the contemporary psychology literature, along with the practical outcomes they are associated with. Next, we review the variety of ways teacher effectiveness is measured, along with historical concerns about studying

*Corresponding author:* H. Kell, E-mail: hkell@ets.org

how it is related to personality. We then describe our literature review strategy and highlight its most important results. The final section of the paper discusses the implications of these findings and makes recommendations for future research and practice based on the relationship between teachers' personality traits and their performance.

## Personality Psychology: An Overview

### Personality Traits

Personality traits are “[r]elatively stable individual differences in consistent patterns of behavior” (Matsumoto, 2009, p. 549). Evidence for human beings' interest in personality traits is voluminous and can be found in documents produced in ancient China, Greece, Israel, and Rome (Bowman, 1989; Dumont, 2010; Revelle, 2008; Vandenberg, 1956). The ancient physicians Hippocrates and Galen believed personality to be rooted in biology, and traces of their theories remain in the form of adjectives still in use in everyday speech (e.g., *melancholic*, *sanguine*) (Dumont, 2010). These conceptions of personality found their way into medieval Muslim culture through the work of Avicenna, who elaborated them in terms of mental abilities, attitudes, dreams, and self-awareness (Fountoulakis, 2015). In the 19th century, phrenologists attempted to localize personality characteristics in specific areas of people's brains, with these loci indicated by bumps on their skulls (Butcher, 2010). The more structured, scientific study of personality traits using a wide variety of techniques (e.g., interviews, perceptual tests, rating scales) was inaugurated in the early 20th century (Fernald, 1912; Heymans & Wiersma, 1906) and continues to the present day (see Barenbaum & Winter, 2008; Dumont, 2010; R. Hogan, 1991; McAdams, 1997).

The term *trait* comes with some unhelpful connotations (Guion, 1987). Well-respected educational psychologist Richard Snow noted, “The term ‘trait’ was borrowed by early theorists from biology, where it still often refers to hereditary, permanent, and physically based characteristics. It is a treacherous term” (Snow, Kyllonen, & Marshalek, 1984, p. 53). Nobel laureate James Heckman remarked that the term conveys “a sense of immutability or permanence” (Heckman & Kautz, 2014, p. 6) and some popular sources explicitly treat traits as being “natural born” (Gregoire, 2014). Such perceptions of personality traits are at odds with the nuanced accounts that dominate contemporary personality psychology.

Contrary to William James's (1890, p. 126) intuition that personality traits eventually “set like plaster, and will never soften again,” they can be altered by major life experiences (J. J. Jackson, Thoemmes, Jonkmann, Lüdtke, & Trautwein, 2012; Lüdtke, Roberts, Trautwein, & Nagy, 2011; Roberts, Caspi, & Moffitt, 2003; Specht, Egloff, & Schmukle, 2011) and are not perfectly consistent across the life span (Roberts & DelVecchio, 2000; Roberts, Walton, & Viechtbauer, 2006). There is also emerging evidence that personality can be altered through intervention, both clinical (Roberts et al., 2017; Tang et al., 2009) and nonclinical (Hudson & Fraley, 2015; J. J. Jackson, Hill, Payne, Roberts, & Stine-Morrow, 2012; Magidson, Roberts, Collado-Rodriguez, & Lejuez, 2014). Consequently, when considering potential associations between personality traits and teaching effectiveness, it is critical to keep in mind that the traits in question are not immutable, innate, or inborn.

### The Big Five Personality Traits

By the 1960s, personality research across educational (Ross & Stanley, 1954), occupational (Gibby & Zickar, 2008), and laboratory (Cattell, 1957) settings had produced thousands of inventories (R. Hogan, 1998) and identified hundreds of disparate traits (for only a partial list, see Goldberg, 1971, pp. 296–303). The field lacked an agreed-upon framework to organize and integrate these findings. The absence of an overarching taxonomy had many deleterious effects, including promotion of a reliance on ad hoc and atheoretical measures (Block, 1977), pursuit of fads rather than concerted lines of research (Sechrest, 1976), inattention to the appropriateness and dimensionality of the criteria to be predicted (Barrick, Mount, & Judge, 2001; Hough & Schneider, 1995), and the splitting of broad traits into progressively narrower constructs with often scant practical relevance (Goldberg, 1971). These problems eventually undermined the entire discipline and led some to doubt the ability of personality traits to predict outcomes of interest to the world at large (e.g., Guion & Gottier, 1965; Kendrick, 1964; Mischel, 1968; Peterson, 1968; Vernon, 1964). This period has been called “the dark days of personality testing” (Zickar & Kostek, 2013, p. 182).

The 1980s and early 1990s witnessed the development of the taxonomy of personality traits the field had lacked. Research findings consistently coalesced around five major personality dimensions, with converging evidence for the Big

Five accumulating across multiple sources, including self-ratings using adjectives and preexisting questionnaires (Digman & Inouye, 1986; Digman & Takemoto-Chock, 1981; Goldberg, 1990; McCrae & Costa, 1985; Noller, Law, & Comrey, 1987), knowledgeable observers' ratings (McCrae & Costa, 1987), strangers' ratings (Borkenau & Liebler, 1992a, 1992b, 1993a, 1993b), and interpersonal interactions (Wiggins & Trapnell, 1997). Reviews of the academic personality literature and retrospective analyses that have adopted the Big Five perspective (or the five-factor model; McCrae & Costa, 2008) have found confirming evidence for the presence of these traits in studies dating from the 1910s to the 1940s (Deary, 1996; Fiske, 1949; Goldberg, 1990, 1995; Thurstone, 1934; Webb, 1915). Each of the Big Five traits is briefly described below. For comprehensive treatments, see chapters dedicated to each trait in John, Robins, and Pervin (2008), Leary and Hoyle (2009), and Widiger (2017a).

- **Agreeableness:** Adjectives that describe highly agreeable people include warm, kind, and cooperative; those for highly disagreeable people include cold, selfish, and distrustful. Agreeableness is positively associated with tendencies to use negotiation (versus retaliation) to resolve conflicts and engage in helping behaviors and negatively associated with aggression, prejudice, and competitiveness. Highly agreeable people tend to effectively regulate the frustration that sometimes arises during interpersonal interactions, experience empathic concern when they observe people in distress, and be motivated to maintain harmonious relations with others.
- **Conscientiousness:** Adjectives that describe highly conscientious people include thorough, hardworking, and responsible; those for highly unconscientious people include lazy, careless, and negligent. Conscientiousness is related to many practically important variables: It is positively associated with longevity, educational attainment, job performance, and marital stability and negatively associated with criminality, smoking, and unemployment. Highly conscientious people tend to follow social norms, feel guilt and shame when they fail to meet others' expectations, and be able to delay gratification in order to achieve long-term goals.
- **Extraversion:** Adjectives that describe highly extraverted people include talkative, bold, and energetic; those for highly introverted people include timid, unadventurous, and inactive. Extraversion is positively associated with positive emotionality, number of mates over the lifetime, and mortality and negatively associated with depression, anxiety, and feelings of insecurity. Highly extraverted people tend to strive for interdependence and intimacy, create positive social environments in the course of their interactions with others, and be biased toward attending to positive stimuli.
- **Neuroticism:** Adjectives that describe highly neurotic people include nervous, discontented, and tense; those for highly emotionally stable people include relaxed, at ease, and calm. Neuroticism is positively associated with cardiovascular disease, alcohol abuse, and the presence of many types of psychopathology (e.g., eating disorders, schizophrenia) and negatively associated with self-efficacy, subjective well-being, and relationship satisfaction. Highly neurotic people tend to feel self-conscious and insecure, act impulsively when upset, and be prone to finding minor frustrations emotionally overwhelming.
- **Openness to Experience:** Adjectives that describe people scoring high on openness include imaginative, creative, and curious; those for people scoring low on openness include unsophisticated, unreflective, and uninquisitive. Openness is positively associated with appreciating art, divergent thinking, and political liberalism and negatively associated with right-wing authoritarianism, racism, prejudice, and religiosity. Highly open people tend to seek out novelty and originality, be competent in recognizing others' emotions, and be primarily attracted to other highly open people.

### The Practical Importance of the Big Five

The settling of consensus on the Big Five had enormous implications. It allowed for the development of theoretically informed personality assessments, which tend to demonstrate more validity than measures constructed without being based on overarching theory (Block, 1977; R. Hogan, 1991; McCrae & Costa, 1985; Salgado, 2003). It led to a greater focus on the constructs that personality assessments measure rather than on the inconsistent labels promiscuously applied to them (Barrick et al., 2001; Ghiselli, 1973). And it led to the theoretically informed classification and clustering of seemingly disparate traits, which revealed practically important relationships with criteria that did not appear when only extremely narrow dimensions were considered (R. Hogan, 1991; Hough & Schneider, 1995).

Evidence for the practical importance of the Big Five rapidly accumulated in the wake of these developments. In particular, 1991 was a watershed year (Oswald & Hough, 2011; Zickar & Kostek, 2013), marking the publication of two meta-analyses (Barrick & Mount, 1991; Tett, Jackson, & Rothstein, 1991) presenting powerful evidence for the ability of personality traits to predict job performance. Since then, the Big Five have been reliably linked to many other important social phenomena, including career success (Ng, Eby, Sorensen, & Feldman, 2005), job satisfaction (Judge, Heller, & Mount, 2002), leadership (Judge, Bono, Ilies, & Gerhardt, 2002), academic performance in high school and college (Poropat, 2009; Richardson, Abraham, & Bond, 2012; Schneider & Preckel, 2017), and health variables such as smoking, alcohol intake, and longevity (Bogg & Roberts, 2004; Friedman & Kern, 2014; Goodwin & Friedman, 2006; Ozer & Benet-Martínez, 2006; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). Associations between workers' personality traits and their job performance have been documented across a very broad array of occupations, ranging from police officer to salesperson to lawyer (Barrick & Mount, 1991; Huang, Ryan, Zabel, & Palmer, 2014; Hurtz & Donovan, 2000; Shultz & Zedeck, 2011). The systematic accumulation of evidence supporting the practical importance of personality helped stimulate interest in "noncognitive skills" more broadly, which has grown rapidly in the last 15 years in psychology (e.g., Duckworth & Seligman, 2005), education (Duckworth & Yeager, 2015; Klassen et al., 2018) and economics (Heckman, Humphries, & Kautz, 2014).

### The Challenge of Measuring Teachers' Effectiveness and Its Implications

Effectiveness in many jobs can be evaluated adequately by asking supervisors to rate their subordinates on one or two performance dimensions (Motowidlo & Kell, 2013; Viswesvaran, Schmidt, & Ones, 2005). In contrast, it is widely agreed that the performance domain of K–12 teachers is complex, and teachers' effectiveness can vary across many different dimensions (Darling-Hammond, 2010; Harris & Rutledge, 2010). As a consequence, multiple assessments are required in order to adequately appraise the short-term effect teachers have on students. The quality of teachers' practices is often directly assessed by classroom observations conducted by principals or other experts (S. L. Campbell & Ronfeldt, 2018; Harris & Sass, 2014) and indirectly by examining teachers' effects on students' outcomes (Harris & Rutledge, 2010). Most often, these latter effects are indexed by academic achievement measures, but there have been recent efforts to measure the influence of teachers on students' noncognitive characteristics and outcomes as well (e.g., absences, suspensions; C. K. Jackson, 2012). Ratings-based protocols are also sometimes employed, wherein principals or other supervisory figures evaluate teachers based on their proficiency in specific areas (e.g., organizational skill, skill in teaching a subject area) and their overall effectiveness, or both, but based on that figure's overall impression of the teacher rather than behavior during a specific classroom-based episode (Harris, Ingle, & Rutledge, 2014). Student surveys are also sometimes used when appraising teachers' performance, along with records of the teachers' own tardiness and absenteeism (Cheng & Zamarro, 2016; Jacob, Rockoff, Taylor, Lindy, & Rosen, 2016). Although not explicitly treated as part of teachers' effectiveness, their retention–attrition is also sometimes considered, as turnover and teacher experience are related to important outcomes (e.g., student achievement; Ronfeldt, Loeb, & Wyckoff, 2013), and the cost of hiring and training new teachers is substantial (Barnes, Crowe, & Schaefer, 2007).

