INTRODUCTION

Performance is a critically important outcome that predicts success in life. Students who perform better in school are more likely to enter prestigious programs and jobs, employees who perform well at work are more likely to be promoted and receive raises, and people who perform well in other cognitive and social domains are more likely to receive a variety of material and social rewards. Further, poor performance can lead to school remediation, job termination, or social rejection. Given the extraordinary stakes of performance, scientists in a variety of disciplines including psychology, education, and management have long been fascinated by whether basic personality traits can be used to predict who will perform well (Hurtz & Donovan, 2000; Penney et al., 2011; Richardson et al., 2012). In the present research, we quantitatively aggregate over a century of research examining the extent to which personality traits
are associated with performance across a wide array of performance types, samples, and settings. The present report, therefore, represents one of the most comprehensive analyses of the personality—performance connection in the literature to date.

Although other categorization schemes exist, the Five-Factor Model or Big Five has emerged as the dominant framework in personality science (Costa & McCrae, 1992; John et al., 2008; John & Srivastava, 1999). According to this model, personality can be reduced to five basic traits: (1) extraversion, which reflects qualities such as warmth, gregariousness, assertiveness, activity, and positive emotion, (2) agreeableness, which reflects trust, altruism, compliance, and modesty, (3) conscientiousness, which reflects order, dutifulness, achievement-striving, self-discipline, and deliberation, (4) neuroticism, which reflects anxiety, hostility, depression, self-consciousness, and vulnerability, and (5) openness, which reflects curiosity, creativity, resourcefulness, and a willingness to consider unconventional ideas. The five-factor conceptualization of personality generalizes across measures, cultures, and rating sources (McCrae & John, 1992). Further, in part because personality traits are influenced by genetic factors (Briley & Tucker-Drob, 2014) and early life experiences (Fraley et al., 2013), they are relatively stable across time (Roberts & DelVecchio, 2000).

In addition to widespread agreement regarding the nature of personality, behavioral scientists largely agree on the definition of performance. Specifically, performance is conceptualized as a behavior or action that is (1) directly observed, (2) relevant to important goals, (3) controlled at least in part by the individual, and (4) used to rank individuals in terms of proficiency (Campbell, 1990; Campbell et al., 1993; see also, Carpini et al., 2017; Griffin et al., 2007). Performance outcomes span a wide array of cognitive, motor, psychomotor, or interpersonal outcomes (Bartram, 2005). Given its emphasis on explicit behavioral outcomes, the prediction of performance is considered the “holy grail” of several disciplines (O’Boyle et al., 2011). Some scholars have focused on contextual variables that may influence performance, such as stereotype threat (Spencer et al., 2016), practice (Ericsson, 2008), homework (Cooper et al., 2006), and ego-depletion (Hagger & Chatzisarantis, 2016). Alternatively, other scholars have sought to identify individual difference variables that may be associated with performance, such as age (Rhodes, 2004), gender (Hyde, 2014), socioeconomic status (Westrick et al., 2015), and most relevant to the current report, personality.

1.1 | Personality and performance

Since the early 1900s, thousands of studies have explored the association of personality traits with performance. The widespread adoption of the Five-Factor Model in the mid-1980s dramatically increased the frequency of these studies and provided a useful framework for the interpretation of earlier studies. As the number of individual studies continued to accumulate, scholars increasingly began to use meta-analysis to estimate the average size of the association between personality and performance. Prominent meta-analyses, for example, examined the association of Big Five traits with job performance (Barrick & Mount, 1991) and academic performance (Poropat, 2009) as well as performance on intelligence (Ackerman & Heggestad, 1997) and emotional intelligence tests (Joseph & Newman, 2010). A major benefit of meta-analysis is that it is based on multiple independently conducted studies and thus generates more reliable effect size estimates than individual studies, which are often plagued by low statistical power and unduly focused on the outcome of (often arbitrary) statistical significance tests (Paterson et al., 2016).

As of present, dozens of unique meta-analyses have been published on the association of Big Five traits with performance, with each varying to some extent in the personality measures, performance outcomes, or samples included. Along these lines, meta-analyses have examined the association of Big Five traits and performance when examining specific personality inventories (Hogan & Holland, 2003), forced-choice measures (Salgado et al., 2015), and other-ratings as opposed to self-ratings of personality (Connelly & Ones, 2010). Further, meta-analyses have examined the association of Big Five traits with a wide array of performance types, including job and academic performance, but also computer programming (Gnams, 2015), creativity (Puryear et al., 2017), interpersonal sensitivity (Hall et al., 2009), leadership effectiveness (Judge et al., 2002), and negotiation (Sharma et al., 2013). Finally, some meta-analyses on the personality—performance connection have focused on specific populations, such as children (Poropat, 2014), university students (McAbee & Oswald, 2013), salespersons (Vinchar et al., 1998), Europeans (Salgado, 1997), or South Africans (van Aarde et al., 2017).

Nonetheless, despite the longstanding interest in the association of personality with overall or aggregate performance (Barrick et al., 2001), research has yet to integrate these disparate findings. Therefore, the fundamental question of how well the Big Five predict performance in general (i.e., a latent variable indicated by shared variance in performance in specific domains) has yet to be systematically tested. Moreover, despite a common interest in personality and performance, there has been little communication among scholars in different research traditions examining this topic. The present research capitalizes upon the now large database of meta-analyses on personality—performance associations to provide the first...
direct test of the association of personality with overall performance, thereby filling a major gap in the personality literature. In other words, rather than focusing on the personality—performance association when examining a specific measure, performance domain, or population, we adopt a broader “birds-eye” perspective to provide an up-to-date synthesis of this literature as a whole, assess its strengths and weaknesses, and inform the next generation of research on Big Five—performance associations.

