Received: 26 February 2019 | Revised: 2 September 2020 | Accepted: 5 September 2020

DOI: 10.1111/ijsa.12314

FEATURE ARTICLE

International Journal of Selection and Assessment
WILEY

The relationship between cognitive ability and personality scores in selection situations: A meta-analysis

Michael Schilling | Nicolas Becker | Magdalena M. Grabenhorst | Cornelius J. König

Universität des Saarlandes, Saarbrücken, Germany

Correspondence
Michael Schilling, Arbeits- & Organisationspsychologie, Universität des Saarlandes, Campus A1 3, Saarbrücken 66123, Germany.
Email: michael.schilling@uni-saarland.de

Abstract

Several faking theories have identified applicants' cognitive ability (CA) as a determinant of faking—the intentional distortion of answers by candidates—but the corresponding empirical findings in the area of personality tests are often ambiguous. Following the assumption that CA is important for faking, we expected applicants with high CA to show higher personality scores in selection situations, leading in this case to significant correlations between CA and personality scores, but not in nonselection situations. This meta-analysis (66 studies, k = 115 individual samples, N = 46,265) showed this pattern of results as well as moderation effects for the study design (laboratory vs. field), the response format of the personality test, and the type of CA test.

KEYWORDS

Big Five, cognitive ability, faking, meta-analysis, personality, personnel selection, self-presentation

1 | INTRODUCTION

Faking—the intentional distortion of answers by applicants—is a frequently occurring phenomenon found when personality tests are used for personnel selection (e.g., Anglim et al., 2018; Birkeland et al., 2006; Galić et al., 2012; Griffin & Wilson, 2012). In this context, interindividual differences in faking behavior are particularly problematic, as they can affect the applicants' rank order and thus the validity of selection decisions (König et al., 2011; McFarland & Ryan, 2000; Mueller-Hanson et al., 2006; Raymark & Tafero, 2009). The majority of faking theories attribute these differences in applicants' faking to two factors in particular (Ellingson & McFarland, 2011; Marcus, 2009; McFarland & Ryan, 2006; Snell et al., 1999): (a) applicants' motivation to present themselves in a highly favorable way in order to improve their chances within the selection process, and (b) the abilities needed to manage the image they convey to the organization by distorting the answers in the required direction.

Regarding the abilities aspect, some authors (e.g., Johnson & Hogan, 2006; Kleinmann et al., 2011) even suggested that part of the criterion validity of personality tests may be attributed to the fact that such abilities necessary for faking are also of great relevance in today's working world. In line with this argument, several theoretical models have identified applicants' cognitive ability as a crucial determinant of the occurrence and magnitude of faking behavior (e.g., Marcus, 2009; Snell et al., 1999; Tett & Simonet, 2011). However, previous empirical results were inconclusive, while a substantial proportion of studies found a corresponding effect (Grubb & McDaniel, 2007; Levashina et al., 2014; Pauls & Crost, 2005), others did not (Furnham et al., 2008; Levashina et al., 2009; Mudgett, 2000; Schilling et al., 2020). Not only are results inconclusive, it is also unclear why there are such inclusive results. For example, faking of personality tests has been studied in the field and in the lab, and both research strategies have their advantages and disadvantages that could also matter for the relationship of cognitive ability (e.g., van Hooft & Born, 2012; Ryan et al., 1998). Similarly, some researchers have used Likert scaled personality tests, whereas others have used forced-choice tests (e.g., Hausknecht, 2010; MacKenzie et al., 2010), and cognitive abilities have been measured with different kinds of operationalizations (e.g., verbal vs. nonverbal cognitive ability tests, e.g., Hale & Padgett, 2014; Klehe et al., 2012). Thus, it remains open how much which cognitive ability is related to faking and what variables moderates this relationship. An answer to these questions will help to provide a better understanding of the phenomenon of faking and thus of the consequences of faking for the construct and criterion validity of personality tests.

The goal of the current study was, therefore, to provide aggregated results concerning the relationship of participants' cognitive ability and their faking in personality tests during selection situations. Given that only a minority of studies directly report a corresponding effect size, our meta-analysis focused on the correlation between cognitive ability and personality scores. We compared this meta-analytic correlation between selection samples and nonselection samples because it can be found in different kinds of studies, irrespective of whether they were conducted in the laboratory or in the field or how they operationalized faking. In the following, we explain how the relationship between cognitive ability and faking should affect corresponding correlations, and we introduce possible moderating variables in this context.

1.1 | Personality tests in selection and cognitive ability

The majority of faking theories have agreed that successful faking requires not only the motivation to present oneself in a highly favorable way, but also the ability to behave and answer appropriately (Levashina et al., 2009; Marcus, 2009; Mueller-Hanson et al., 2006; Roulin et al., 2016; Snell et al., 1999). Applicants, for instance, have to identify the expectations of hiring organizations (i.e., the criteria on which they are being assessed) to subsequently show self-benefitting behavior (Kleinmann et al., 2011; König et al., 2006). For personality tests, this means that applicants could try to figure out which pole of a personality test item the hiring organizations consider the positive one so that they can move their response toward this pole (Klehe et al., 2012). Applicants with higher cognitive ability should handle this mainly analytical task more easily, which may, in turn, lead to more faking behavior among these applicants (Tett & Simonet, 2011).