The relationships among evaluations of different facets of teachers' performance (e.g., student noncognitive skills, value-added) are complex. Little research has been conducted into how consistent teachers' performance is across different measures of effectiveness (Harris & Sass, 2014), and what research has been conducted suggests the relationships between different measures of teachers' effectiveness are relatively weak (Harris et al., 2014). For example, teachers who are successful in increasing students' value-added scores tend *not* to be the same teachers who are successful in enhancing students' noncognitive skills, and vice versa (Blazar & Kraft, 2017; Gershenson, 2016). The multidimensional nature of teachers' performance, in addition to the complex interrelations among its components, means that any predictor of teachers' effectiveness can only be properly evaluated when its association with the many different aspects of teachers' performance is taken into account; some personality traits may be related to one element of effectiveness (e.g., absenteeism, increasing students' noncognitive skills) but not others (e.g., retention, value-added). Further, some traits could be positively related to one aspect of teachers' performance but negatively related to a second. Consequently, a comprehensive review of the literature examining the association between teachers' personality traits and their effectiveness necessitates taking into account both the full scope of teachers' performance and a broad range of personality traits.

## Teacher Effectiveness and Personality Traits

As evidenced by article titles such as “Personality Tests and Teaching Ability” (Tyler, 1949), “What Are the Personality Traits of the Successful Teacher?” (Dodge, 1943), and “An Analysis of the Personality Traits of the Effective Teacher” (Witty, 1947), interest in the relationship between teachers' personalities and their performance was widespread in the first half of the 20th century. Nonetheless, in the first edition of the *Handbook of Research on Teaching*, Getzels and Jackson (1963) concluded their seminal chapter reviewing the relationship between teachers' personality characteristics and performance with the following statement:

Despite the critical importance of the problem and a half-century of prodigious research effort, very little is known for certain about the nature and measurement of teacher personality, or about the relation between teacher personality and teaching effectiveness. The regrettable fact is that many of the studies so far have not produced significant results. Many others have produced only pedestrian findings. (p. 574)

This chapter has been credited with effectively ending research into the association between teachers' personality traits and important outcomes (G. Sykes, personal communication, February 2, 2017). This assertion is borne out by recent statements such as “There is further little evidence that differences in [teachers'] personality characteristics are associated with differences in student achievement” (Klette, 2007, p. 151), and “Though the quest has been pursued by many ... the identification of a particular ‘teacher personality’ has eluded researchers and theorists” (Pajak, 2012, p. 1197), and “Very little is known about the role of person-level qualities, or personality, in the teacher labor market” (B. K. Jones, 2016, p. 1).

Getzels and Jackson (1963) identified the major reason their review failed to identify recurrent patterns in the association between teachers' personalities and their effectiveness as the fact that the studies they reviewed were conducted in a “theoretical vacuum” (p. 575). They noted that sound theory provides a taxonomic framework for generating research designs, interpreting research findings in a meaningful way, guiding a long-range research agenda for investigating a phenomenon, and ultimately venturing theoretical explanations for observations. Further, failure to situate research within a prevailing theory can lead to the ad hoc selection of research instruments—especially problematic in the personality domain, where the authors noted the number of measures available was “legion” (Getzels & Jackson, 1963, p. 574). Indeed, they specifically noted that just because two measures are labeled as personality tests does not mean the data they generate are necessarily comparable. These criticisms are highly similar to those leveled at the field of personality in general in the 1960s and 1970s. As noted previously, the Big Five paradigm has largely been successful in addressing these criticisms, providing a taxonomic framework for developing theoretically informed personality measures, classifying and interpreting the results of preexisting personality studies that studied seemingly disparate traits, and selecting appropriate personality measures for predicting chosen criteria based on proper theoretical alignment.

A second major weakness Getzels and Jackson (1963) noted in the literature they reviewed was the failure to appropriately, and consistently, define and measure teacher effectiveness. As indicated previously, this is a problem that still bedevils the field (Harris et al., 2014; Harris & Rutledge, 2010), with discrepancies across investigations' findings sometimes being attributable to differences in the definition and assessment of effectiveness (e.g., classroom observations vs. value-added scores; C. K. Jackson, Rockoff, & Staiger, 2014). Failure to devote proper attention to criterion variables and choose personality traits likely to predict them based on underlying theory was also identified as a reason for the low confidence in the practical importance of personality traits in the 1960s and 1970s (Barrick et al., 2001; Hough & Schneider, 1995).

The “solutions” to the problems identified by Getzels and Jackson (1963) in the personality–teacher effectiveness area generated in personality psychology in the 1980s and 1990s have largely not penetrated the K–12 teacher research space. In turn, most researchers and practitioners have seemingly—and understandably, given their perception of the state of the field—not studied the topic. Cheng and Zamarro (2016) attributed the lack of findings to the fact that data about teachers' personality characteristics are usually simply not collected, making conducting even retrospective analyses difficult.

## The Current Investigation

This study aims to provide a comprehensive review<sup>2</sup> of the empirical research literature in the K–12 teacher field that examines the association between personality traits and variables indicative of teachers' effectiveness, broadly defined. It

is specifically interested in examining literature in a circumscribed sample of potential data sources. The nature of the articles is described in the following, along with rationales for why these criteria for inclusion have been selected.

### Time Period

This review is limited to studies conducted from 1990 to the present. We limited the review to this period for several reasons. First, by 1990, personality psychology had begun to coalesce around the Big Five taxonomy (Digman, 1990), and 1 year later, the two first meta-analyses (Barrick & Mount, 1991; Tett et al., 1991) of the association between personality traits and job performance were published. Second, this period encompasses the enactment of No Child Left Behind in the United States, which has been described as “arguably the most significant federal intervention into education in the United States over the last 40 years” (C. K. Jackson, 2012, p. 3). Finally, the rationale for not reviewing studies conducted prior to 1990 is that conclusions drawn from data gathered across too broad a period could court overgeneralization across time. Eminent educational psychologist Lee Cronbach (1975, pp. 122–123) once stated, “[g]eneralizations decay. At one time a conclusion describes the existing situation well; at a later time it accounts for rather little variance, and ultimately it is valid only as history.” This philosophical perspective is practically echoed by sentiments about validity decay by industrial/organizational (I/O) psychologists (e.g., Tenopyr, 1999), who caution individuals not to assume that the association between personnel selection devices and job performance is invariant over long periods of time. Plonsky and Oswald (2014) stated, more bluntly, that there is no reason to assume that effect sizes will be stable over long periods of time.

### Sample Composition

The behavioral sciences have been criticized for overgeneralizing findings based primarily on participants drawn from Western, educated, industrialized, rich, and democratic (WEIRD) populations (e.g., Arnett, 2008; D. Jones, 2010; Henrich, Heine, & Norenzayan, 2010a, 2010b); overgeneralizing results can have negative practical and theoretical consequences (Greenwald, Pratkanis, Leippe, & Baumgardner, 1986; Simons, Shoda, & Lindsay, 2017). To avoid making inappropriately broad inferences across locations and contexts about the association between personality traits and teacher effectiveness, this review is purposely limited in scope to in-service classroom teachers working in the kindergarten through 12th-grade domain in the United States. We have purposely excluded preservice teachers and teacher candidates, as it is possible that people scoring particularly high or low on various personality traits may not move from these stages of their careers into the teaching profession itself. We also do not include studies with college instructors: These individuals likely constitute a distinct population relative to K–12 teachers, as the path to this profession differs greatly from that followed by prospective K–12 teachers, and many college instructors are less focused on teaching than other aspects of their careers (e.g., research).

### Sample Size

Small sample sizes can produce misleading estimates of effects in the population of interest, with samples of over 1,000 recommended for generating highly accurate estimates of population effect sizes (Maxwell, Kelley, & Rausch, 2008). Simulation studies suggest that correlation coefficients do not stabilize until sample sizes equal or exceed approximately 150 data points (Schönbrodt & Perugini, 2013), and a sample size of 150 is considered reasonable when validating predictors for personnel selection (Oh & Roth, 2017). Similarly, a minimum sample size equal to or greater than  $104 + m$  (where  $m$  is the number of predictors) is recommended when testing individual predictors in multiple regression (equivalent to a bivariate correlation) (Green, 1991; Wilson Van Voorhis & Morgan, 2007); the lower limit for sample size should be 105, given the minimum number of predictors is 1. Consequently, to avoid including small studies whose results might contribute misleading effect size estimates, we include only studies with sample sizes of 100 or greater, dipping slightly below the minimum recommended sample size of 105 to potentially include studies whose investigators may have used the round number 100 as a rule of thumb for determining their sample size.

### Nature of the Criterion

As discussed in the “Introduction,” defining and operationalizing teacher performance, effectiveness, productivity, or quality is challenging, and there are multiple models for conceptualizing how teachers contribute positively to student



and organizational outcomes (Harris & Rutledge, 2010), many of which treat the domain as multidimensional (Cheng & Zamarro, 2016; Harris et al., 2014). Teacher effectiveness has been indexed in many ways, including student value-added measures (Ballou & Springer, 2015), tardiness and absenteeism (Jacob et al., 2016), student surveys (Bacher-Hicks, Chin, Kane, & Staiger, 2017), classroom observations (Steinberg & Garrett, 2016), and influence on students' noncognitive skills (Cheng & Zamarro, 2016). Although not explicitly considered as part of effectiveness, teacher attrition is also considered an outcome worthy of study and has been shown to be related to some measures of teacher performance (e.g., student achievement; Ronfeldt et al., 2013) and has been implicated in creating unstable learning environments (Redding & Henry, 2018).

As a consequence of the difficulties of defining and measuring teacher performance, we adopt the very broad definition of *effectiveness* offered by Harris and Rutledge (2010, p. 917): "the degree to which workers produce outcomes related to the objectives of their organizations." As applied to teachers, this broad definition allows for the inclusion of objective (e.g., absenteeism, value-added scores) and subjective (e.g., principals' ratings, student surveys) measures of teacher effectiveness, in addition to variables that may not necessarily be construed as direct indicators of teachers' performance (e.g., turnover). Especially in this exploratory stage, it is necessary to study multiple criteria, as personality traits may be associated with some aspects of teachers' effectiveness but not others. Indeed, the full promise of personality traits for predicting job performance was not realized until the criterion space was expanded to include motivational and interpersonal variables (Borman & Motowidlo, 1993; Hertz & Donovan, 2000).

Not included as viable criteria are self-reported variables—from effectiveness itself to more distinctly psychological variables such as burnout, satisfaction, and self-efficacy. Common method bias (D. T. Campbell & Fiske, 1959) artificially inflates the observed association between personality traits and scores on other self-report measures. Further, although variables such as burnout and self-efficacy may influence teacher effectiveness, they do not constitute effectiveness itself.