1.2 Research aims

Four primary aims motivated this research. First, by aggregating across an enormous database of studies cataloged in prior meta-analyses, we sought to derive comprehensive estimates of the association between each Big Five trait and performance. These estimates allowed us to determine whether Big Five associations exist, and if yes, how large they are. We anticipated that each of the Big Five would yield at least a small association with overall performance. More specifically, extraversion, agreeableness, conscientiousness, low neuroticism, and openness were each expected to be conducive to performance (Judge & Ilies, 2002). Although mechanisms underlying personality—performance associations have received relatively little theoretical and empirical attention, the benefits of extraversion may come via increased assertiveness, positive emotions, and activity levels, agreeableness via the ability to manage conflict or other aversive situations and thus avoid counterproductive behaviors, conscientiousness via enhanced work ethic and self-control, low neuroticism via increased motivation and self-efficacy, and openness via interest in intellectual, imaginative, and other creative pursuits (Gerhardt et al., 2007; Hampson, 2012; Judge & Zapata, 2015).

Second, we sought to compare Big Five—performance associations to each other to determine their relative magnitude. Conscientiousness is theorized to be a uniquely important trait, given its connection with qualities like self-discipline, grit, planning, and sustained goal-directed behavior over time (Roberts et al., 2014; Schneider & Preckel, 2017). Indeed, initial meta-analyses on performance in educational and vocational settings indicated that conscientiousness yielded an association with performance that was at least 40% larger and generally more robust to a variety of potential moderating factors than other Big Five traits (Barrick & Mount, 1991; Poropat, 2009). Prior theory and research, however, provide less clarity about the relative effects of the remaining traits. Given that the effects of extraversion, agreeableness, neuroticism, and openness might be less durable than conscientiousness and more specific to performance settings that fit the unique features of these traits—for example, interpersonal settings for extraversion and agreeableness (Mount et al., 1998; Salgado & Táuriz, 2014) and intellectual or creative settings for openness (Ackerman & Heggestad, 1997; Puryear et al., 2017)—we anticipated that each would exert a smaller overall association with performance than conscientiousness.

Third, we provide the first direct comparison of the size of Big Five—performance associations across different categories of performance, particularly job and academic performance. Although some have proposed that personality—performance associations should be similar across these two high-stakes settings (e.g., Lounsbury et al., 2004), we anticipated that the pattern of associations would vary substantially across settings. Conscientiousness has been viewed as particularly important in academic settings, because it helps students maintain focus on academic tasks even when these tasks are of low personal importance or conflict with other goals (Duckworth et al., 2019; Kappe & van der Flier, 2012). Conversely, extraversion, agreeableness, and neuroticism may be particularly important in work settings, especially those in which relationships with clients or co-workers are paramount (e.g., sales; Mount et al., 1998). Nonetheless, little research has directly compared the associations of Big Five traits with performance across academic and vocational contexts. Further, prior theory and research has yet to synthesize personality—performance effects that fall outside these two cardinal domains. Thus, potential effect sizes as well as the number and nature of the meta-analyses conducted in these other areas remains unclear.

Fourth, in the cases where multiple meta-analyses on the same topic were conducted by independent research teams (i.e., meta-analyses examining the same personality—performance association in the same population), we capitalize on these data to explore whether meta-analytic estimates of Big Five—performance associations are replicable. Meta-analysts make many decisions—for example, how to operationalize the variables of interest, which studies to include or exclude, how individual-study effect sizes are calculated or converted, and how data are aggregated across studies—each of which could be influenced by preferences of the research team. Thus, if meta-analytic estimates of personality—performance associations are consistent across independently conducted meta-analyses, this would suggest that the obtained results are likely valid estimates of population effects and not unduly influenced by the idiosyncratic decisions of any research team. Although concerns about replicability of individual studies have increasingly occupied behavioral scientists (Klein et al., 2018; Open Science Collaboration, 2015), little work has examined the replicability of meta-analyses. Given the large number of meta-analyses conducted, research on personality and
performance provides a unique opportunity to address this question.

1.3 | Theoretical contribution

For the first time in the personality literature, the present research uses quantitative aggregation procedures to examine the association of each Big Five trait with performance across the vast collection of topics studied in prior meta-analyses. Our work builds upon the landmark research of Barrick et al. (2001), who synthesized findings from 8 partially overlapping meta-analyses to estimate the association of personality with job performance. When examining effects corrected for measurement error and range restriction, each of the Big Five yielded substantive associations with job performance, with the smallest effects emerging for openness (0.07) and agreeableness (0.11), intermediate effects for extraversion (0.15) and neuroticism (−0.15), and the largest effect emerging for conscientiousness (0.24); meta-analyses restricted to 5 independent (i.e., non-overlapping) meta-analyses yielded similar results. This innovative study was the first to aggregate findings across personality—job performance meta-analyses in effort to estimate the overall size and variability of Big Five effects. Nonetheless, many meta-analyses on personality and performance, both in the area of job performance and elsewhere, have been conducted since 2001. As we describe below, there are now over 50 unique meta-analyses on personality and performance, including over 30 in the job performance domain. Thus, a comprehensive analysis of the personality—performance association in light of these new data is in order.