However, when empirically analyzing the relationship between applicants' cognitive ability and their faking in personality tests in the form of a meta-analysis, the diversity of previous studies posed a particular challenge. Thus, large differences in the study design and in the operationalization of faking posed hurdles that prevent a direct aggregation of the previous study results. Part of the research in this area consisted of field studies which examined faking only at a group level, thus not allowing for the calculation of individual-level correlations between cognitive ability and faking (e.g., MacKenzie et al., 2010). The other part of the studies were laboratory studies, which predominantly used a within-person design, where the participants took the personality test twice, once honest, and once with the instruction to respond as an applicant. However, even within this group of laboratory studies, aggregation of results turned out quite hard as the operationalizations of faking differed substantially (for a detailed overview of different operationalizations of faking, see Burns & Christiansen, 2011). Thus, some studies operationalized faking as the difference between the participant personality scores in both conditions (e.g., Peterson et al., 2009), others measure faking as the within-person correlation on item level (e.g., Mersman & Shultz, 1998) or modeled faking as a latent factor in a structural equation model, which loads on all personality dimensions under selection condition but not under honest condition (e.g., Wrensen & Biderman, 2005). As a result, only a small fraction of the laboratory studies report the direct relationship between cognitive ability and faking required for meta-analytic calculations (for instance, less than 20 percent of the laboratory studies included in this meta-analysis reported a corresponding effect-size). In addition, there were also a number of laboratory studies that did not use within-person design and instead measured faking with social desirability or impression management scales (e.g., Robie et al., 2010).

In order to aggregate the results of as many of these studies as possible, our analysis is based on what, we believe, is the most basic indicator of the corresponding relationship: the correlation between applicants' cognitive ability and their personality scores in selection situations. If cognitive ability is a determinate of faking, it should also be a predictor of personality scores in selection situations. Thereby, the necessary correlations were provided in almost all studies, whether they were field or laboratory studies, whether it was a within or between-person design, whether faking was operationalized as difference between personality scores under two conditions, as a score on an impression management scale or was modeled as a latent variable in a structural equation model. Finally, we will compare our results with those from samples in which the personality test was not completed under the pressure of a selection situation (abbreviated asnonselection samples) so that we can conclude the relationship between cognitive ability and faking. Previous meta-analyses without a focus on selection situations showed no, or rather low, correlations between cognitive ability and personality scores (Lange, 2013; Poropat, 2009).[1] Assuming that applicants with higher cognitive ability are more successful at faking (e.g., Marcus, 2009; McFarland & Ryan, 2006; Snell et al., 1999) and that successful faking usually leads to higher scores on personality tests, we expect to find higher correlations between cognitive ability and personality scores in selection samples than in nonselection samples.

Hypothesis 1 *Correlations between cognitive ability and personality scores are higher in selection samples than in nonselection samples.*

Based on the diversity of faking research, there is much to suggest that the relationship between cognitive ability and personality

TABLE 1 English search terms

Cognitive ability	Personality	Selection	Faking
Cognitive abilit[a]	Personality	Applica[a]	Fak[a]
Cognitive skill[a]	Integrity	Selection	Malinger
Mental abilit[a]	Emotional intelligence	Incumbent	Cheat
Mental skill[a]	Emotional quotient	Assessment	Self-presentation
Intelligence	Self-report	Hire	Impression Management
Reasoning	Self-description	Candidate	Distort[a]
IQ	Noncognitive	Allocation	Self enhanc[a]
Raven	Big 5	Assignment	Ideal employee
Wonderlic	Big Five	Job	ATIC
GMA	Hexaco		Identify criteria
	NEO		
	IPIP		

Wildcard.[a]

also varies systematically. In the following sections, we introduce three meta-analytic moderators, which address the diversity of study designs, differences in the personality tests used, and differences in the type of cognitive ability tests employed.

2 | MODERATOR HYPOTHESES

2.1 | Study design (laboratory vs. field)

In general, most studies in the area of faking research can be assigned to two categories (Birkeland et al., 2006): (a) studies conducted in the field, with real applicants in actual selection situations and (b) studies conducted in the laboratory with participants who are put in a simulated selection situation or who are instructed to fake. Previous meta-analyses about faking in field and laboratory studies found significantly higher faking effects in the laboratory than in field studies (Birkeland et al., 2006; Hooper, 2007). Regarding this difference, some authors have argued that the processes underlying faking likely differ between the two types of studies (Ones et al., 1996): In field studies, the applicants' motivation to fake depends on many individual factors, including subjective considerations and situational circumstances. In laboratory studies, the faking motivation arises rather from the concrete instruction or from the cover story that is used to induce the application situation; this should lead to a similarly high faking motivation for all participants, which, in turn, may lead to individual differences in the ability to convert this motivation into faking behavior becoming more evident. In this line, we, therefore, expect higher correlations between cognitive ability and personality scores in laboratory studies than in field studies.