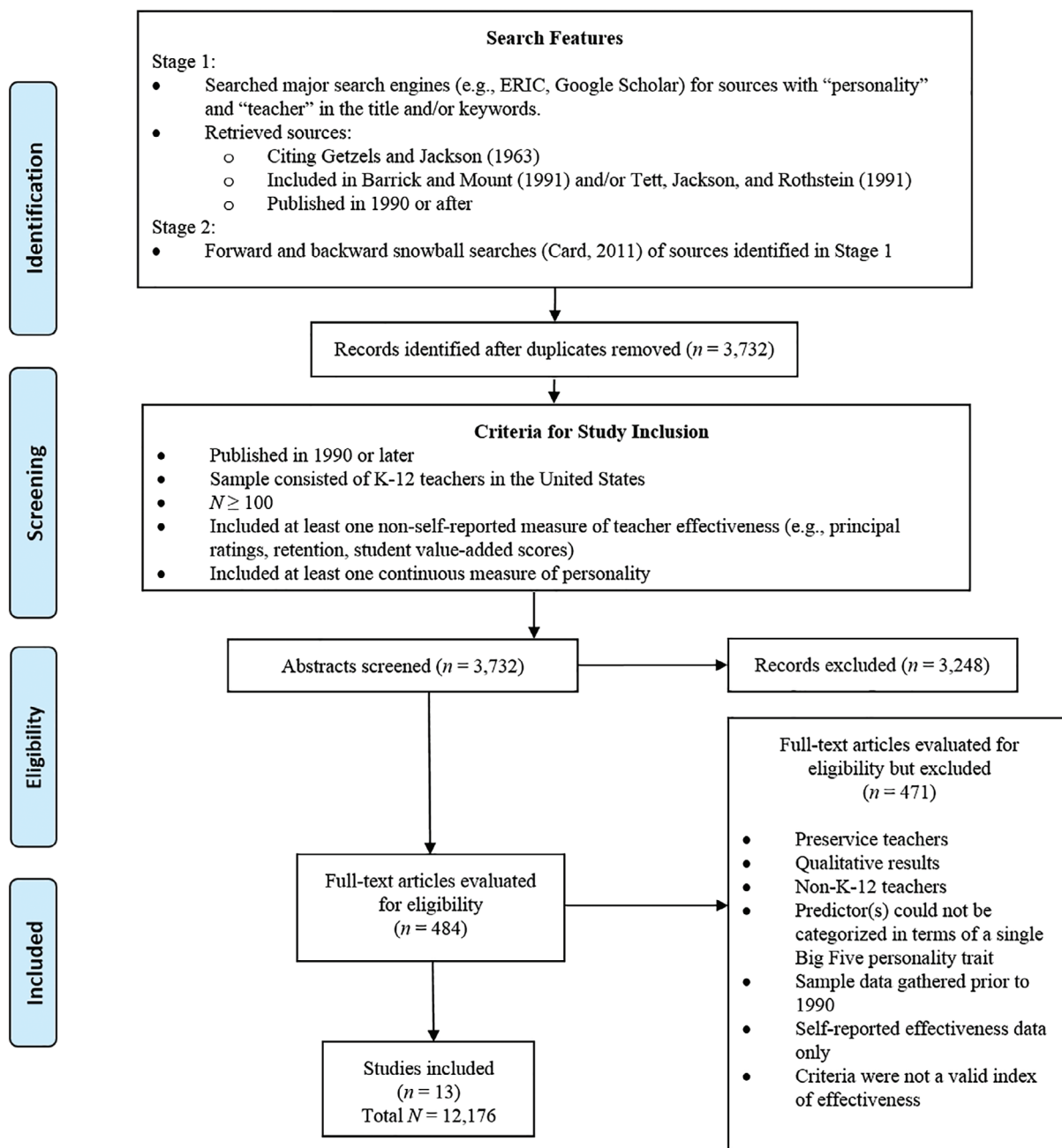
### **Nature of the Predictor: Continuous Traits**

Although the trait-based Big Five approach to personality is currently dominant within personality psychology, outside personality psychology it is arguably more common to think of personality in terms of discrete categories or "types" that people can be classified into. Some of the earliest conceptualizations of personality (e.g., Galen's temperaments, Hippocrates' humors) are typological, and the tendency to sort individuals into discrete personality types is evident in the stock characters of Greek, Roman, and Italian theater and the descriptions of people based on their astrological birth signs (Eysenck, 1997; Revelle, Wilt, & Condon, 2011). In the 20th century, Spranger, Kretschmer, Lombroso, and C. G. Jung created prominent personality typologies, with Jung's serving as the basis for the Myers-Briggs Type Indicator (R. Hogan, 1991; Kagan, 2005). Despite the long-standing and intuitive appeal of personality types, modern empirical research has generally found that continuous personality traits are a more informative means of describing people. Although some personality types can be decomposed into individuals' profiles on the Big Five traits (McCrae & Costa, 1989), the dichotomization needed to derive typological scores entails a loss of information about individuals' standings across the entire trait continuum (Pittenger, 2004) and can distort experimental and correlational results (cf. MacCallum, Zhang, Preacher, & Rucker, 2002). Accordingly, typological scores cannot be straightforwardly transformed into scores on the Big Five. Personality types have also been shown to offer no information beyond the Big Five, be difficult to replicate, and not clearly distinguish individuals across categories (Costa & McCrae, 2017). Given the empirical and theoretical limitations of the typological approach to personality, this review focuses solely on measures of continuous personality traits.

## **Method**

### **Literature Search**

We conducted the literature search according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, PRISMA Group, 2009). The PRISMA flowchart comprises Figure 1 and documents how we progressively narrowed the search results to 13 final studies. Due to long-standing concerns about statistical significance and its (mis)interpretation (Wasserstein & Lazar, 2016), all effect sizes were summarized, regardless of their significance.



**Figure 1** Literature search process depicted in terms of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram. Adapted from “Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement” by D. Moher, A. Liberati, J. Tetzlaff, & D. G. Altman, D. G. (2009) *PLoS Medicine*, 6(7) 267. Copyright 2009 Moher et al.

We divided the search process into two stages. In Stage 1, the research team used five electronic databases (ERIC, Google Scholar, JSTOR, PsycINFO, Web of Knowledge) to search for sources that included the terms *teacher* and *personality* in their titles or keywords or cited the seminal Getzels and Jackson article (1963). We also identified articles meta-analyzed by Barrick and Mount (1991) and/or Tett et al. (1991) that included teachers (the first two meta-analyses of personality traits and job performance). For all three searches we restricted results to studies published from 1990 to 2018.

Two researchers reviewed the abstracts and titles of data sources identified by these searches; if at least one researcher believed the title and/or abstract indicated the source could yield data relevant to the review, the full text of that source was retrieved. We used these sources as a basis for Stage 2 of the literature search, which constituted a forward and

backward snowball search (Card, 2011). To identify additional potentially relevant studies, for each of the full-text data sources retrieved in Stage 1, we examined their references sections and conducted Google Scholar citation searches. Two researchers reviewed the titles and abstracts of the potential sources identified by the snowball search. If at least one researcher flagged the title and/or abstract as indicative of a source that could contain data relevant to the association between personality traits and teacher effectiveness, the full text of that source was retrieved. To these full-text sources, we applied the selection criteria for sample composition, time period, continuous trait measure, and criterion nature described previously. This search process ultimately yielded 13 investigations meeting the inclusion criteria.

### Sorting Personality Traits Into the Big Five

Some of the 13 studies identified their predictors in terms of the Big Five personality traits, but not all. For those studies whose authors did not explicitly align their personality variables with the Big Five, three PhD-holding psychologists with extensive knowledge of the personality literature and over 5 years of experience with applied and basic research read definitions of those constructs and categorized them in terms of one of the Big Five traits or a sixth category (unclear or belongs to two or more Big Five domains). A construct was retained for the literature review only if the majority of coders classified it in terms of one of the Big Five traits. This process led to the exclusion of seven constructs investigated in the 13 studies reviewed.

## Results

Tables 1 and 2 summarize the major characteristics of the 13 studies in terms of the Big Five personality traits assessed and teacher effectiveness measures used. The mean sample size across the 13 investigations was 671 ( $SD = 745$ ), with a range of 118 to 2,671. Given that this review purposely excludes studies with  $N$ s less than 100, the standard deviation underestimates the variability of the literature reviewed, as some studies uncovered during the search had  $N$ s below 20. Samples were highly variable in terms of the grades in which teachers taught, the subject(s) they taught, and the regions of the country in which they worked. In some cases, no information about these sample characteristics was provided. Four samples comprised Teach for America (TFA) participants, and a fifth consisted solely of teachers seeking certification from the National Board for Professional Teaching Standards. As these samples are likely unrepresentative of the larger teacher population, the results of the studies they feature in may lack generalizability.

Table 2 summarizes information about the personality traits in the studies reviewed. The most commonly assessed personality trait was conscientiousness (92% of studies), followed by extraversion (54%). Less than a quarter (23%) of the studies assessed all the Big Five traits. Personality was primarily assessed by self-report (62%), followed by observers' ratings (23%); no studies measured personality using multiple methods. In the early stages of research, the consistency of the direction of effect sizes is more important than the magnitude of those effects (Cook, Campbell, & Peracchio, 1990), and totals for positive associations are noted accordingly. For four of the Big Five traits, the majority of the effect sizes are positive, although for extraversion and emotional stability, that majority is small. Only one third of the effect sizes for agreeableness are positive.

Table 3 includes a column for zero-order effect sizes because many studies reported only effect sizes (e.g., standardized beta weights, odds ratios) that controlled for various covariates (e.g., gender, ethnicity, prior achievement, school characteristics). Moreover, studies differed in the variables they controlled for and did not report the results of incremental validity analyses (e.g.,  $\Delta R^2$ ) examining the contribution of personality variables above and beyond more typically used predictors of teachers' performance. Consequently, it is appropriate to summarize and compare only the results of studies reporting zero-order effects (i.e., not controlling for covariates). Of the 38 zero-order effect sizes, 79% fell between  $-.10$  to  $.10$  (as indexed by a correlation coefficient). Except for one medium effect ( $r = .31$ , conscientiousness & retention), all effect sizes were small when interpreted according to traditional benchmarks (Cohen, 1992).

Table 4 summarizes details of the measurement of teachers' effectiveness. Teachers' performance was assessed in a wide variety of ways, including value-added scores, observers' ratings, retention, and students' noncognitive skills. The majority of studies (62%) used multiple methods or metrics to index teachers' effectiveness.

Interest in the relationship between teachers' personality traits and their effectiveness is apparently increasing. Only a single study meeting the inclusion criteria was published in the 1990s, as was the case between 2000 and 2005 and from 2006 and 2010. From 2011 to 2015, six studies were published, and from 2016 and 2018, four studies were published.