More recently, two major research syntheses have examined the association of a single Big Five trait with job and academic performance. Conscientiousness yielded a medium-sized corrected association with performance across 48 partially overlapping meta-analyses (0.17; Wilmot & Ones, 2019) and extraversion yielded a small corrected association with performance across 45 partially overlapping meta-analyses (0.10; Wilmot et al., 2019). These studies provide robust estimates of the association of conscientiousness and extraversion with performance. The present research significantly expands upon this work by examining the association of the other Big Five traits (agreeableness, neuroticism, and openness) with performance. Further, rather than focus on a single trait, the present research simultaneously examined all of the Big Five, which enabled direct comparisons of the size of their associations with performance under equivalent measurement conditions. Finally, the present research incorporates data from any performance domain, not just those on academic or vocational performance, which facilitated novel estimates of the association of personality with overall or aggregated performance.

For theory building, it is important to understand the overall size of personality—performance associations as well as the degree to which these associations fluctuate across performance domains. If personality effects are consistent across domains, it would suggest that there is a fundamental core to performance that overlaps with personality regardless of domain. Conversely, if personality effects vary across domains, it would suggest that personality—performance associations are context specific, and that associations in one domain (e.g., academic) should not be assumed to occur in other domains (vocational). By aggregating data from over 50 meta-analyses, the present research is uniquely positioned to address these questions and to identify potential sources of variation in personality—performance effects.

2 | METHOD

We followed recommended procedures for second-order meta-analysis (Cooper & Koenka, 2012; Zell & Krizan, 2014). First, we conducted an exhaustive literature search to identify potentially relevant papers. Second, we sorted these papers using formal inclusion criteria into a final set of relevant meta-analyses. Third, effect sizes and other details were extracted from meta-analyses to facilitate data analysis. Fourth, quantitative aggregation procedures were used to assess the association of Big Five traits with performance, both overall and in specific performance categories. Data and supplemental materials for this project are publicly available at https://osf.io/6z7wk/?view_only=c8333258e5f8417e96ce0e20a567b0f6.

2.1 | Article search and inclusions

We searched three scholarly databases (i.e., PsychInfo, Business Source Complete, and ERIC) for relevant meta-analyses. Specifically, databases were searched for records that mentioned Big Five, performance, and meta-analysis in the title, abstract, or keywords. Big Five search terms included broad labels (e.g., personality, traits, temperament), names of Big Five measures and models (e.g., Big Five, Five Factor, BFI, and NEO-PI) and names of the 5 factors (e.g., extraversion, agreeableness). Performance search terms included broad labels (e.g., performance, effectiveness, productivity) and terms that reflect specific types of performance (e.g., job performance, GPA, emotional intelligence). The search was restricted to full-text articles published in English. Additional search details are provided in the Supplemental Materials.
After removing duplicates and a paper that was retracted, the search yielded 1050 potentially relevant articles. An additional 178 published articles were identified by scanning the reference lists of three major reviews examining the association of personality traits with organizational and educational outcomes (Schneider & Preckel, 2017; Wilmot & Ones, 2019; Wilmot et al., 2019), resulting in a total of 1228 potentially relevant articles.

We next evaluated articles to determine whether they were eligible for inclusion in our quantitative synthesis (see the Supplemental Materials for a flow chart). To be included, articles had to provide meta-analytic effect sizes indexing the relation between performance and all Big Five traits. Thus, papers were excluded if they were not a meta-analysis, if they did not examine all Big Five traits, or if they did not examine performance \((n = 1107)\). Big Five traits were defined broadly to include self-ratings and other-ratings of personality as well as ratings on any Big Five measure or measure that can be interpreted using a Big Five framework (e.g., the Hogan Personality Inventory; Hogan & Hogan, 2007). Performance was also defined broadly to include any behavioral index of individual productivity, achievement, or effectiveness, typically assessed via objective measures (e.g., sales volume, GPA) or other external criteria (e.g., supervisor, teacher, or peer ratings of performance in general or in specific contexts). Thus, we excluded papers examining team or group performance as well as papers focusing on correlates of performance rather than direct measures of performance per se (e.g., performance satisfaction, motivation, counterproductive behaviors, and leadership style).

The remaining papers were further evaluated and excluded if they did not provide meta-analytic effect sizes indexing the zero-order correlation between performance and all Big Five traits or if they merely reported (and reanalyzed) effect sizes obtained in prior meta-analyses \((n = 67)\). After these exclusions, a total of 54 meta-analyses on the relation between Big Five traits and performance remained (see the Supplemental Materials for a list). The entire database was reviewed by the second author, who screened out articles that clearly were not meta-analyses or did not include measures of the Big Five or performance \((n = 1077)\). The remaining articles were then evaluated for inclusion by the first author \((n = 151)\). All included meta-analyses met a minimum threshold of \(k \geq 3\) or \(N \geq 500\) for each Big Five effect (Wilmot et al., 2019).

### 2.2 Effect extraction and coding

Effect sizes indexing the overall correlation between each Big Five measure and performance were extracted from included meta-analyses. All meta-analyses provided effects that were weighted by sample size (i.e., used fixed-effect, random-effects, or other approaches). Where possible, we extracted effect sizes that were corrected for measurement unreliability and range restriction, but as we note below, some meta-analyses did not correct effects. When meta-analyses reported effect sizes for multiple performance outcomes \((m = 17)\), these effects were averaged so that each meta-analysis contributed only one effect and was therefore weighted equally in data analyses. Further, when meta-analyses reported separate effect sizes for multiple indicators of a Big Five trait (e.g., ambition and sociability for extraversion; \(m = 4\)), these effects were averaged before entry into the model.