Hypothesis 2 *Correlations between cognitive ability and personality scores are higher in laboratory studies than in field settings.*

2.2 | Type of personality test

There are two main types of personality tests used in personnel selection (Vasilopoulos et al., 2006): (a) forced-choice personality tests, in which participants have to choose between statements representing different personality dimensions for each single item, and (b) single-stimulus personality tests, in which each item belongs to one personality dimension, for which the participants have to express their rejection or approval or something in between. Forced-choice tests can be considered as fairly robust against faking (Cao & Drasgow, 2019; Martin et al., 2002), mainly because it should be more difficult to answer in a socially desirable manner if one item includes two equally desirable dimensions (Vasilopoulos et al., 2006). In this case, applicants who are motivated to fake are faced with the task of determining which of the corresponding dimensions is most relevant for a future employer. This analytical task is difficult because in contrast to the single-stimulus personality tests, the social desirability of the items provides the applicants with fewer hints for successful faking. Given this increased difficulty with regard to faking in forced-choice personality tests, cognitive ability should be even more important when this type of test is used. Therefore, we expect higher correlations between cognitive ability and personality scores in samples completing forced-choice tests than in samples completing single-stimulus tests.

Hypothesis 3 *Correlations between cognitive ability and personality scores are higher in studies employing forced-choice personality tests than in studies employing single-stimulus personality tests.*

2.3 | Type of cognitive ability test

The type of cognitive ability test is also a potential moderator of the correlation between cognitive ability and personality scores. Previous studies showed higher correlations between verbal cognitive ability and faking than between nonverbal cognitive ability and faking (Grieve & Mahar, 2010; MacCann, 2013). The authors of these studies argued that a deeper understanding of the items is beneficial for effective faking, which underlines the importance of verbal cognitive ability. Following MacCann (2013) as well as Grieve and Mahar (2010), we thus expect higher correlations in samples completing verbal cognitive ability tests than in samples completing nonverbal cognitive ability tests.

Hypothesis 4 *Correlations between cognitive ability and personality scores are higher if cognitive ability is measured with verbal than with nonverbal cognitive ability tests.*

Studies that cannot be classified into one of the aforementioned categories will be summarized in a *mixed* category. We have no further hypotheses regarding this mixed category.

3 | METHODS

3.1 | Literature search

Four strategies were used to identify studies for this meta-analysis: (a) We conducted an extensive literature search using the databases Scopus, Science Direct, PsycINFO, PsycArticles, ERIC, EconLit, PSYNDEX, MEDLINE, Web of Science, and ProQuest Dissertations and Theses Database. Search queries were constructed from four lists containing broad search terms for cognitive ability (tests), personality (tests), selection, and faking (see Table 1 for full lists) and were carried out in English and German. The literature search was conducted over the course of 2016 and yielded 610 published articles, dissertations, book chapters, and unpublished reports. (b) Additionally, we searched in the program booklets of the annual conferences of the Society for Industrial and Organizational Psychology (SIOP) for abstracts containing the four key topics mentioned above (for the years 1999–2015). A further 58 studies were detected with this approach. (c) After applying our inclusion and exclusion criteria (which we will describe in the next section) to our results up to this point, we contacted 32 experts by e-mail, with experts being defined as study authors or dissertation supervisors of at least three studies. These authors were asked whether they knew of other studies that may meet our criteria. In addition, we asked this question to all 26 authors of studies that lacked the information needed for our meta-analysis (this group of authors overlaps with the expert group). This approach resulted in another three studies. (d) Finally, we conducted a backward search based on the bibliographies of the studies found so far and of the publication lists of the aforementioned 32 experts; this delivered 30 additional studies.