**Table 1** Summary of Major Characteristics of Investigations of the Association Between Teachers' Personality Traits and Effectiveness

Study	Big Five trait(s)	Effectiveness outcome(s)	Effect size	Results
Cutchin (1999)	A, C, E, N, O	Student evaluations of two behavioral aspects of teacher performance; two principal ratings	Correlation coefficient ( <i>r</i> )	Skill/rapport (student) $A = -.03$ ; $C = .02$ ; $E = .06$ ; $N = -.03$ ; $O = .04$ Assignments/materials (student): $A = -.01$ ; $C = .12$ ; $E = -.01$ ; $N = -.05$ ; $O = .02$ Overall effectiveness (principal): $A = -.13$ ; $C = .12$ ; $E = -.07$ ; $N = .10$ ; $O = -.05$ Global rating (principal): $A = -.13$ ; $C = .10$ ; $E = -.04$ ; $N = .06$ ; $O = .02$ Performance: $A = -.09$ ; $C = .01$ ; $E = .00$ ; $N = .05$ ; $O = .03$ Effectiveness: 1.31
Emmerich, Rock, and Trapani (2006)	A, C, E, N, O	Teacher performance (composite of 10 NBPTS exercises)	Correlation coefficient ( <i>r</i> )	Mathematics achievement: .02 English achievement: -.01
Duckworth, Quinn, and Seligman (2009)	C	Effectiveness (ratings by TFA administrators)	Odds ratio	Math achievement: $C = .02$ ; $E = .001$
Dobbie (2011)	C	Student achievement in mathematics and English	Standardized regression coefficient ( $\beta$ )	Classroom observations: $C = .20$ ; $E = .22$
Rockoff et al. (2011)	C, E	Students' math achievement; classroom observations by teachers' mentors; teacher absences; returned to an NYC school; returned to the same NYC school	Standardized regression coefficient ( $\beta$ )	Teacher absences: $C = .05$ ; $E = .03$ Returned to NYC: $C = -.003$ ; $E = -.002$ Returned to same school: $C = .01$ ; $E = .03$
Maurer (2012)	C, E, N	Median rate of MAP; percentage typical growth of MAP; principal rating	Correlation coefficient ( <i>r</i> )	Median rate of growth: $C = .07$ ; $E = .07$ ; $N = -.03$ Percent making typical growth: $C = .01$ ; $E = .02$ ; $N = .01$
Bastian (2013)	C	Students' test scores on end-of-grade math and reading exams (elementary school) or STEM and non-STEM end-of-course exams (high school); principals' ratings of five competencies	Standardized regression coefficient ( $\beta$ )	Principal rating: $C = .09$ ; $E = .17$ ; $N = -.07$ Elementary math grades: .01 Elementary reading grades: .02 High school end-of-course exams: .01 STEM subjects: .01 Non-STEM subjects: .01 <i>Principal ratings:</i> Leadership: 1.11 Respectful environment: 1.16 Content knowledge: 1.18 Facilitate learning: 1.08 Reflect on practice: .83

Table 1 Continued

Study	Big Five trait(s)	Effectiveness outcome(s)	Effect size	Results
Harris et al. (2014)	A, C, E	Inclusion in high or low category of overall effectiveness according to principals' ratings; inclusion in high or low category of students' value-added scores	Difference between mean personality scores of teachers identified as the most effective and least effective according to principals' evaluations and value-added measures	Math (principal evaluations): A = 2.11; C = 2.87; E = 2.68 Math (value-added): A = .49; C = .55; E = .64; Reading (principal evaluations): A = 1.96; C = 2.75; E = 2.53 Reading (value-added): A = .06; C = .23; E = .36
Robertson-Kraft and Duckworth (2014)	C	Study 1: One-year retention; effectiveness (rated by regional supervisors) Study 2: Effectiveness (rated by regional supervisors)	Correlation coefficient ( <i>r</i> )	Study 1: Retention: .31 Effectiveness: .20
Cheng and Zamarro (2016)	C	Teacher-level outcomes: Classroom observations based on videotapes of classroom lessons; students' ratings of teaching effectiveness; principals' overall evaluations of teachers	Standardized regression coefficient ( $\beta$ )	Study 2: Effectiveness: .17 Teacher-level outcomes: Classroom observations: -.10 Student ratings: -.09 Principal ratings: -.03 Student-level outcomes: Math value-added: .01 English value-added: .04 Math test scores: -.03 English test scores: -.01 Student grit: -.01 Student effort: .03 Student conscientiousness (survey effort): .03 Job performance: -.02
Jacob et al. (2016)	E	Job performance (combination of value-added, classroom observations, principals' ratings)	Standardized regression coefficient ( $\beta$ )	Commitment: .00
B. K. Jones (2016)	C	Commitment to teaching (number of years as a teacher and years plan to continue teaching)	Correlation coefficient ( <i>r</i> )	Value-added: A = .89; C = 1.07; E = .99; N = 1.01; O = 1.01
Bastian et al. (2017)	A, C, E, N, O	Student achievement value-added; NCEES teacher evaluation ratings; one-year retention	Odds ratio	Evaluations: A = .94; C = 1.23; E = .96; N = .94; O = 1.03 Retention: A = .97; C = 1.25; E = 1.08; N = 1.1; O = .98

Note. A = Agreeableness; C = Conscientiousness; E = Extraversion; G = Grit; MAP = Median average proficiency; NBPTS = National Board for Professional Teaching Standards; N = Neuroticism; NYC = New York City; NCEES = North Carolina Educator Effectiveness System; O = Openness to experience; STEM = Science, technology, engineering, mathematics; TFA = Teach for America.

**Table 2** Summary of Sample Characteristics, Sample Size, and Personality Assessment Method for 13 Studies Retrieved by Literature Review

Study	Sample characteristics	N	Personality assessment method
Cutchin (1999)	High school teachers in Indiana	118 – 126	Self-report
Emmerich et al. (2006)	Primary and secondary school teachers seeking certification from National Board for Professional Teaching Standards (NBPTS)	2,671	Self-report
Duckworth et al. (2009)	First- and second-year Teach for America (TFA) teachers	390	Self-report
Dobbie (2011)	First-year TFA teachers in New York City, students in third through eighth grades	267 – 303	Ratings from TFA applicant interview
Rockoff et al. (2011)	First-year math teachers in New York City fourth through eighth grades	1,117 – 1,608	Self-report
Maurer (2012)	NHA teachers with less than 6 years of teaching experience	313	Self-report
Bastian (2013)	Elementary and high school North Carolina TFA teachers	249	Ratings from TFA applicant interview
Harris et al. (2014)	K – 12 teachers from a mid-sized Florida school district	279	Principal ratings
Robertson-Kraft and Duckworth (2014)	First- and second-year teachers	154 (Study 1), 307 (Study 2)	Coded from work experience and college activities listed on résumés
Cheng and Zamorro (2016)	Teachers participating in the Measuring of Effective Teaching Project; teachers were randomly assigned to teach in fourth through ninth-grade classrooms	813 – 1,852	Survey effort
Jacob et al. (2016)	Washington, DC public school teachers hired in 2011 through the Teach DC process	292 – 314	Self-report
B. K. Jones (2016)	Graduates of teacher licensure and master's program	118	Self-report
Bastian, McCord, Marks, and Carpenter (2017)	First-year North Carolina public school teachers	1,230 – 1,779	Self-report

## Discussion

Our comprehensive review of the empirical literature uncovered only 13 studies examining the association between teachers' personality traits and performance that also met our inclusion criteria. Based on the findings of the studies summarized, along with the relevant knowledge and practices that have accumulated in the personality and I/O psychology literatures over the past several decades, we explore the following implications of our review and set forth several recommendations for future research and practice.

### Small Effects in Perspective

Nearly all the effect sizes observed were small according to conventional guidelines (Cohen, 1992). The effects were somewhat smaller than those found in the I/O psychology literature when they are the result of cumulating findings across many jobs (Barrick & Mount, 1991; Huang et al., 2014) but comparable to or larger than the effect sizes observed for

**Table 3** Summary of Study Characteristics Related to Personality Traits

Personality trait	<i>N</i> studies	Total effect sizes	Total zero-order effect sizes	<i>N</i> zero-order positive associations with criteria
Agreeableness	4	12	5	0
Conscientiousness	13	61	14	12
Extraversion	7	21	8	4
Neuroticism	3	8	8	3
Openness	3	8	5	4

**Table 4** Summary of Study Characteristics Related to Teacher Effectiveness

Criterion	<i>N</i> studies	Total effect sizes	Total zero-order effect sizes	<i>N</i> zero-order positive associations with personality
Supervisor ratings	8	37	20	12
Student ratings	2	11	10	5
Classroom observations	2	3	0	–
Retention	4	12	0	–
Student achievement	7	35	9	7
Absences	1	2	0	–
Other	3	9	5	3

some specific jobs (e.g., salespeople, skilled & semiskilled laborers; Hurtz & Donovan, 2000). It is important to realize that guidelines for interpreting the magnitude of statistical associations are somewhat arbitrary, and *small* does not mean *trivial* (Cohen, 1992). Further, small statistical effects can cumulate over time and, as a result, have major consequences (Abelson, 1985; Roberts et al., 2007).

Only a small percentage of effect sizes were interpretable in terms of traditional guidelines, as the majority of studies reported results only after controlling for many covariates. Future research would benefit from reporting zero-order associations between personality traits and effectiveness in order for results across studies to be more directly comparable. These analyses should be accompanied by subsequent computations that control for the influence of important covariates, however, as the relations between performance and personality traits may be greater or smaller after accounting for other associations.

### Multiple Traits, Multiple Criteria, Multiple Methods

As noted previously, the teacher performance domain is multidimensional, and teachers who are highly effective in one domain are not necessarily those who are highly effective in others (C. K. Jackson, 2012). This was reflected in the studies reviewed, as evidenced by the multiple ways in which teacher performance was measured and how personality traits were differentially related to performance across dimensions. For instance, extraversion was shown to be positively associated with growth of students' test scores (Maurer, 2012) but negatively associated with principals' overall effectiveness ratings (Cutchin, 1999). The only trait that demonstrated highly consistent associations with most outcomes was agreeableness—and the sign of these (zero-order) associations was always negative. This finding suggests some principals' intuitions regarding the appropriate characteristics to look for when hiring new teachers may be incorrect, as one of the characteristics that principals believe to be most important among teachers is *caring* (Engel & Finch, 2015; Harris, Rutledge, Ingle, & Thompson, 2010)—a key element of agreeableness. Given the differential associations among different traits and different dimensions of teachers' performance observed in the review, the promise of using personality traits to predict and understand teachers' effectiveness can only be met if all the Big Five are being measured, along with multiple elements of the teacher quality domain.

It is important for multiple dimensions of teachers' performance to be assessed in future investigations, as the sign and magnitude of the relationships between personality traits and effectiveness observed will be influenced by what elements of teacher performance are assessed; researchers and practitioners must be precise in delineating how they define effectiveness in their studies, as different definitions will lead to different operationalizations in terms of measurement and,

consequently, associations with predictors. In many of the studies we reviewed, teacher effectiveness was clearly conceptualized as multidimensional (e.g., Cheng & Zamarro, 2016; Harris et al., 2014), and accordingly, researchers used multiple indices of teacher effectiveness.

The majority of the studies reviewed assessed teachers' performance using multiple methods, but no study examined personality traits using multiple methods; future research would benefit from doing so. Personality traits manifest across a wide variety of assessments, including self-report questionnaires, interviews, observer ratings, survey effort measures, and life history records. Research in personality and I/O psychology has consistently demonstrated that others' ratings of personality traits are better predictors of many criteria than self-reports (Oh, Wang, & Mount, 2011), supporting their continued use in linking teachers' personality traits to various indicators of their effectiveness. In several of the studies reviewed (e.g., Bastian, 2013; Harris et al., 2010), observers' measures of personality traits were most strongly related to *other* ratings-based indices of effectiveness and only weakly related to "objective" indicators, such as students' test scores and value-added measures. This discrepancy underscores the fact that it is important to consider not only which traits are being selected as predictors, but also the method with which they are assessed *and* how teachers' effectiveness is being defined.

Finally, it may be worthwhile to examine the association between personality traits and outcomes that are not necessarily within the domain of teacher effectiveness but nonetheless are practically important. For example, Bastian, McCord, Marks, and Carpenter (2015) found that people higher on openness to experience were more likely to choose to work in high-need schools (e.g., high-poverty, high-minority, low-performing). Similarly, it is possible that teachers' personality traits may be related to their preferences for teaching different subjects or students of different ages, working in urban versus rural school districts, and engaging in culturally responsive practices.