The final set of meta-analyses was also coded to determine whether they examined job, academic, or other performance types. Job performance was broadly defined to include any measure of proficiency or effectiveness in the workplace including overall indicators of job performance as well as contextual performance, personnel data, performance during interviews, training, and assessment center exercises, performance on workplace (i.e., situational) judgment tests, sales and entrepreneurial performance, and leadership effectiveness. Academic performance was defined by grade-point-average (GPA) and course grades in any grade level (i.e., primary, secondary, postsecondary). Performance dimensions that did not fit into the above categories (e.g., ability-based emotional intelligence, computer programming, creativity, intelligence or general mental ability, interpersonal sensitivity, and negotiation skill) or that spanned multiple performance dimensions (e.g., combined job and academic performance) were labelled as “other”.

Additional coding of each meta-analysis was done to determine methodological characteristics (i.e., whether or not the meta-analysis included a database search, unpublished studies, and corrections for unreliability and range restriction), sample characteristics (i.e., age and country), and personality measurement (i.e., whether personality was measured via self-ratings or other-ratings and via a single scale or multiple scales across studies). Meta-analysis coding was done by the first and second author; disagreements were resolved by discussion (all \(k > 0.85\)).

Lastly, reference sections for meta-analyses on related topics (e.g., that examined similar performance outcomes, personality measures, or samples) were compared to estimate the percentage of studies that were overlapping across them. The degree of sample overlap \((SO)\) was estimated for all pairs of related meta-analyses, and the highest amount of overlap was used as an estimate of each paper’s uniqueness.
### 2.3 Analytic approach

As have second-order meta-analyses on related topics (e.g., personality and health; Strickhouser et al., 2017), we used the unweighted average of personality—performance effects provided by prior meta-analyses to estimate population effects ($\rho$). Past research found that unweighted averages in meta-analysis perform as well or better in predicting population effects than averages that weight by study sample size or other criteria (Bonett, 2009; Shuster, 2010). Nonetheless, we also report sensitivity analyses that weight meta-analytic effects by their respective sample size ($N$) and number of effects ($k$). Overall effects are interpreted using recommended guidelines for individual differences research ($0.05 = \text{very small}, 0.10 = \text{small}, 0.20 = \text{medium}, 0.30 = \text{large}, 0.40 = \text{very large}$; Funder & Ozer, 2019; Gignac & Szordorai, 2016).

Moreover, we included all meta-analyses in the final model, even those that were reliant upon overlapping samples, for two reasons. First, even when meta-analyses shared many samples, there were still significant differences in the samples or research questions across them, such that each meta-analysis generated unique data (i.e., focused on a unique set of samples, personality measures, and/or performance outcomes). Second, inclusion of overlapping meta-analyses allowed us to explore whether personality—performance associations replicate across related meta-analyses. Nonetheless, we conducted a sensitivity analysis that inverse weighted meta-analyses by their degree of sample overlap (i.e., subtracted $SO$ from 100) to examine the degree to which it influenced our results.

### 3 RESULTS

#### 3.1 Overview and sample overlap

When examining the final set of 54 partially overlapping meta-analyses, the total number of effects ranged from 1539 to 2028 and the total sample size ranged from 406,696 to 554,778 across Big Five traits. Most meta-analyses aggregated studies from many independent research teams ($m = 52$), included unpublished studies ($m = 45$) and adult samples ($m = 49$), and examined self-ratings of personality ($m = 51$) across studies that collectively included variety of different Big Five measures ($m = 47$). The location of the included samples was reported in 19 meta-analyses. Meta-analyses primarily focused on samples from North America and Europe, but a few included samples from Asia ($m = 5$), Australia ($m = 4$), and Africa ($m = 3$) as well. Additional methodological details are provided in the Supplemental Materials.

The obtained meta-analyses generally included samples that were non-overlapping with other meta-analyses (see the Supplemental Materials). Specifically, the amount of sample overlap across meta-analyses was typically low ($M = 36\%$). However, there was considerable variability in sample overlap ($SD = 36\%$), with estimates ranging from 0% ($m = 18$) to 100% ($m = 8$). Cases of high overlap reflected the tendency for meta-analyses to examine a targeted subset of studies from another meta-analysis. Thus, even when they overlapped, each meta-analysis contributed a unique estimate of the association between personality and performance in specific conditions or contexts.

#### 3.2 Primary model and sensitivity analyses

After aggregating across the full set of 54 meta-analyses, there were robust associations of each of Big Five trait with performance (see Figure 1). Specifically, extraversion (0.10), agreeableness (0.10), neuroticism (−0.12), and openness (0.13) all yielded small associations with performance that were comparable in size and in the predicted direction (i.e., negative for neuroticism and positive for the others). However, the association of conscientiousness with performance was medium in size (0.19), at least 46% larger than the other Big Five traits, and had a confidence interval that was non-overlapping with the other traits (see Table 1 for 95\% CI). These results suggest that, when viewed independently, each Big Five trait is a meaningful
### Table 1: Effect sizes from the quantitative synthesis