3.2 | Inclusion and exclusion criteria

To be included in our current meta-analysis, studies had to meet the following criteria: (a) Studies had to include some kind of selection situation. This was the case for field studies with a real selection situation (e.g., applying for a job or to a university) or laboratory studies with a simulated selection situation (e.g., induced by an instruction such as 'Imagine you are applying for a job as a…'). This criterion was also applied during our search for control data from nonselection samples. Even though this approach drastically limited the number of corresponding samples, it was the only practicable way to follow a uniform and consistent search strategy. (b) There had to be some motivation for the participants to fake and present themselves in a favorable way. Accordingly, we excluded field studies in which it was clear to the participants that the personality test score would not be used for selection purposes (e.g., Merkulova et al., 2014) and studies in which it was unclear whether the participants would be motivated to present themselves favorably (e.g., in a compulsory military service examination; Boss et al., 2015). Furthermore, we excluded laboratory studies in which some tests were filled out under selection conditions but the personality test was not (e.g., Peeters & Lievens, 2005). Moreover, we excluded studies that measured faking solely as overclaiming (e.g., Ackerman & Ellingsen, 2014) or as a fraud in objective tests (e.g., Wright et al., 2014). (c) Personality had to be measured by self-report, and the personality scales must belong (or at least be assignable to) the Five-Factor model of personality. (d) Studies had to include some objective measurement of cognitive ability in the form of an intelligence or ability test. We also included studies reporting college admission test scores, for example from the Scholastic Aptitude Test (SAT) or American College Testing (ACT) as both have been found to be valid measures of cognitive ability (Frey & Detterman, 2004; Jackson & Rushton, 2006; Koenig et al., 2008). However, studies which only reported academic achievements, such as the grade point average (GPA), were excluded as these measurements are rather considered to be the outcome of cognitive ability, personality, and other factors and not as a measure of cognitive ability itself (e.g., Komarraju et al., 2013; Poropat, 2009). (e) Furthermore, studies had to report the correlation between personality test scores and cognitive ability, as we used this correlation as the effect size. If the latter precondition was not met, we contacted the author(s) and requested the corresponding data. (f) In a final step, we carried out a sensitivity analysis, comparing the results with and without the largest five percent of the samples. In this way, we identified seven individual samples (e.g., Arthur et al., 2014; De Fruyt et al., 2006; Levashina et al., 2014), that were at least three times larger than the largest remaining study and at least 40 times larger than the average sample size of the remaining studies. The seven samples consisted exclusively of field data from personnel selection providers or were directly taken from the personnel selection of large companies or public authorities. Thereby, the meta-analytic calculations with and without these large samples showed very similar results pattern leading to the same conclusions. For instance, the true score correlation between cognitive ability and Conscientiousness in field studies was $\rho = 0.079$ based on all samples and $\rho = 0.084$ when the large samples were excluded. However, the results obtained including these large studies corresponded to the values from these very large studies down to the 3rd decimal place. However, the results including these large samples were, except for the 3rd decimal place, identical to the results obtained by performing the meta-analytical calculations only on the basis of these seven studies. Since the inclusion of these samples would have completely overshadowed the meta-analytical results, we excluded these results from our analysis to avoid overemphasizing these single sample effect sizes. Tables S3–S6 of the Supporting Information show the meta-analytical results if these samples are not excluded, Table S7 gives a more detailed overview of the excluded samples.