### Measure the Basic Big Five Personality Traits Explicitly

or featured personality constructs that could be classified into one of the Big Five dimensions with a relatively high degree of agreement. Due to this strict inclusion criterion, studies that investigated the relationship between what were ambiguous "blends" of the Big Five and effectiveness were excluded, as were some effect sizes in several studies that were reviewed. For example, Goldhaber, Grout, and Huntington-Klein (2014) examined the association between teachers' effectiveness and constructs including *interpersonal skills*, *cultural competency*, and *classroom management*—all constructs that are surely related to the Big Five traits but are not straightforwardly interpretable in terms of single Big Five dimensions.<sup>3</sup> Similarly, the TFA dimensions *respect* ("holds high expectations for individuals in low-income communities," Dobbie, 2011, p. 6) and *motivational ability* ("an individual's ability to use interpersonal skills to motivate and lead others toward a common goal," Bastian, 2013, p. 12) bear resemblance to openness and conscientiousness but are not directly categorizable in terms of either broad trait. Commercial teacher selection instruments, such as the Teacher Perceiver Interview (TPI) and Haberman Star Teacher Pre-Screener, also assess attributes that appear to be specifiable in terms of mixtures of the Big Five (e.g., fallibility, gestalt, listening, organization and planning; Hartlep, Hansen, & Horn, 2015; Metzger & Wu, 2008).

The practical utility of compound traits—aggregates of the more "basic" Big Five—is well established (Hough & Schneider, 1995), but they can be difficult to interpret theoretically. Further, if measures of different traits are combined inappropriately, the correlation between the compound trait and criterion may be smaller than if separate analyses were conducted at the level of the Big Five (Paunonen & Ashton, 2001; Vinchur, Schippmann, Switzer, & Roth, 1998). Consequently, future research would benefit from discretely measuring each of the Big Five and reporting zero-order associations between those traits and teachers' performance, even if the personality scales are later empirically combined to maximize their predictive validity (cf. Hough & Schneider, 1995). Additionally, investigating the associations between well-established measures or constructs that are associated with teachers' effectiveness (e.g., Haberman Star Teacher Pre-Screener, TFA selection criteria, TPI) and the Big Five could provide information that would aid in interpreting prior research employing those established measures and constructs.

### Increase Measurement of the Big Five in Low-Stakes Teacher Research

That researchers have given only a small amount of attention to the relationship between personality and teachers' effectiveness has been partially attributed to the fact that personality is rarely assessed in large-scale studies of the teaching



profession (Cheng & Zamarro, 2016). It would be relatively easy to begin the process of rectifying this oversight, given that personality traits can be quickly and efficiently assessed in low-stakes settings using self-report inventories and that many high-quality, theoretically grounded measures of personality traits are readily accessible. The International Personality Item Pool (IPIP; Goldberg et al., 2006) is a free online resource that features dozens of personality scales with detailed accompanying technical information. One, or several, measures of the Big Five derived from the IPIP could be incorporated into large-scale surveys of the teaching profession, as they often take only minutes for participants to complete.

Inclusion of such measures in low-stakes teacher research could generate many important findings beyond simply allowing for the large-scale investigation of the relationship between teachers' personality characteristics and various indices of their effectiveness: Doing so would substantially enhance understanding of potential differences among the central tendencies and variances of teachers' personality characteristics across subject areas, grade levels, school districts (e.g., high versus low poverty), and regions of the country, in addition to allowing more precise estimation of the association of the Big Five with teachers' demographic characteristics. It would also allow for the estimation of associations between personality and more traditional predictors of teachers' performance (e.g., certification, college GPA, SAT<sup>®</sup> score). Adding self-report personality surveys to large-scale studies of the teaching profession could produce results with important theoretical and practical implications and would cost little to add to large-scale surveys in terms of survey length, development time, and money. If those results were sufficiently promising, they would also provide an argument for eventually examining teachers' personality traits using other, more effort-intensive methods such as interviews and observer-based ratings.

### **Exercise Caution If Basing High-Stakes Decisions on Personality Scores**

When there are few to no consequences contingent on teachers' responses to personality questionnaires, as when they are completing surveys, presumably they will respond to those questionnaires honestly and their self-ratings will be valid. When the stakes are high, however, and responses to a self-report personality survey may have major ramifications (e.g., being hired, promoted, terminated), teachers will have a strong incentive to present themselves in the best light, likely leading many to distort their responses. Concerns about personality questionnaire respondents answering dishonestly under high-stakes conditions have existed since nearly the inception of the self-report method (Ellis, 1946) and were a deciding factor in the College Board's decision not to adopt personality measures for college admissions purposes in the 1960s (College Entrance Examination Board Committee on Research, 1963; Kendrick, 1964).

The debate about the extent to which faking impacts the validity of self-report personality questionnaires in general, and specifically those administered in the context of personnel selection, is vigorous and ongoing (e.g., J. Hogan, Barrett, & Hogan, 2007; McGrath, Mitchell, Kim, & Hough, 2010; Paunonen & LeBel, 2012). Even those who do not believe that response distortion degrades the validity of personality surveys acknowledge that faking can cause shifts in job applicants' rank ordering (e.g., Mueller-Hanson, Heggstad, & Thornton, 2003; Rosse, Stecher, Miller, & Levin, 1998), potentially leading to the ethically untenable outcome of different applicants being hired than would have been the case if faking had not occurred (Morgeson et al., 2007). Moreover, there is some evidence that individuals who exhibit a higher likelihood of distorting their responses to personality surveys tend to score lower on morality, honesty, and humility and be more prone to engaging in deviant behaviors in the workplace (O'Neill et al., 2013). School principals and other administrators should be aware of these controversies before considering the administration of self-report personality inventories to prospective teacher hires. If school personnel desire to base their hiring decisions explicitly on candidates' personality traits, they should consider alternative measures that are not as amenable to response distortion, such as biodata forms (Harold, McFarland, & Weekley, 2006), situational judgment tests (Lievens et al., 2018), and structured interviews developed with the intention of eliciting the personality constructs of interest (Van Iddekinge, Raymark, & Roth, 2005).

Ironically, despite these cautions, personality traits are likely already playing an important role in who becomes a teacher. When hiring prospective teachers, principals favor candidates who score relatively high on a variety of personality traits (e.g., caring [agreeableness], creativity [openness], enthusiasm [extraversion], and motivation [conscientiousness]; Engel & Finch, 2015; Harris et al., 2010). The great majority of teacher hiring practices entail a job interview (Delli, 2010; Rutledge, Harris, Thompson, & Ingle, 2008), and the Big Five traits can be inferred from behaviors exhibited during job interviews (Van Dam, 2003; Van Iddekinge et al., 2005). Principals and school officials are

likely partially basing their hiring decisions about teacher candidates, albeit implicitly, on those candidates' personality traits.

### Resist Premature Causal Inferences

As noted in the introduction, the term *trait* comes with historical baggage and may summon inappropriate connotations of immutability and innateness (Guion, 1987; Snow et al., 1984). Further, there are differences in the extent to which scientists and practitioners make causal inferences based on the same data; see Tryon (1979) and responses. Given there is already evidence that some favor the perspective of teachers being “born rather than made” (Scott & Dinham, 2008), finding recurrent associations between teachers' personality traits and their effectiveness may quickly lead some to infer that the degree to which teachers are likely to be effective is predetermined and inborn. This tendency should be resisted. There is no distinct boundary differentiating innate and acquired characteristics (Bateson & Mameli, 2007), and the simple dichotomy of nature versus nurture was rejected long ago (Gould, 1996); current trait models acknowledge the enormous complexity of the processes and their interplay, from which individuals' personality traits emerge (e.g., Roberts, 2018). Further, at this early stage of research, the most appropriate course is to detect the presence of phenomena, explore and describe them, and establish the extent to which they exhibit regularity (Hacking, 1983; Rozin, 2001, 2009)—and under what conditions (Cacioppo, Semin, & Berntson, 2004). Only when, and if, regular statistical relations between teachers' personality traits and their effectiveness are consistently observed will it be appropriate to move forward with making strong causal inferences (cf. Roberts et al., 2007). The very fact that scientists and philosophers have grappled with the notion of causality for thousands of years (Gower, 1997) should serve as a warning that making causal inferences is no simple matter.

### Conclusion

There is rising interest in studying noncognitive variables that are observable prior to hiring and that predict teachers' effectiveness (Bastian et al., 2017; Corcoran & O'Flaherty, 2016); one such suite of variables is personality traits (Kylloinen, Lipnevich, Burrus, & Roberts, 2014). Given that this interest is relatively newfound—or, more properly, growing (cf. Getzels & Jackson, 1963)—researchers have acknowledged that the extent of the relationship between teachers' personality traits and their effectiveness is largely unknown (Robertson-Kraft & Duckworth, 2014). This review endeavored to fill this gap in the research literature by summarizing the empirical literature from 1990 to 2018 examining the association. We organized findings according to the Big Five personality trait taxonomy, the predominant model used by personality psychologists (Widiger, 2017b).

Given that only 13 studies were suitable for review, and with substantial variation among them in how teacher effectiveness was operationalized, it is premature to draw strong conclusions about the direction and magnitude of the associations between teachers' personality traits and their performance. However, judging solely by the zero-order relationships observed across these studies, the associations appear to be small by statistical standards but also highly variable, with associations potentially being moderated by which personality traits are measured, how they are measured, and how teacher effectiveness is indexed. Due to this complex interplay, care must be exercised in designing studies in this area, as inappropriate consideration of these factors could lead to over- or underestimates of the relationship between teachers' personalities and their effectiveness in a given research or practice setting.

There have been recent calls to purposefully consider prospective teachers' personalities during the hiring process, both in the United States (Kennedy, 2012; Rose et al., 2014) and other countries (e.g., Australia, United Kingdom; Bowles, Hattie, Dinham, Scull, & Clinton, 2014; Kim, Dar-Nimrod, & MacCann, 2017; Klassen & Tze, 2014; Klassen et al., 2018). In many countries, for those in favor of selecting teachers based on their personality traits, the Big Five is a viable taxonomy for organizing those characteristics, as investigations across over 50 societies support the robust cross-cultural replicability of the Big Five (McCrae, 2002; McCrae, Terracciano, & 79 Members of the Personality Profiles of Cultures Project, 2005; Schmitt et al., 2007), and they have also proven to be predictive of job performance in many different jobs in many different nations (Salgado, 1997). Nonetheless, we purposely limited our review to investigations featuring K–12 teachers in the United States to avoid inappropriately generalizing our findings to other countries. Every nation has a unique educational system, and separate reviews and investigations of the predictive validity of the Big Five for teachers' performance should

be conducted in whatever country they are being considered for use, especially in light of the variability of the effect sizes observed in the studies conducted within the United States.

The results of this review suggest it is unlikely that considering teachers' personalities during hiring or placement processes would be a revolutionary educational system reform that would substantially decrease teacher absences, increase the retention of qualified teachers, and improve student outcomes. Nonetheless, the findings of the studies summarized, coupled with the robust evidence for the real-world significance of personality traits in psychology and economics, suggest that continuing to investigate how teachers' personality traits are related to their performance will advance the field practically, as well as theoretically.

### Acknowledgments

I thank Natalie DeAngelo, Jennifer Klafehn, and Michelle Martin-Raugh for their assistance with this paper and Heather Buzick, David Klieger, and Gary Sykes for their helpful comments on earlier versions of this manuscript.

### Notes

- 1 In this paper, we use the terms *performance*, *quality*, *effectiveness*, and *productivity* interchangeably to represent the enumeration of teachers' short-term impact on the school-related outcomes of a particular group(s) of students at a given time and in a specific context.
- 2 Klassen and Tze (2014) conducted a meta-analysis of the association between teachers' aggregate personalities and their performance, finding a correlation of .08 ( $k = 10$ ). Aggregating all personality traits to form a single composite can provide a useful "first glance" at their ability to predict teachers' performance, but doing so can also obscure differential relations between personality traits and teacher performance. Moreover, many theoretical objections to the existence of a "general factor" of personality (which such a composite represents) have been lodged (e.g., Revelle & Wilt, 2013). Further, due to our inclusion criteria (e.g., date of publication, sample size), only two of the studies summarized in the current review overlap with those meta-analyzed by Klassen and Tze, meaning the two reviews make largely independent contributions to the literature.
- 3 That complex constructs such as *motivational ability*, *classroom management*, and *interpersonal skills* are not directly reducible to the Big Five personality traits supports the idea that there is no single "effective teacher personality" (cf. Hines, 2010): Equal scores on such complex constructs can presumably be attained through different combinations of the Big Five traits.