<table>
<thead>
<tr>
<th>Analysis</th>
<th>E</th>
<th>A</th>
<th>C</th>
<th>N</th>
<th>O</th>
<th>m</th>
<th>k</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary model</strong></td>
<td>0.10 (0.08, 0.13)</td>
<td>0.10 (0.07, 0.13)</td>
<td>0.19 (0.16, 0.21)</td>
<td>-0.12 (-0.15, -0.10)</td>
<td>0.13 (0.10, 0.15)</td>
<td>54</td>
<td>1539–2028</td>
<td>406,696–554,778</td>
</tr>
<tr>
<td><strong>Standard deviation</strong></td>
<td>0.10</td>
<td>0.11</td>
<td>0.10</td>
<td>0.10</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Minimum effect</strong></td>
<td>-0.11</td>
<td>-0.04</td>
<td>-0.04</td>
<td>-0.47</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maximum effect</strong></td>
<td>0.35</td>
<td>0.51</td>
<td>0.50</td>
<td>0.08</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sensitivity analyses</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted by k</td>
<td>0.08</td>
<td>0.08</td>
<td>0.19</td>
<td>-0.10</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted by N</td>
<td>0.07</td>
<td>0.09</td>
<td>0.20</td>
<td>-0.08</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted by SO</td>
<td>0.11</td>
<td>0.09</td>
<td>0.18</td>
<td>-0.12</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Correction type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unreliability &amp; range</td>
<td>0.14 (0.11, 0.16)</td>
<td>0.12 (0.08, 0.15)</td>
<td>0.22 (0.20, 0.24)</td>
<td>-0.16 (-0.19, -0.12)</td>
<td>0.12 (0.10, 0.15)</td>
<td>21</td>
<td>694–900</td>
<td>121,654–157,725</td>
</tr>
<tr>
<td>Unreliability</td>
<td>0.09 (0.07, 0.12)</td>
<td>0.09 (0.07, 0.12)</td>
<td>0.18 (0.15, 0.21)</td>
<td>-0.10 (-0.12, -0.08)</td>
<td>0.14 (0.11, 0.17)</td>
<td>25</td>
<td>624–842</td>
<td>211,762–297,437</td>
</tr>
<tr>
<td>None</td>
<td>0.05 (0.02, 0.08)</td>
<td>0.08 (0.06, 0.11)</td>
<td>0.13 (0.10, 0.15)</td>
<td>-0.10 (-0.13, -0.07)</td>
<td>0.12 (0.10, 0.13)</td>
<td>8</td>
<td>206–422</td>
<td>73,280–99,632</td>
</tr>
<tr>
<td><strong>Performance category</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic</td>
<td>-0.01 (-0.02, 0.01)</td>
<td>0.07 (0.07, 0.07)</td>
<td>0.28 (0.26, 0.31)</td>
<td>-0.03 (-0.05, -0.02)</td>
<td>0.14 (0.11, 0.18)</td>
<td>7</td>
<td>299–372</td>
<td>144,117–165,339</td>
</tr>
<tr>
<td>Job</td>
<td>0.14 (0.12, 0.16)</td>
<td>0.11 (0.08, 0.14)</td>
<td>0.20 (0.18, 0.22)</td>
<td>-0.15 (-0.18, -0.13)</td>
<td>0.11 (0.09, 0.14)</td>
<td>32</td>
<td>782–1099</td>
<td>145,834–252,343</td>
</tr>
<tr>
<td>Other</td>
<td>0.08 (0.05, 0.11)</td>
<td>0.09 (0.07, 0.12)</td>
<td>0.12 (0.09, 0.15)</td>
<td>-0.10 (-0.13, -0.07)</td>
<td>0.16 (0.13, 0.19)</td>
<td>15</td>
<td>434–721</td>
<td>108,290–162,725</td>
</tr>
</tbody>
</table>

*Note:* E, A, C, N and O refer to their respective Big Five traits. 95% CI of effect sizes are in parentheses. m = number of meta-analyses, k = number of effects, N = sample size, and SO = sample overlap (k and N values reflect the minimum and maximum across Big Five traits).
correlate of performance, but conscientiousness is the most robust correlate. Sensitivity analyses found that these effects were very similar when weighted by the total number of participants ($N$) or effects ($k$) in each meta-analysis, and when inverse weighted by the sample overlap of each meta-analysis ($SD$). Additionally, analyses yielded similar results after the removal of outliers (i.e., effects that were $\pm 3 SD$ from the mean; see the Supplemental Materials).

In general, the 54 meta-analytic effects were in the predicted direction for each Big Five trait. That is, most effects were positive for extraversion ($0.14$), agreeableness ($0.11$), conscientiousness ($0.12$), and openness ($0.11$), and most were negative for neuroticism ($−0.15$). However, there was considerable variability in the size of these effects across the 54 meta-analyses (see Table 1). Specifically, excluding conscientiousness, the standard deviation of each Big Five effect was numerically similar to the effect size itself. Further, meta-analytic effects ranged from a minimum value that was close to zero or even trending in the opposite direction to a maximum value that was very large and in the predicted direction.

### 3.3 Effects by correction strategy

Meta-analytic effects varied by the statistical correction strategy used in prior meta-analyses. Specifically, meta-analyses that corrected for both unreliability and range restriction (absolute effect sizes from $0.12$ to $0.22$, $m = 21$) generally yielded larger effects than meta-analyses that only corrected for unreliability (absolute effect sizes from $0.09$ to $0.18$, $m = 25$) or that did not utilize a correction procedure (absolute effect sizes from $0.05$ to $0.13$, $m = 8$). Therefore, overall effect size estimates obtained in the present research may be slight underestimates of the true association between Big Five traits and performance. The effect of correction strategy within each of the three performance categories is explored in the Supplemental Materials.

### 3.4 Effects by performance category

Many of the meta-analyses that we obtained related broadly to job performance ($m = 32$). These meta-analyses examined a variety of different types of work performance (e.g., overall performance, training, interview, contextual), performance in different locations (e.g., Europe, South Africa) and samples (e.g., military, police, expatriates, leaders, entrepreneurs, salespersons), as well as associations of performance with different Big Five measures (e.g., forced-choice, other-ratings). As in the primary model, overall effect sizes for each Big Five trait were small to medium for job performance. Once again, conscientiousness had the largest effect ($0.20$), but effects for extraversion ($0.14$), agreeableness ($0.11$), neuroticism ($−0.15$), and openness ($0.11$) were robust as well (see Figure 2).