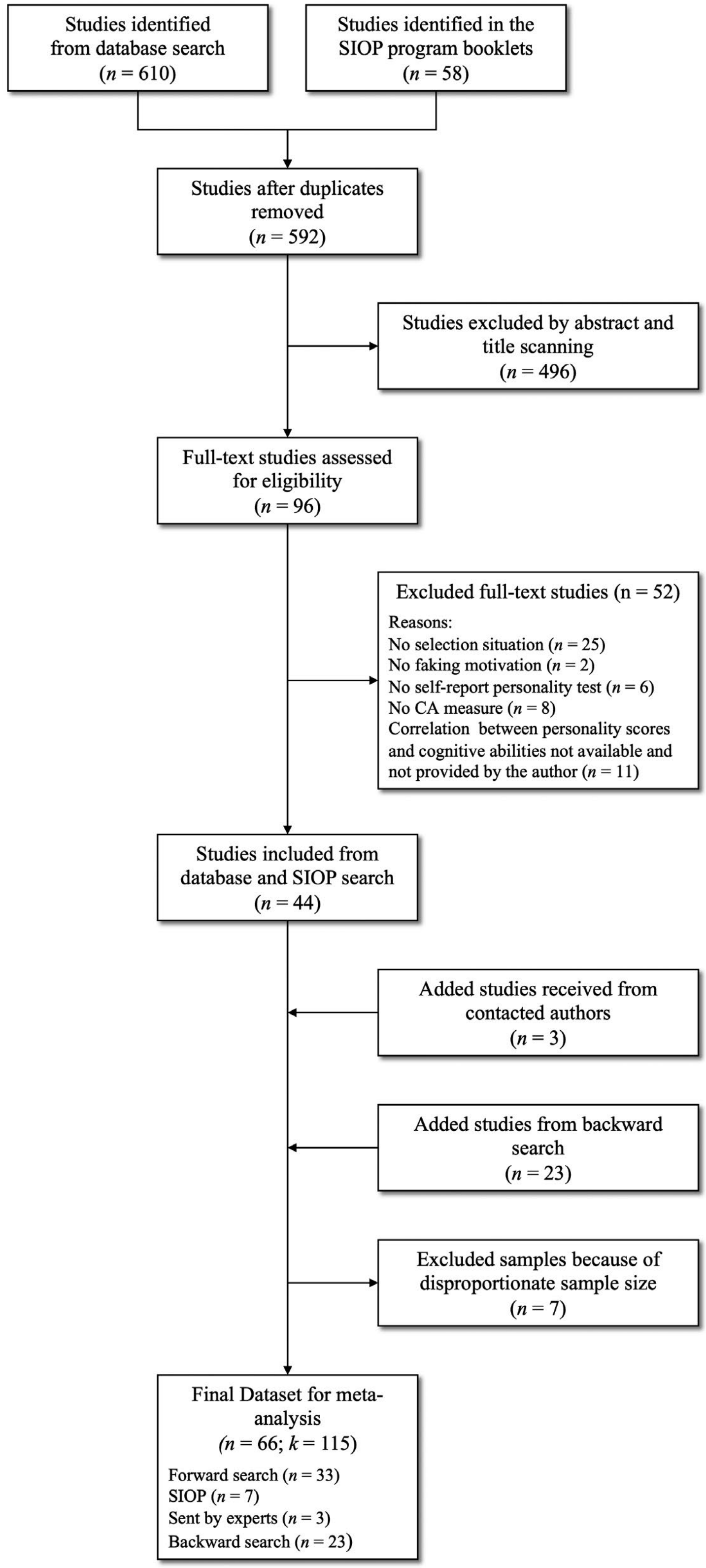

FIGURE 1 Flowchart showing the process of identifying and selecting studies

3.3 | Final data set

The final data set consisted of 66 studies with 115 independent samples, 825 effect sizes (in total for all dimensions of the Five-Factor model of personality), and a total N = 46,265 participants (N = 2,289,508 without excluding the three disproportionately large studies, see exclusion criteria). Of these 66 studies, 33 came from the forward database search, and seven were SIOP conference articles. Authors from our expert list made us aware of three further studies, and another 23 arose during the backward search. The oldest included study was published in 1957; the most recent studies were conducted in 2016. The entire process up to this final data set is shown in Figure 1. Table S1 of the Supporting Information gives an overview of all studies included, Table S2 gives an overview of the resulting independent samples.

3.4 | Coding of studies

Personality scales not based on the Five-Factor model of personality were grouped into the model based on the work of Salgado and Táuriz (2014). If a specific dimension was not mentioned in their overview, we used a strategy developed by Barrick and Mount (1991). Five raters, all psychology graduates (three with a PhD and two with a Master's degree or equivalent) categorized the leftover dimensions into the Five-Factor model of personality. The classification had to be accomplished with a 75% majority; abstentions were not counted. If a scale could not be clearly classified, we excluded it from our analysis. For the purpose of simplification and better interpretation, *Neuroticism* was reverse-coded as *Emotional stability*.

We distinguished three types of cognitive ability tests: verbal tests (e.g., the Word Comprehension subtest of the Wechsler Adult Intelligence Scale; Wechsler, 2014), nonverbal tests (e.g., Raven's Progressive Matrices; Raven, 1938), and mixed tests (e.g., Wonderlic Personnel Test; Wonderlic, 1996). Categorization occurred primarily according to the information provided in the corresponding article and was carried out by two raters independently (both with a Master's degree or equivalent). If the authors did not provide the relevant information in the article, the categorization was conducted with the help of the test manuals. To be able to compare effect sizes in studies that provided correlations for verbal and nonverbal cognitive ability tests, we calculated—as far as all required data were available—the composite scores and the corresponding reliabilities according to Schmidt and Hunter (Schmidt & Hunter, 2014, p. 442). If the required correlations between the variables that should be aggregated were not documented, we calculated the arithmetic mean using Fisher's *Z*-values.

Some studies also reported several independent correlations between the variables of interest for a single sample (e.g., there were two or more correlations between personality scales that were categorized as the same personality dimension and the cognitive ability measurement, or studies provided only the correlation of two or more cognitive ability subtests that were both categorized as verbal/nonverbal). In these cases, we also used the aggregation procedure laid out in the last paragraph.

3.5 | Meta-analytic procedures

We followed the procedures for psychometric meta-analysis described by Schmidt and Hunter (2014). Mean correlations between cognitive ability and personality dimensions were estimated by sample size-weighted individual correlation coefficients (see equation 3.1 in Schmidt & Hunter, 2014, p. 95). These 'bare bones' correlations are comparable with the results from methods in the tradition of Hedges and colleagues (Hedges & Vevea, 1998). Furthermore, psychometric meta-analysis provides the option to correct for the unreliability of measurement scales and range restriction, yielding the population correlation ρ. As not all studies reported the required information, we were unable to correct correlations individually and thus used artifact distribution meta-analysis (Schmidt & Hunter, 2014) instead. Unreliability of measurement scales was corrected for cognitive ability (the predictor) and for the personality scales (the criterion). We also corrected for the indirect range restriction of cognitive ability, since many samples may have already been preselected on the basis of cognitive ability or related constructs (e.g., students in laboratory studies who are selected based on their Scholastic Assessment Test). The artifact information was extracted from the associated studies, if reliabilities were not documented we took the information from the corresponding test manuals of the cognitive ability respectively personality tests. Depending on the analysis, the artifact distributions were based on reliability information for 0.0%–50.0% of the coefficients for the cognitive ability tests and 33.3%–93.2% of the coefficients for the personality scales. Information on indirect range restriction was available for 0.0%–61.5% of the samples. Tables 3–6 provide additional information about the average reliabilities, average range restriction, corresponding variances, and number of included coefficients for each artifact distributions used in the meta-analytical calculation. If no information at all was available for a meta-analytical calculation regarding predictor reliability, criterion reliability or range restriction, the corresponding artifact distribution was specified as a uniform distribution with a mean of 1.00 and no variance (Schmidt & Hunter, 2014). In such a case, the corresponding aspect could not be corrected in the meta-analytic calculation. For the meta-analytic calculations, we used the metafor package (Viechtbauer, 2010) in R 3.3.3 (R Core Team, 2017) and the Schmidt and Le meta-analysis program (Schmidt & Le, 2005).