### References

- Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, *97*, 129–133. <https://doi.org/10.1037/0033-2909.97.1.129>
- Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist*, *63*, 602–614. <https://doi.org/10.1037/0003-066X.63.7.602>
- Bacher-Hicks, A., Chin, M. J., Kane, T. J., & Staiger, D. O. (2017). *An evaluation of bias in three measures of teacher quality: Value-added, classroom observations, and student surveys* (NBER Working Paper No. 23478). Cambridge, MA: National Bureau of Economic Research.
- Ballou, D., & Springer, M. G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher*, *44*, 77–86. <https://doi.org/10.3102/0013189X15574904>
- Barenbaum, N. B., & Winter, D. G. (2008). History of modern personality theory and research. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 3–26). New York, NY: Guilford Press.
- Barnes, G., Crowe, E., & Schaefer, B. (2007). *The cost of teacher turnover in five school districts: A pilot study*. Washington, DC: National Commission on Teaching and America's Future.
- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*, 1–26. <https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- 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? *International Journal of Selection and Assessment*, *9*, 9–30. <https://doi.org/10.1111/1468-2389.00160>
- Bastian, K. C. (2013, March). *Do teachers' non-cognitive skills and traits predict effectiveness and evaluation ratings?* Paper presented at the Association for Education Finance and Policy 38th Annual Conference, New Orleans, LA.
- Bastian, K. C., McCord, D. M., Marks, J. T., & Carpenter, D. (2015). *Do personality traits impact beginning teacher performance and persistence?* Chapel Hill: University of North Carolina at Chapel Hill.

- Bastian, K. C., McCord, D. M., Marks, J. T., & Carpenter, D. (2017). A temperament for teaching? Associations between personality traits and beginning teacher performance and retention. *AERA Open*, 3, 1–17. <https://doi.org/10.1177/2332858416684764>
- Bateson, P., & Marnell, M. (2007). The innate and the acquired: Useful clusters or a residual distinction from folk biology? *Developmental Psychobiology*, 49, 818–831. <https://doi.org/10.1002/dev.20277>
- Blazar, D., & Kraft, M. A. (2017). Teacher and teaching effects on students' attitudes and behaviors. *Educational Evaluation and Policy Analysis*, 39, 146–170. <https://doi.org/10.3102/0162373716670260>
- Block, J. (1977). Advancing the psychology of personality: Paradigmatic shift or improving the quality of research? In D. Magnusson & N. S. Endler (Eds.), *Personality at the cross-roads: Current issues in interactional psychology* (pp. 37–63). Hillsdale, NJ: Erlbaum.
- Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, 130, 887–919. <https://doi.org/10.1037/0033-2909.130.6.887>
- Borkenau, P., & Liebler, A. (1992a). The cross-modal consistency of personality: Inferring strangers' traits from visual or acoustic information. *Journal of Research in Personality*, 26, 183–204. [https://doi.org/10.1016/0092-6566\(92\)90053-7](https://doi.org/10.1016/0092-6566(92)90053-7)
- Borkenau, P., & Liebler, A. (1992b). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, 62, 645–657. <https://doi.org/10.1037/0022-3514.62.4.645>
- Borkenau, P., & Liebler, A. (1993a). Consensus and self-other agreement for trait inferences from minimal information. *Journal of Personality*, 61, 477–496. <https://doi.org/10.1111/j.1467-6494.1993.tb00779.x>
- Borkenau, P., & Liebler, A. (1993b). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology*, 65, 546–553. <https://doi.org/10.1037/0022-3514.65.3.546>
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98). San Francisco, CA: Jossey-Bass.
- Bowles, T., Hattie, J., Dinham, S., Scull, J., & Clinton, J. (2014). Proposing a comprehensive model for identifying teaching candidates. *The Australian Educational Researcher*, 41, 365–380. <https://doi.org/10.1007/s13384-014-0146-z>
- Bowman, M. L. (1989). Testing individual differences in ancient China. *American Psychologist*, 44, 576–578. <https://doi.org/10.1037/0003-066X.44.3.576.b>
- Butcher, J. N. (2010). Personality assessment from the nineteenth to the early twenty-first century: Past achievements and contemporary challenges. *Annual Review of Clinical Psychology*, 6, 1–20. <https://doi.org/10.1146/annurev.clinpsy.121208.131420>
- Cacioppo, J. T., Semin, G. R., & Berntson, G. G. (2004). Realism, instrumentalism, and scientific symbiosis: Psychological theory as a search for truth and the discovery of solutions. *American Psychologist*, 59, 211–223. <https://doi.org/10.1037/0003-066X.59.4.214>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105. <https://doi.org/10.1037/h0046016>
- Campbell, S. L., & Ronfeldt, M. (2018). Observational evaluation of teachers: Measuring more than we bargained for? *American Educational Research Journal* <https://doi.org/10.3102/0002831218776216>, 55, 1233, 1267
- Card, N. A. (2011). *Applied meta-analysis for social science research*. New York, NY: Guilford Press.
- Cattell, R. B. (1957). *Personality and motivation structure and measurement*. Oxford, UK: World Book Co.
- Chamberlain, G. E. (2013). Predictive effects of teachers and schools on test scores, college attendance, and earnings. *Proceedings of the National Academy of Sciences*, 110, 17176–17182. <https://doi.org/10.1073/pnas.1315746110>
- Cheng, A., & Zamarro, G. (2016). *Measuring teacher conscientiousness and its impact on students: Insight from the measures of effective teaching longitudinal database* (EDRE Working Paper No. 2016-05). Retrieved from <http://www.uaedreform.org/downloads/2016/04/edre-working-paper-2016-05.pdf>
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104, 2633–2679. <https://doi.org/10.1257/aer.104.9.2633>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- College Entrance Examination Board Committee on Research. (1963). A statement on personality testing. *College Board Review*, 51, 11–13.
- Cook, T. D., Campbell, D. T., & Peracchio, L. (1990). Quasi experimentation. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 491–576). Palo Alto, CA: Consulting Psychologists Press.
- Corcoran, R. P., & O'Flaherty, J. (2016). Personality development during teacher preparation. *Frontiers in Psychology*, 7, 1677. <https://doi.org/10.3389/fpsyg.2016.01677>
- Costa, P. T., & McCrae, R. R. (2017). The NEO Inventories as instruments of psychological theory. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 11–37). Oxford, England: Oxford University Press.
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116–127. <https://doi.org/10.1037/h0076829>
- Cutchin, G. C. (1999). *Relationships between the Big Five personality factors and performance criteria for in-service high-school teachers* (Unpublished doctoral dissertation). Purdue University, Lafayette, IN.

- Darling-Hammond, L. (2010). Teacher education and the American future. *Journal of Teacher Education*, 61, 35–47. <https://doi.org/10.1177/0022487109348024>
- Deary, I. J. (1996). A (latent) Big Five personality model in 1915? A reanalysis of Webb's data. *Journal of Personality and Social Psychology*, 71, 992–1005. <https://doi.org/10.1037/0022-3514.71.5.992>
- Delli, D. A. (2010). Hiring decisions. In M. Kennedy (Ed.), *Teacher assessment and the quest for teacher quality* (pp. 149–162). San Francisco, CA: Jossey-Bass.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417–440. <https://doi.org/10.1146/annurev.ps.41.020190.002221>
- Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, 50, 116–123. <https://doi.org/10.1037/0022-3514.50.1.116>
- Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research*, 16, 149–170. [https://doi.org/10.1207/s15327906mbr1602\\_2](https://doi.org/10.1207/s15327906mbr1602_2)
- Dobbie, W. (2011). *Teacher characteristics and student achievement: Evidence from teach for America*. Retrieved from <http://blogs.edweek.org/edweek/teacherbeat/teachercharacteristicsjuly2011.pdf>
- Dodge, A. F. (1943). What are the personality traits of the successful teacher? *Journal of Applied Psychology*, 27, 325–337. <https://doi.org/10.1037/h0062404>
- Duckworth, A. L., Quinn, P. D., & Seligman, M. E. (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology*, 4, 540–547. <https://doi.org/10.1080/17439760903157232>
- Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, 16, 939–944. <https://doi.org/10.1111/j.1467-9280.2005.01641.x>
- Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44, 237–251. <https://doi.org/10.3102/0013189X15584327>
- Dumont, F. (2010). *A history of personality psychology*. Cambridge, England: Cambridge University Press. <https://doi.org/10.1017/CBO9780511676093>
- Ellis, A. (1946). The validity of personality questionnaires. *Psychological Bulletin*, 43, 385–440. <https://doi.org/10.1037/h0055483>
- Emmerich, W., Rock, D. A., & Trapani, C. S. (2006). Personality in relation to occupational outcomes among established teachers. *Journal of Research in Personality*, 40, 501–528. <https://doi.org/10.1016/j.jrp.2005.04.001>
- Engel, M., & Finch, M. A. (2015). Staffing the classroom: How urban principals find teachers and make hiring decisions. *Leadership and Policy in Schools*, 14, 12–41. <https://doi.org/10.1080/15700763.2014.983131>
- Eysenck, H. J. (1997). *Rebel with a cause* (revised and expanded edition). New Brunswick, NJ: Transaction.
- Fernald, G. G. (1912). The defective delinquent class differentiating tests. *American Journal of Psychiatry*, 68, 523–594. <https://doi.org/10.1176/ajp.68.4.523>
- Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology*, 44, 329–344. <https://doi.org/10.1037/h0057198>
- Fountoulakis, K. N. (2015). The role of temperament. In K. N. Fountoulakis (Ed.), *Bipolar disorder* (pp. 163–179). New York, NY: Springer.
- Friedman, H. S., & Kern, M. L. (2014). Personality, well-being, and health. *Annual Review of Psychology*, 65, 719–742. <https://doi.org/10.1146/annurev-psych-010213-115123>
- Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy*, 11, 125–149. [https://doi.org/10.1162/EDFP\\_a\\_00180](https://doi.org/10.1162/EDFP_a_00180)
- Getzels, J. W., & Jackson, P. W. (1963). The teacher's personality and characteristics. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 506–582). Chicago, IL: Rand McNally.
- Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26, 461–477. <https://doi.org/10.1111/j.1744-6570.1973.tb01150.x>
- Gibby, R. E., & Zickar, M. J. (2008). A history of the early days of personality testing in American industry: An obsession with adjustment. *History of Psychology*, 11, 164–184. <https://doi.org/10.1037/a0013041>
- Goldberg, L. R. (1971). A historical survey of personality scales and inventories. In P. McReynolds (Ed.), *Advances in psychological assessment* (pp. 293–336). Palo Alto, CA: Science and Behavior Books.
- Goldberg, L. R. (1990). An alternative “description of personality”: The big-five factor structure. *Journal of Personality and Social Psychology*, 59, 1216–1229. <https://doi.org/10.1037/0022-3514.59.6.1216>
- Goldberg, L. R. (1995). What the hell took so long? Donald Fiske and the Big-Five factor structure. In P. E. Shrout & S. T. Fiske (Eds.), *Personality research, methods, and theory: A Festschrift honoring Donald W. Fiske* (pp. 29–43). Hillsdale, NJ: Erlbaum.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Robert Cloninger, C., ... Trobst, K. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40, 84–96. <https://doi.org/10.1016/j.jrp.2005.08.007>