Several of the remaining meta-analyses broadly related to academic performance ($m = 7$). These papers examined academic performance at the primary level (i.e., elementary education; $m = 1$), at the tertiary level (i.e., postsecondary or university; $m = 5$), or across primary to tertiary levels ($m = 1$). A strikingly different pattern of results obtained for academic performance. Specifically, conscientiousness yielded a large association with academic performance ($0.28$) that was noticeably bigger than associations for the other Big Five traits. Openness placed second with a small effect ($0.14$). Of the remaining traits, agreeableness yielded a small association with performance ($0.07$), but extraversion ($−0.01$) and neuroticism ($−0.03$) yielded effects that were close to zero.

Finally, several meta-analyses examined “other” performance dimensions. These papers examined a combination of academic and job performance ($m = 3$), performance on cognitive ability ($m = 4$) or emotional intelligence tests ($m = 4$), or performance in other areas (i.e., computer programming, creativity, interpersonal sensitivity, and negotiation; $m = 4$). Overall, each Big Five trait yielded a robust association with “other” performance types. Perhaps because several of these meta-analyses examined intellectual performance dimensions, the largest effect obtained for openness ($0.16$). Effects for extraversion ($0.08$), agreeableness ($0.09$), conscientiousness ($0.12$), and neuroticism ($−0.10$) were slightly smaller.

When comparing across the three performance categories, the small effect of agreeableness and openness largely replicated across all three performance categories (i.e., had $95\% CI$ that were generally overlapping). However, the small effects of extraversion and neuroticism that obtained for job performance and other performance types were significantly reduced when examining academic performance (i.e., had $95\% CI$ that were non-overlapping). Further, the effect of conscientiousness was large for academic performance, medium for job performance, and small for other performance types, with significant differences in each pairwise comparison. Analyses that were weighted by sample overlap for job performance, academic performance, and other performance types yielded a similar pattern of results (see the Supplemental Materials).

### 3.5 Effects by performance subcategory

We identified three outcomes in which personality-performance associations were examined in multiple,
partially overlapping meta-analyses conducted by largely independent research teams: university performance ($m = 5$), emotional intelligence ($m = 4$), and overall job performance ($m = 4$). For each of these outcomes, meta-analyses examined the same research question in the same population, that is, the association of self-ratings of personality with performance in samples that were not restricted to a particular region or occupation.

### 3.5.1 University performance

There was striking similarity in personality-performance associations across meta-analyses on university performance (see Table 2). In all 5 meta-analyses, conscientiousness yielded by far the largest effect, with an overall effect that was medium-to-large in size (0.25). Further, whereas openness (0.09) and agreeableness (0.07) consistently had TABLE 2 Effect sizes by performance subcategory

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>E</th>
<th>A</th>
<th>C</th>
<th>N</th>
<th>O</th>
<th>SO</th>
<th>k</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>University</strong></td>
<td>0.25</td>
<td>0.26</td>
<td>0.25</td>
<td>-0.01</td>
<td>0.09</td>
<td>170</td>
<td>211</td>
<td>81,008–88,707</td>
</tr>
<tr>
<td>McAbee and Oswald (2013)</td>
<td>-0.03</td>
<td>0.07</td>
<td>0.26</td>
<td>0.00</td>
<td>0.08</td>
<td>56%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’Connor and Paunonen (2007)</td>
<td>-0.05</td>
<td>0.06</td>
<td>0.24</td>
<td>-0.03</td>
<td>0.06</td>
<td>70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Richardson et al. (2012)</td>
<td>-0.03</td>
<td>0.06</td>
<td>0.23</td>
<td>0.01</td>
<td>0.09</td>
<td>48%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trapmann et al. (2007)</td>
<td>-0.06</td>
<td>0.06</td>
<td>0.27</td>
<td>0.00</td>
<td>0.13</td>
<td>28%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vedel (2014)</td>
<td>0.00</td>
<td>0.08</td>
<td>0.26</td>
<td>-0.01</td>
<td>0.07</td>
<td>71%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Emotional intelligence</strong></td>
<td><strong>0.09</strong></td>
<td>0.21</td>
<td>0.11</td>
<td><strong>-0.12</strong></td>
<td>0.16</td>
<td><strong>101–107</strong></td>
<td>20,367–21,281</td>
<td></td>
</tr>
<tr>
<td>Joseph and Newman (2010)</td>
<td>0.11</td>
<td>0.19</td>
<td>0.16</td>
<td>-0.12</td>
<td>0.13</td>
<td>53%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’Boyle et al. (2011)</td>
<td>0.11</td>
<td>0.26</td>
<td>0.11</td>
<td>-0.16</td>
<td>0.18</td>
<td>42%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van der Linden et al. (2017)</td>
<td>0.06</td>
<td>0.20</td>
<td>0.11</td>
<td>-0.11</td>
<td>0.18</td>
<td>41%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Rooy et al. (2005)</td>
<td>0.09</td>
<td>0.18</td>
<td>0.06</td>
<td>-0.08</td>
<td>0.14</td>
<td>63%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.02</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall job</strong></td>
<td><strong>0.13</strong></td>
<td><strong>0.10</strong></td>
<td>0.23</td>
<td><strong>-0.11</strong></td>
<td><strong>0.05</strong></td>
<td><strong>219–375</strong></td>
<td><strong>46,067–92,064</strong></td>
<td></td>
</tr>
<tr>
<td>Barrick and Mount (1991)</td>
<td>0.13</td>
<td>0.07</td>
<td>0.22</td>
<td>-0.08</td>
<td>0.04</td>
<td>42%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hurtz and Donovan (2000)</td>
<td>0.09</td>
<td>0.11</td>
<td>0.20</td>
<td>-0.13</td>
<td>0.06</td>
<td>44%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Judge et al. (2013)</td>
<td>0.18</td>
<td>0.15</td>
<td>0.28</td>
<td>-0.11</td>
<td>0.08</td>
<td>27%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Judge and Zapata (2015)</td>
<td>0.11</td>
<td>0.06</td>
<td>0.21</td>
<td>-0.12</td>
<td>0.04</td>
<td>25%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: E, A, N and O refer to their respective Big Five traits. SO = sample overlap, k = number of effects, N = sample size ($k$ and $N$ values reflect the minimum and maximum across Big Five traits). SD = standard deviation of effect size across meta-analyses. Bolding reflects higher-order effects.
small positive effects, extraversion consistently had a very small negative effect (−0.03), and neuroticism consistently had a near-zero effect (−0.01). Further attesting to the similarity in effect sizes across meta-analyses, the SD of each Big Five effect was very small and ranged from 0.01 to 0.03. Sample overlap among meta-analyses on university performance ranged from 28% to 71% (M = 55%, SD = 18%). Therefore, despite being on the same topic, each meta-analysis contributed unique samples to the primary model.