We report 80% credibility intervals around ρ to provide analysis of the homogeneity of the corrected effect sizes as well as the percentages of the variance in effect size explained by artifacts (Schmidt & Hunter, 2014). In this regard, based on the '75% rule' (Schmidt & Hunter, 2014), less than 75% variance reduction by artifact correction indicates the presence of additional moderators. For moderator analysis, we calculated 95% confidence intervals around ρ to locate meaningful moderating effects (Schmidt & Hunter, 2014; Whitener, 1990), using the formula reported by Whitener (1990).

In order to test the robustness of our findings, we calculated fail-safe *N*s as the number of null results that would have to be added to the studies in our data set to reduce the meta-analytic outcome to a trivial average effect size (Orwin, 1983). Following the recommendations of Schmidt and Hunter (2014) as well as McNatt (2000), we regarded correlations of $r = 0.05$ and below as trivial. For additional analysis of file drawer bias (Light & Pillemer, 1984), we created funnel plots of the included effect sizes for all of our meta-analytic calculations using the R-package metafor (Viechtbauer, 2010). These funnel plots were

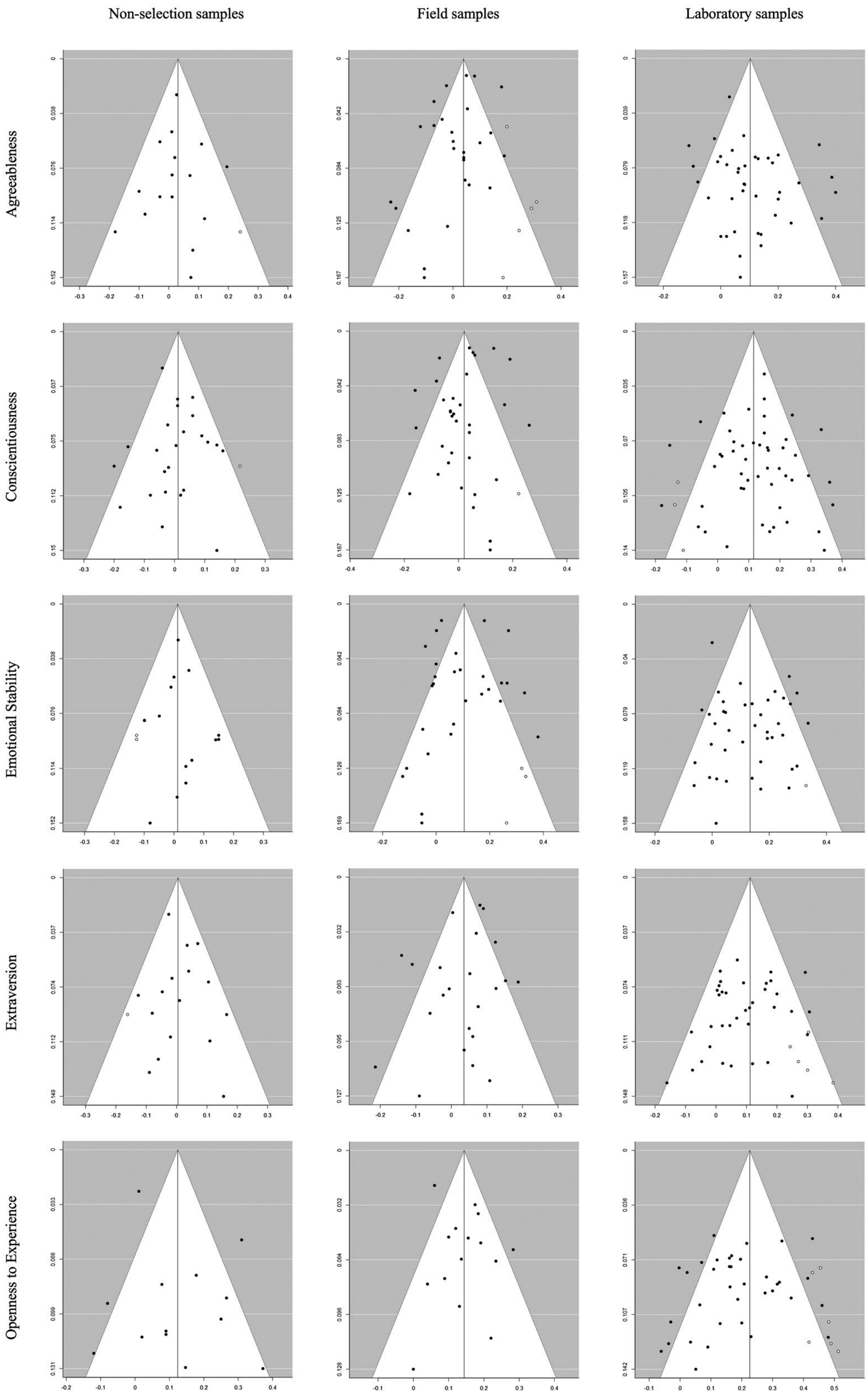

FIGURE 2 Funnel plots for nonselection samples, field samples, and laboratory samples based on Bare Bones analysis. Correlations employed in the analysis are represented by black dots, correlations complemented by the trim and fill method are represented as white dots. The 95%-standard error of the mean correlation is represented as a dotted line

adjusted for missing studies using the trim and fill method (Duval & Tweedie, 2000). All plots were relatively symmetrical, indicating that our meta-analysis did not seem to prioritize the consideration of statistically significant effects over nonsignificant effects. As an example, Figure 2 shows the funnel plots for the nonselection samples, field samples, and laboratory samples.