- Goldhaber, D., Grout, C., & Huntington-Klein, N. (2014). *It's selective, but is it effective? Exploring the predictive validity of teacher selection tools*. Seattle, WA: Center for Education Data & Research.
- Goodwin, R. D., & Friedman, H. S. (2006). Health status and the five-factor personality traits in a nationally representative sample. *Journal of Health Psychology, 11*, 643–654. <https://doi.org/10.1177/1359105306066610>
- Gould, S. J. (1996). *The mismeasure of man*. New York, NY: Norton & Company.
- Gower, B. (1997). *Scientific method: An historical and philosophical introduction*. New York, NY: Routledge.
- Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26*, 499–510. [https://doi.org/10.1207/s15327906mbr2603\\_7](https://doi.org/10.1207/s15327906mbr2603_7)
- Greenwald, A. G., Pratkanis, A. R., Leippe, M. R., & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress? *Psychological Review, 93*, 216–229. <https://doi.org/10.1037/0033-295X.93.2.216>
- Gregoire, C. (2014, October 20). *Seven habits of natural leaders*. Retrieved from [http://www.huffingtonpost.com/2014/10/20/traits-that-make-a-leader\\_n\\_5959298.html](http://www.huffingtonpost.com/2014/10/20/traits-that-make-a-leader_n_5959298.html)
- Guion, R. M. (1987). Changing views for personnel selection research. *Personnel Psychology, 40*, 199–213. <https://doi.org/10.1111/j.1744-6570.1987.tb00601.x>
- Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18*, 135–164. <https://doi.org/10.1111/j.1744-6570.1965.tb00273.x>
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge, England: Cambridge University Press. <https://doi.org/10.1017/CBO9780511814563>
- Harold, C. M., McFarland, L. A., & Weekley, J. A. (2006). The validity of verifiable and non-verifiable biodata items: An examination across applicants and incumbents. *International Journal of Selection and Assessment, 14*, 336–346. <https://doi.org/10.1111/j.1468-2389.2006.00355.x>
- Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal, 51*, 73–112. <https://doi.org/10.3102/0002831213517130>
- Harris, D. N., & Rutledge, S. A. (2010). Models and predictors of teacher effectiveness: A comparison of research about teaching and other occupations. *Teachers College Record, 112*, 914–960.
- Harris, D. N., Rutledge, S. A., Ingle, W. K., & Thompson, C. C. (2010). Mix and match: What principals really look for when hiring teachers. *Education Finance and Policy, 5*, 228–246. <https://doi.org/10.1162/edfp.2010.5.2.5205>
- Harris, D. N., & Sass, T. R. (2014). Skills, productivity and the evaluation of teacher performance. *Economics of Education Review, 40*, 183–204. <https://doi.org/10.1016/j.econedurev.2014.03.002>
- Hartlep, N. D., Hansen, C. M., & Horn, B. R. (2015). *Measuring the dispositions that make teachers effective*. Cambridge, MA: Scholars Strategy Network.
- Heckman, J. J., Humphries, J. E., & Kautz, T. (Eds.). (2014). *The myth of achievement tests: The GED and the role of character in American life*. Chicago, IL: University of Chicago Press.
- Heckman, J. J., & Kautz, T. (2014). Achievement tests and the role of character in American life. In J. J. Heckman, J. E. Humphries, & T. Kautz (Eds.), *The myth of achievement tests: The GED and the role of character in American life* (pp. 3–54). Chicago, IL: University of Chicago Press.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences, 33*, 111–135. <https://doi.org/10.1017/S0140525X10000725>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). Most people are not WEIRD. *Nature, 466*, 29. <https://doi.org/10.1038/466029a>
- Heymans, G., & Wiersma, E. (1906). Beiträge zur speziellen psychologie auf grund einer massenuntersuchung. *Zeitschrift für Psychologie, 43*, 81–127.
- Hines, L. M. (2010). Disposition assessment: The science and psychology of teacher selection. In M. Kennedy (Ed.), *Teacher assessment and the quest for teacher quality* (pp. 43–65). San Francisco, CA: Jossey-Bass.
- Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology, 92*, 1270–1285. <https://doi.org/10.1037/0021-9010.92.5.1270>
- Hogan, R. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 873–919). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Hogan, R. (1998). Reinventing personality. *Journal of Social and Clinical Psychology, 17*, 1–10. <https://doi.org/10.1521/jscp.1998.17.1.1>
- Hough, L. M., & Schneider, R. J. (1995). Personality traits, taxonomies, and applications in organizations. In K. R. Murphy (Ed.), *Individuals and behavior in organizations* (pp. 31–88). San Francisco, CA: Jossey-Bass.
- Huang, J. L., Ryan, A. M., Zabel, K. L., & Palmer, A. (2014). Personality and adaptive performance at work: A meta-analytic investigation. *Journal of Applied Psychology, 99*, 162–179. <https://doi.org/10.1037/a0034285>

- Hudson, N. W., & Fraley, R. C. (2015). Volitional personality trait change: Can people choose to change their personality traits? *Journal of Personality and Social Psychology, 109*, 490–507. <https://doi.org/10.1037/pspp0000021>
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*, 869–879. <https://doi.org/10.1037/0021-9010.85.6.869>
- Jackson, C. K. (2012). *Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina* (NBER Working Paper No. 18624). Cambridge, MA: National Bureau of Economic Research.
- Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (2014). Teacher effects and teacher-related policies. *Annual Review of Economics, 6*, 801–825.
- Jackson, J. J., Hill, P. L., Payne, B. R., Roberts, B. W., & Stine-Morrow, E. A. L. (2012). Can an old dog learn (and want to experience) new tricks? Cognitive training increases openness to experience in older adults. *Psychology and Aging, 27*, 286–292. <https://doi.org/10.1037/a0025918>
- Jackson, J. J., Thoemmes, F., Jonkmann, K., Lüdtke, O., & Trautwein, U. (2012). Military training and personality trait development. *Psychological Science, 23*, 270–277. <https://doi.org/10.1177/0956797611423545>
- Jacob, B., Rockoff, J.E., Taylor, E.S., Lindy, B., & Rosen, R. (2016). *Teacher applicant hiring and teacher performance: Evidence from DC public schools*. Cambridge, MA: National Bureau of Economic Research. <https://doi.org/10.3386/w22054>
- James, W. (1890). *The principles of psychology* (Volume 1). Cambridge, MA: Harvard University Press.
- John, O. P., Robins, R. W., & Pervin, L. A. (Eds.) (2008). *Handbook of personality: Theory and research* (3rd ed.). New York, NY: Guilford Press.
- Jones, B. K. (2016). Enduring in an “impossible” occupation: Perfectionism and commitment to teaching. *Journal of Teacher Education, 67*, 437–446. <https://doi.org/10.1177/0022487116668021>
- Jones, D. (2010). A WEIRD view of human nature skews psychologists' studies. *Science, 328*, 1627. <https://doi.org/10.1126/science.328.5986.1627>
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*, 765–780. <https://doi.org/10.1037/0021-9010.87.4.765>
- 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. <https://doi.org/10.1037/0021-9010.87.3.530>
- Kagan, J. (2005). Personality and temperament: Historical perspectives. In M. Rosenbluth, S. H. Kennedy, & R. M. Bagby (Eds.), *Depression and personality: Conceptual and clinical challenges* (pp. 3–18). Washington, DC: American Psychiatric Publishing, Inc.
- Kendrick, S. A. (1964). The personality testing tangle. *College Board Review, 54*, 26–30.
- Kennedy, M. M. (2012). Sorting out teacher quality. *The Phi Delta Kappan, 90*, 59–63. <https://doi.org/10.1177/003172170809000115>
- Kim, L. E., Dar-Nimrod, I., & MacCann, C. (2017). Teacher personality and teacher effectiveness in secondary school: Personality predicts teacher support and student self-efficacy but not academic achievement. *Journal of Educational Psychology, 110*, 309–323. <https://doi.org/10.1037/edu0000217>
- Klassen, R. M., Durksen, T. L., Al Hashmi, W., Kim, L. E., Longden, K., Metsäpelto, R. L., ... & Györi, J. G. (2018). National context and teacher characteristics: Exploring the critical non-cognitive attributes of novice teachers in four countries. *Teaching and Teacher Education, 72*, 64–74. <https://doi.org/10.1016/j.tate.2018.03.001>
- Klassen, R. M., & Tze, V. M. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review, 12*, 59–76. <https://doi.org/10.1016/j.edurev.2014.06.001>
- Klette, K. (2007). Trends in research on teaching and learning in schools: Didactics meets classroom studies. *European Educational Research Journal, 6*, 147–160. <https://doi.org/10.2304/eeerj.2007.6.2.147>
- Kyllonen, P. C., Lipnevich, A. A., Burrus, J., & Roberts, R. D. (2014). *Personality, motivation, and college readiness: A prospectus for assessment and development*. (Research Report No. RR-14-06). Princeton, NJ: Educational Testing Service. <https://doi.org/10.1002/ets2.12004>
- Leary, M. R., & Hoyle, R. H. (Eds.) (2009). *Handbook of individual differences in social behavior*. New York, NY: Guilford Press.
- Lievens, F., Lang, J. W., De Fruyt, F., Corstjens, J., Van de Vijver, M., & Bledow, R. (2018). The predictive power of people's intraindividual variability across situations: Implementing whole trait theory in assessment. *Journal of Applied Psychology, 103*, 753–771. <https://doi.org/10.1037/apl0000280>
- Lüdtke, O., Roberts, B. W., Trautwein, U., & Nagy, G. (2011). A random walk down university avenue: Life paths, life events, and personality trait change at the transition to university life. *Journal of Personality and Social Psychology, 101*, 620–637. <https://doi.org/10.1037/a0023743>
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19–40. <https://doi.org/10.1037/1082-989X.7.1.19>
- Magidson, J. F., Roberts, B. W., Collado-Rodriguez, A., & Lejuez, C. W. (2014). Theory-driven intervention for changing personality: Expectancy value theory, behavioral activation, and conscientiousness. *Developmental Psychology, 50*, 1442–1450. <https://doi.org/10.1037/a0030583>