3.5.2 | Emotional intelligence

Results also replicated across 4 meta-analyses on emotional intelligence. Agreeableness had the largest effect in each meta-analysis and yielded an overall effect that was medium in size (0.21). Further, openness consistently yielded a small-to-medium effect (0.16), extraversion consistently yielded a small positive effect (0.09), and neuroticism consistently yielded a small negative effect (−0.12). Finally, conscientiousness had a small overall effect (0.11) that fluctuated somewhat across meta-analyses (range = 0.06 to 0.16). Once again, the SD of each Big Five effect across meta-analyses was relatively small, ranging from 0.02 to 0.04. Sample overlap among meta-analyses on emotional intelligence ranged from 41% to 63% (M = 50%, SD = 10%).

3.5.3 | Overall job performance

Our last analysis once again suggested that meta-analytic estimates of personality—performance relations are replicable. Specifically, when examining meta-analyses on the association of personality with overall job performance, conscientiousness consistently yielded the largest effect (0.23) and openness consistently yielded the smallest effect (0.05). Moreover, extraversion (0.13) and agreeableness (0.10) had small positive effects that fluctuated slightly across meta-analyses and neuroticism consistently had a small negative effect (−0.11). The SD of each Big Five effect, which ranged from 0.02 to 0.04, further suggests that fluctuations in effect sizes across meta-analyses were generally small. Meta-analyses on overall job performance typically had relatively low sample overlap, ranging from 25% to 44% (M = 35%, SD = 10%).

4 | GENERAL DISCUSSION

The possibility that some people, merely as a function of their basic personality, generally perform better than others has occupied scholars in a variety of disciplines for over a century. In effort to address this fundamental question, dozens of unique meta-analyses have examined the association of Big Five personality traits with performance in specific conditions or contexts. Nonetheless, the broad question of how well Big Five traits predict overall or aggregate performance remains unresolved. Addressing this gap, the present research synthesized data from 54 partially overlapping meta-analyses, collectively including over 2000 studies and 550,000 participants, to estimate associations of Big Five traits with performance. Our analysis yielded several novel insights that significantly advance the personality—performance literature.

First, Big Five traits yielded detectable, albeit generally small associations with overall performance, with each effect size crossing the 0.10 threshold. In documenting these effects, our analysis goes beyond prior research syntheses, which obtained relatively few meta-analyses (Barrick et al., 2001) or were restricted to a single Big Five trait (Wilmot & Ones, 2019; Wilmot et al., 2019). Although small, we believe that the overall associations of personality with performance should be viewed optimistically. Statistically small effects of personality traits may yield practically important consequences (Roberts et al., 2007; Strickhouser et al., 2017), especially in the performance domain, where small gains in performance bring tangible rewards such as increased entry into selective academic programs and job promotion. In sum, by showing that each Big Five trait is a valid correlate of performance, our analysis uniquely bolsters the Five Factor Model of personality (Costa & McCrae, 1992; John et al., 2008) and suggests that the model is a useful framework for the prediction of performance in high stakes settings, such as school and work (Barrick & Mount, 1991; Poropat, 2009).

Second, for the first time in the literature, we directly compared the size of personality—performance associations across many meta-analyses and found that conscientiousness yielded a significantly larger effect than the other Big Five traits, with an overall effect size that was at least 46% larger than each of its counterparts. These data are consistent with emerging theory postulating that conscientiousness is a uniquely important correlate of performance, given its overlap with critical dimensions like self-control, grit, and planning (Duckworth et al., 2019; Roberts et al., 2014). Moreover, the slightly smaller effects of extraversion, agreeableness, and openness are consistent with the notion that these effects may be linked to specific performance contexts, and thus less robust overall than conscientiousness (Puryear et al., 2017; Salgado & Táuriz, 2014). Finally, although initial meta-analyses suggested that neuroticism yields relatively large and durable effects that rival conscientiousness (e.g., Barrick & Mount, 1991; Salgado, 1997), our more comprehensive
synthesis suggests that neuroticism effects are generally smaller than conscientiousness, but comparable to the remaining Big Five traits.