4 | RESULTS

4.1 | Validation of the representativeness of our control data

First, we checked whether the meta-analytic outcome for our nonselection samples corresponds with the findings of other meta-analyses on the relationship between personality and cognitive ability (i.e., with the results of Lange; Lange, 2013; Poropat, 2009). Table 2 shows the results of this comparison. As can be seen, the confidence intervals for all five dimensions show substantial overlap, indicating that our meta-analytic data for nonselection samples replicate the current state of research.

4.2 | Main analysis

Table 3 shows the meta-analytic results regarding the differences between selection and nonselection situations. The overall true-score correlations between cognitive ability and Five-Factor personality test scores in the selection samples were (nonselection correlations in parentheses for direct comparison): $\rho_{\text{Conscientiousness}} = 0.130$ (0.001), $\rho_{\text{Emotional stability}} = 0.181$ (0.041), $\rho_{\text{Agreeableness}} = 0.105$ (0.055), $\rho_{\text{Extraversion}} = 0.117$ (0.001), and $\rho_{\text{Openness to experience}} = 0.262$ (0.158). Confidence intervals did not include zero for any of the dimensions, and for all dimensions, the confidence intervals did not overlap with the corresponding confidence intervals of the nonselection samples.

TABLE 2 Comparison of our nonselection samples with already published meta-analysis

	Poropat (2009)	Lange (2013)	Nonselection samples
Agreeableness			
ρ	−0.019	0.01	0.010
95% CI	[−0.052 to 0.014]	[−0.02 to 0.04]	[−0.011 to 0.032]
Conscientiousness			
ρ	0.002	−0.04	−0.019
95% CI	[−0.058 to 0.062]	[−0.08 to 0.00]	[−0.037 to −0.001]
Emotional stability			
ρ	0.039	0.09	0.058
95% CI	[0.018 to 0.061]	[0.07 to 0.11]	[0.037 to 0.080]
Extraversion			
ρ	−0.012	−0.02	−0.005
95% CI	[−0.040 to 0.016]	[−0.06 to 0.02]	[−0.026 to 0.016]
Openness to experience			
ρ	0.125	0.19	0.208
95% CI	[0.100 to 0.151]	[0.16 to 0.22]	[0.186 to 0.229]

Note: The results of Poropat are not contents of the corresponding paper but were provided to us by the author on request.
Abbreviation: 95% CI, 95%-confidence interval [lower bound to upper bound].

As assumed in Hypothesis 1, significantly higher correlations were found in the selection samples for the remaining four dimensions. In the selection samples, only 29%–62% of the variance in effect size was explained by artifacts, which indicates that further moderator effects are likely (Schmidt & Hunter, 2014).

4.3 | Moderator analysis 1: Study design (laboratory vs. field)

Table 4 summarizes the meta-analytic results for the relationship between cognitive ability and the Five-Factor dimensions separately for nonselection samples, samples derived from field studies, and samples derived from laboratory studies. For all five dimensions, the same order of moderator levels emerged: True-score correlations were smallest for nonselection samples (ρ = 0.001–0.151), slightly higher in field studies (ρ = 0.078–0.206), and clearly higher in laboratory studies (ρ = 0.161–0.350). For Conscientiousness, Emotional Stability, and Extraversion, none of the moderator-level confidence intervals overlapped. For Agreeableness and Openness to Experience, the confidence intervals did not overlap for nonselection and laboratory studies but did overlap for nonselection and field studies. In summary, a moderation by the study design, as assumed in Hypothesis 2, was found for all dimensions of the Five-Factor model of personality. For both groups of studies, laboratory and field, the percentage of reduced variance in effect size through artifacts did not meet the traditional 75% criterion, indicating further moderator effects.

4.4 | Moderator analysis 2: Type of personality test

Table 5 presents our meta-analytical results separately for the use of forced-choice and single-stimulus personality tests. Correlations for single-stimulus tests were very similar to the results for selection situations of the main analysis. Results for forced-choice tests showed very wide confidence intervals for Agreeableness, Extraversion, and Openness to Experience (each including the corresponding confidence interval on the single-stimulus moderator level). True-score correlations were higher for forced-choice tests than for single-stimulus tests when Conscientiousness was measured but were lower when Emotional Stability was measured. In both cases, confidence intervals were not overlapping. In summary, the results only partially support Hypothesis 3, but it must also be noted that the analysis for the moderator level of the forced-choice tests was based on a small number of samples and participants: In contrast to the previous analyses, the results are based on only three to seven samples and at most 1,375 participants. Moreover, only for the dimension Conscientiousness one of the included studies provided information about the reliability of the forced-choice personality scale, accordingly, the calculations could not be corrected for the unreliability of the forced-choice personality scales. For the forced-choice tests, the percentage of reduced variance in effect size through artifacts