- Matsumoto, D. (Ed.). (2009). *The Cambridge dictionary of psychology*. Cambridge, England: Cambridge University Press.
- Maurer, A. B. (2012). *The impact of positive traits on teacher performance within a systems driven charter school management organization* (Unpublished master's thesis). University of Pennsylvania, Philadelphia, PA.
- Maxwell, S. E., Kelley, K., & Rausch, J. R. (2008). Sample size planning for statistical power and accuracy in parameter estimation. *Annual Review of Psychology*, *59*, 537–563. <https://doi.org/10.1146/annurev.psych.59.103006.093735>
- McAdams, D. P. (1997). A conceptual history of personality psychology. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 3–39). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-012134645-4/50002-0>
- McCrae, R. R. (2002). NEO-PI-R data from 36 cultures: Further intercultural comparisons. In R. R. McCrae & J. Allik (Eds.), *The five-factor model of personality across cultures* (pp. 105–125). New York, NY: Kluwer Academic/Plenum Publishers. [https://doi.org/10.1007/978-1-4615-0763-5\\_6](https://doi.org/10.1007/978-1-4615-0763-5_6)
- McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, *49*, 710–721. <https://doi.org/10.1037/0022-3514.49.3.710>
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81–90. <https://doi.org/10.1037/0022-3514.52.1.81>
- McCrae, R. R., & Costa, P. T. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality*, *57*, 17–40. <https://doi.org/10.1111/j.1467-6494.1989.tb00759.x>
- McCrae, R. R., & Costa, P. T. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 157–180). New York, NY: Guilford Press.
- McCrae, R. R., Terracciano, A., & 79 members of the Personality Profiles of Cultures Project. (2005). Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology*, *89*, 407–425. <https://doi.org/10.1037/0022-3514.89.3.407>
- McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, *136*, 450–470. <https://doi.org/10.1037/a0019216>
- Metzger, S. A., & Wu, M. J. (2008). Commercial teacher selection instruments: The validity of selecting teachers through beliefs, attitudes, and values. *Review of Educational Research*, *78*, 921–940. <https://doi.org/10.3102/0034654308323035>
- Mischel, W. (1968). *Personality and assessment*. New York, NY: Wiley.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G.; PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7): e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, *60*, 683–729. <https://doi.org/10.1111/j.1744-6570.2007.00089.x>
- Motowidlo, S. J., & Kell, H. J. (2013). Job performance. In N. Schmitt & S. Highhouse (Eds.), *Handbook of psychology, Vol. 12: Industrial and organizational psychology* (2nd ed., pp. 82–103). New York, NY: Wiley.
- Mueller-Hanson, R., Heggstad, E. D., & Thornton III, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology*, *88*, 348–355. <https://doi.org/10.1037/0021-9010.88.2.348>
- Ng, T. W., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, *58*, 367–408. <https://doi.org/10.1111/j.1744-6570.2005.00515.x>
- Noller, P., Law, H., & Comrey, A. L. (1987). Cattell, Comrey, and Eysenck personality factors compared: More evidence for the five robust factors? *Journal of Personality and Social Psychology*, *53*, 775–782. <https://doi.org/10.1037/0022-3514.53.4.775>
- O'Neill, T. A., Lee, N. M., Radan, J., Law, S. J., Lewis, R. J., & Carswell, J. J. (2013). The impact of "non-targeted traits" on personality test faking, hiring, and workplace deviance. *Personality and Individual Differences*, *55*, 162–168. <https://doi.org/10.1016/j.paid.2013.02.027>
- Oh, I.-S., & Roth, P. L. (2017). On the mystery (or myth) of challenging principles and methods of validity generalization (VG) based on fragmentary knowledge and improper or outdated practices of VG. *Industrial and Organizational Psychology*, *10*, 479–485. <https://doi.org/10.1017/iop.2017.45>
- Oh, I.-S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*, *96*, 762–773. <https://doi.org/10.1037/a0021832>
- Oswald, F. L., & Hough, L. M. (2011). Personality and its assessment in organizations: Theoretical and empirical developments. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology: Vol. 2. Selecting and developing members for the organization* (pp. 153–184). Washington, DC: American Psychological Association.
- Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401–421. <https://doi.org/10.1146/annurev.psych.57.102904.190127>
- Pajak, E. F. (2012). Willard Waller's sociology of teaching reconsidered: "What does teaching do to teachers?" *American Educational Research Journal*, *49*, 1182–1213. <https://doi.org/10.3102/0002831212444591>



- Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology, 81*, 524–539. <https://doi.org/10.1037/0022-3514.81.3.524>
- Paunonen, S. V., & LeBel, E. P. (2012). Socially desirable responding and its elusive effects on the validity of personality assessments. *Journal of Personality and Social Psychology, 103*, 158–175. <https://doi.org/10.1037/a0028165>
- Peterson, D. R. (1968). *The clinical study of social behavior*. New York, NY: Appleton-Century-Crofts.
- Pittenger, D. J. (2004). The limitations of extracting typologies from trait measures of personality. *Personality and Individual Differences, 37*, 779–787. <https://doi.org/10.1016/j.paid.2003.10.006>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning, 64*, 878–912. <https://doi.org/10.1111/lang.12079>
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*, 322–338. <https://doi.org/10.1037/a0014996>
- Redding, C., & Henry, G. T. (2018). Leaving school early: An examination of novice teachers' within- and end-of-year turnover. *American Educational Research Journal* <https://doi.org/10.3102/0002831218790542>, 56, 204, 236
- Revelle, W. (2008). Association for research in personality: The home for psychological generalists. *The Online Newsletter for Personality Science, 1*, 1–5.
- Revelle, W., & Wilt, J. (2013). The general factor of personality: A general critique. *Journal of Research in Personality, 47*, 493–504. <https://doi.org/10.1016/j.jrp.2013.04.012>
- Revelle, W., Wilt, J., & Condon, D. (2011). Individual differences and differential psychology: A brief history and prospect. T. Chamorro-Premuzic, A. Furnham, & S. von Stumm (Eds.), *Wiley-Blackwell handbook of individual differences* (pp. 3–38). Oxford, UK: Wiley-Blackwell.
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353–387. <https://doi.org/10.1037/a0026838>
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*, 417–458. <https://doi.org/10.1111/j.1468-0262.2005.00584.x>
- Roberts, B. W. (2018). A revised sociogenomic model of personality traits. *Journal of Personality, 86*, 23–35. <https://doi.org/10.1111/jopy.12323>
- Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology, 84*, 582–593. <https://doi.org/10.1037/0022-3514.84.3.582>
- Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3–25. <https://doi.org/10.1037/0033-2909.126.1.3>
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*, 313–345. <https://doi.org/10.1111/j.1745-6916.2007.00047.x>
- Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin, 143*, 117–141. <https://doi.org/10.1037/bul0000088>
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>
- 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*, 1–27.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review, 94*, 247–252. <https://doi.org/10.1257/0002828041302244>
- Rockoff, J. E., Jacob, B., Kane, T., & Staiger, D. (2011). Can you recognize an effective teacher when you recruit one? *Education, 6*, 43–74.
- Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal, 50*, 4–36. <https://doi.org/10.3102/0002831212463813>
- Rose, D. S., English, A., & Finney, T. G. (2014). *Hire better teachers now*. Cambridge, MA: Harvard Education.
- Ross, C. C., & Stanley, J. C. (1954). *Measurement in today's schools* (3rd ed.). New York, NY: Prentice-Hall.
- Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology, 83*, 634–644. <https://doi.org/10.1037/0021-9010.83.4.634>
- Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review, 5*, 2–14. [https://doi.org/10.1207/S15327957PSPR0501\\_1](https://doi.org/10.1207/S15327957PSPR0501_1)
- Rozin, P. (2009). What kind of empirical research should we publish, fund, and reward? A different perspective. *Perspectives on Psychological Science, 4*, 435–439. <https://doi.org/10.1111/j.1745-6924.2009.01151.x>
- Rutledge, S. A., Harris, D. N., Thompson, C. T., & Ingle, W. K. (2008). Certify, blink, hire: An examination of the process and tools of teacher screening and selection. *Leadership and Policy in Schools, 7*, 237–263. <https://doi.org/10.1080/15700760701822132>

- Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, 82, 30–43. <https://doi.org/10.1037/0021-9010.82.1.30>
- Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76, 323–346. <https://doi.org/10.1348/096317903769647201>
- Schmitt, D. P., Allik, J., McCrae, R. R., Benet-Martínez, V., Alcalay, L., & Ault, L. (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. <https://doi.org/10.1177/0022022106297299>
- Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143, 565–600. <https://doi.org/10.1037/bul0000098>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47, 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Scott, C., & Dinham, S. (2008). Born not made: The nativist myth and teachers' thinking. *Teacher Development*, 12, 115–124. <https://doi.org/10.1080/13664530802038105>
- Sechrest, L. (1976). Personality. *Annual Review of Psychology*, 27, 1–27. <https://doi.org/10.1146/annurev.ps.27.020176.000245>
- Shultz, M. M., & Zedeck, S. (2011). Predicting lawyer effectiveness: Broadening the basis for law school admission decisions. *Law & Social Inquiry*, 36, 620–661. <https://doi.org/10.1111/j.1747-4469.2011.01245.x>
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, 12, 1123–1128. <https://doi.org/10.1177/1745691617708630>
- Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 47–103). Hillsdale, NJ: Erlbaum.
- Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, 101, 862–882. <https://doi.org/10.1037/a0024950>
- 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, 293–317. <https://doi.org/10.3102/0162373715616249>
- Tang, T. Z., DeRubeis, R. J., Hollon, S. D., Amsterdam, J., Shelton, R., & Schalet, B. (2009). Personality change during depression treatment. *Archives of General Psychiatry*, 66, 1322–1330. <https://doi.org/10.1001/archgenpsychiatry.2009.166>
- Tenopyr, M. L. (1999). A scientist-practitioner's viewpoint on the admissibility of behavioral and social scientific information. *Psychology, Public Policy, and Law*, 5, 194–202. <https://doi.org/10.1037/1076-8971.5.1.194>
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703–742. <https://doi.org/10.1111/j.1744-6570.1991.tb00696.x>
- Thornton, B., Peltier, G., & Hill, G. (2005). Do future teachers choose wisely: A study of pre-service teachers' personality preference profiles. *College Student Journal*, 39, 489–496.
- Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, 41, 1–32. <https://doi.org/10.1037/h0075959>
- Tryon, W. W. (1979). The test–trait fallacy. *American Psychologist*, 34, 402–406. <https://doi.org/10.1037/0003-066X.34.5.402>
- Tyler, F. T. (1949). Personality tests and teaching ability. *Canadian Journal of Psychology*, 3, 30–37. <https://doi.org/10.1037/h0084083>
- Van Dam, K. (2003). Trait perception in the employment interview: A five-factor model perspective. *International Journal of Selection and Assessment*, 11, 43–55. <https://doi.org/10.1111/1468-2389.00225>
- Van Iddekinge, C. H., Raymark, P. H., & Roth, P. L. (2005). Assessing personality with a structured employment interview: Construct-related validity and susceptibility to response inflation. *Journal of Applied Psychology*, 90, 536–552. <https://doi.org/10.1037/0021-9010.90.3.536>
- Vandenberg, S. G. (1956). A Chinese classic on individual differences. *American Psychologist*, 11, 714–715. <https://doi.org/10.1037/h0046973>
- Vernon, P. E. (1964). *Personality assessment: A critical survey*. New York, NY: Wiley.
- Vinchur, A. J., Schippmann, J. S., Switzer, F. S., & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology*, 83, 586–597. <https://doi.org/10.1037/0021-9010.83.4.586>
- Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology*, 90, 108–131. <https://doi.org/10.1037/0021-9010.90.1.108>
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on *p*-values: Context, process, and purpose. *American Statistician*, 70, 129–133. <https://doi.org/10.1080/00031305.2016.1154108>
- Webb, E. (1915). *Character and intelligence: An attempt at an exact study of character*. London, England: Cambridge University Press.
- Widiger, T. A. (Ed.) (2017a). *The Oxford handbook of the Five Factor Model*. Oxford, England: Oxford University Press.
- Widiger, T. A. (2017b). Introduction. In T. A. Widiger (Ed.), *The Oxford handbook of the Five Factor Model* (pp. 1–7). Oxford, England: Oxford University Press.

- Wiggins, J. S., & Trapnell, P. D. (1997). Personality structure: The return of the big five. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 737–765). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-012134645-4/50029-9>
- Wilson Van Voorhis, C. R., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorial in Quantitative Methods for Psychology*, 3, 43–50. <https://doi.org/10.20982/tqmp.03.2.p043>
- Witty, P. (1947). An analysis of the personality traits of the effective teacher. *The Journal of Educational Research*, 40, 662–671. <https://doi.org/10.1080/00220671.1947.10881565>
- Zickar, M. J., & Kostek, J. A. (2013). History of personality testing within organizations. In N. D. Christiansen & R. P. Tett (Eds.), *Handbook of personality at work* (pp. 173–190). New York, NY: Routledge.

### **Suggested citation:**

Kell, H. (2019). *Do teachers' personality traits predict their performance? A comprehensive review of the empirical literature from 1990 to 2018* (Research Report No. RR-19-04). Princeton, NJ: Educational Testing Service. <https://doi.org/10.1002/ets2.12241>

**Action Editor:** Heather Buzick

**Reviewers:** Gary Sykes

ETS, the ETS logo, and MEASURING THE POWER OF LEARNING. are registered trademarks of Educational Testing Service (ETS). SAT is a registered trademark of the College Board. All other trademarks are property of their respective owners.

Find other ETS-published reports by searching the ETS ReSEARCHER database at <http://search.ets.org/researcher/>