Third, because prior meta-analyses had mostly focused on the association of personality with either academic or job performance, little was known about whether effect sizes varied across these fundamental performance categories. In contrast with the argument that personality—performance associations are similar across academic and job performance, we found striking differences. Specifically, whereas agreeableness (0.07 vs 0.11) and openness (0.14 vs 0.11) yielded similar effects across academic and job performance categories, conscientiousness yielded a stronger association with academic performance (0.28 vs 0.20). Further, extraversion (−0.01 vs 0.14) and neuroticism (−0.03 vs −0.15) yielded a much stronger association with job performance. These data suggest that personality—performance associations vary across performance categories in profound ways. The present research also uncovered a substantial number of meta-analyses examining personality—performance associations for outcomes that fall outside the traditional areas of academic and job performance (m = 15). Although each trait yielded a detectable effect when aggregating across these meta-analyses, openness yielded the largest effect (0.16), likely due to a focus on intellectual dimensions.

Fourth, and finally, we examined the replicability of personality—performance associations across three sets of partially overlapping meta-analyses in which the same personality—performance relation had been examined by largely independent research teams. In all three instances, we found that the pattern of personality—performance associations across meta-analyses was very similar. These data suggest that estimates of the association of Big Five traits with performance are replicable across meta-analyses even when these meta-analyses have significant differences, such as variations in study inclusion/exclusion criteria as well as variations in the calculation and conversion of effect sizes. Further, they suggest that personality—performance associations derived from prior meta-analyses reflect valid estimates of population effects, as opposed to effects that were biased by idiosyncratic decisions of any individual research team.

### 4.1 | Strengths and weaknesses

Having surveyed the vast terrain of the personality—performance literature, we now evaluate the strengths and weaknesses of this research area, and in doing so, identify pathways for future study. The major strength of this literature is its sheer size. The extremely large number of primary studies on the association of Big Five traits with performance allows definitive statements that are rare in the behavioral sciences. Moreover, the relatively large collection of meta-analyses on this topic facilitates integrative conclusions regarding personality—performance associations both overall and in specific contexts. Although scholars could simply compare one meta-analysis on job performance to another on academic performance, integrating all meta-analyses on these topics before comparison enables highly robust estimates of effect size as well as estimates of the degree to which effects fluctuate across meta-analyses.

The meta-analyses that we obtained generally evidenced substantial methodological rigor. That is, almost all of the meta-analyses conducted a careful search of the literature to identify a large set of potentially relevant studies. Further, most meta-analyses included both published and unpublished studies and corrected effect sizes to account for attenuation in measures of personality and performance. The personality—performance literature should also be lauded for its focus on an explicit, behavioral outcome, which is relatively rare in psychology and related fields (Baumeister et al., 2007).

Nonetheless, we identified several weaknesses that require careful attention in future research. Specifically, basic information about the included samples was often unreported. Thus, we recommend that future meta-analyses specify the gender, age, race-ethnicity, and country of the included samples, and examine whether these variables moderate effect sizes. Given the predominant focus on adults in European-American countries, an important task for future research will be to explore the universality of personality—performance associations as well as their developmental trajectory.

Additionally, several meta-analyses did not specify whether they examined self-ratings or other-ratings of personality. We assumed in these cases that researchers primarily examined self-ratings but recommend that future meta-analyses report this information more precisely and explore whether it moderates effects. Further, we recommend the public sharing of research databases on personality and performance. Such databases may enable a large meta-analysis of individual studies on personality—performance effects and would prevent sample overlap that can occur when researchers integrate multiple related meta-analyses.

Another important caveat is that, because personality—performance research incorporated in the present report was correlational in nature, it remains unclear whether personality has a causal impact on performance. Although personality may cause performance, it is also possible that performance trends initiate changes in personality or that third variables such as socioeconomic status account for both personality and performance. Emerging research...
suggesting that people can intentionally change their personality (Hudson & Fraley, 2015) could be profitably used to assess whether changes in personality result in changes in performance. Further, the mechanisms by which Big Five traits predict performance remain unclear. Additional study is needed to examine, for example, how conscientiousness results in enhanced performance across a variety of contexts (Gerhardt et al., 2007; Hampson, 2012) and why its effects are more pronounced in some contexts than others (e.g., academic settings). Finally, research should continue to examine personality—performance associations not just in academic and vocational contexts, but also in other contexts which have so far received less attention (e.g., sport performance; Allen & LaBorde, 2014).

4.2 | Conclusions

The present research integrated an enormous collection of studies, catalogued in over 50 meta-analyses, to demonstrate that Big Five personality traits yield robust associations with performance and to document how these associations fluctuate across Big Five traits and performance types. Through its synthesis of meta-analyses on personality and performance, the present report provides a comprehensive overview of this field. Nonetheless, future research is needed to examine potential variations in personality—performance associations across populations and to identify mechanisms that underlie personality effects. Additionally, better reporting of sample and methodological characteristics in future meta-analyses is recommended to advance this field. Thus, although much has been learned about the connection between personality and performance, many unanswered questions remain, which should stimulate researchers and theorists for decades to come.

ACKNOWLEDGMENTS

The authors thank Jason E. Strickhouser for assistance with the article search. The authors received no financial support for the research, authorship, and/or publication of this article.

AUTHOR CONTRIBUTIONS

Ethan Zell led data collection, data coding, data analysis, and manuscript writing. Tara L. Lesick contributed to data collection, data coding, and approved the manuscript.

ETHICS APPROVAL STATEMENT

This research involved a secondary analysis of de-identified data/archives and was therefore exempt from ethics approval.

ORCID

Ethan Zell https://orcid.org/0000-0002-8363-8418

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found in the online version of the article at the publisher’s website.

**How to cite this article:** Zell, E., & Lesick, T. L. (2021). Big five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality*, 00, 1–15. https://doi.org/10.1111/jopy.12683