TABLE 3 Main meta-analytic results

			Artifact distribution									Bare-bones eta-analysis		Meta-analysis with full correction for artifacts							
	k	N	r_{xx}	SD_{rxx}	n_{xx}	r_{yy}	SD_{ryy}	n_{yy}	U_x	SD_{Ux}	n_{Ux}	r	SD_r	ρ	σ	VAR (%)	80% CE-lb	80% CE-ub	95% CI-lb	95% CI-ub	N_{FS}
Agreeableness																					
nss	16	3,936	0.84	0.09	4	0.82	0.06	13	0.71	0.13	6	0.026	0.062	0.055	0.000	108	0.055	0.055	0.024	0.086	2
ss	69	28,346	0.81	0.10	10	0.79	0.07	56	0.78	0.13	16	0.055	0.084	0.105	0.123	38	−0.053	0.263	0.094	0.117	77
Conscientiousness																					
nss	24	5,704	0.86	0.09	6	0.81	0.10	16	0.74	0.15	8	0.000	0.071	0.001	0.054	84	−0.069	0.070	−0.025	0.027	−23
ss	85	38,109	0.81	0.09	14	0.83	0.11	61	0.76	0.12	22	0.068	0.091	0.130	0.141	31	−0.050	0.310	0.120	0.140	136
Emotional stability																					
nss	14	3,758	0.86	0.10	3	0.83	0.09	11	0.72	0.18	5	0.021	0.051	0.041	0.000	144	0.041	0.041	0.009	0.073	−2
ss	71	28,785	0.81	0.10	10	0.83	0.09	53	0.78	0.13	17	0.099	0.105	0.181	0.159	29	−0.022	0.384	0.170	0.192	187
Extraversion																					
nss	17	4,358	0.86	0.09	6	0.86	0.05	9	0.74	0.16	6	0.001	0.068	0.001	0.049	84	−0.062	0.065	−0.028	0.031	−16
ss	69	20,376	0.80	0.10	11	0.83	0.10	49	0.77	0.13	16	0.061	0.079	0.117	0.097	59	−0.006	0.241	0.104	0.131	93
Openness to experience																					
nss	14	3,039	0.86	0.10	3	0.74	0.11	13	0.72	0.15	8	0.078	0.116	0.158	0.180	38	−0.073	0.388	0.123	0.192	31
ss	58	14,044	0.82	0.09	8	0.79	0.09	47	0.77	0.13	17	0.140	0.093	0.262	0.102	62	0.131	0.393	0.247	0.277	246

Abbreviations: CE, credibility interval; CI, confidence interval; lb, lower bound; N_{FS}, fail-safe N; nss, nonselection situations; r_{xx}, average reliability of predictor; r_{yy}, average reliability of criterion; ss, selection situations; ub, upper bound; U_x, average restriction in variance; VAR (%), part of variance that can be attributed to artifacts.

TABLE 4 Moderator analysis: Nonselection situations vs. field studies vs. laboratory studies

			Artifact distribution									Bare-bones Meta-analysis		Meta-analysis with full correction for artifacts							
	k	N	r_{xx}	SD_{rxx}	n_{xx}	r_{yy}	SD_{ryy}	n_{yy}	U_x	SD_{Ux}	n_{Ux}	r	SD_r	ρ	σ	VAR (%)	80% CE-lb	80% CE-ub	95% CI-lb	95% CI-ub	N_{FS}
Agreeableness																					
nss	16	3,936	0.84	0.09	4	0.82	0.06	13	71	0.13	6	0.026	0.062	0.055	0.000	108	0.055	0.055	0.024	0.086	2
field	28	21,713	0.86	0.11	3	0.78	0.08	19	0.82	0.20	3	0.048	0.072	0.078	0.097	29	−0.046	0.202	0.065	0.091	16
lab	41	6,633	0.79	0.10	7	0.80	0.06	37	0.77	0.12	13	0.079	0.112	0.161	0.150	54	−0.031	0.352	0.137	0.184	91
Conscientiousness																					
nss	24	5,704	0.86	0.09	6	0.81	0.10	16	0.74	0.15	8	0.000	0.071	0.001	0.054	84	−0.069	0.007	−0.025	0.027	−23
field	35	29,227	0.84	0.08	6	0.82	0.12	20	0.82	0.20	3	0.051	0.080	0.084	0.112	22	−0.006	0.227	0.072	0.095	24
lab	50	8,882	0.79	0.09	8	0.83	0.11	41	0.75	0.11	19	0.122	0.103	0.249	0.122	63	0.093	0.405	0.229	0.268	199
Emotional stability																					
nss	14	3,758	0.86	0.10	3	0.83	0.09	11	0.72	0.18	5	0.021	0.051	0.041	0.000	144	0.041	0.041	0.009	0.073	−2
field	30	22,202	0.86	0.11	3	0.79	0.09	17	0.82	0.20	3	0.096	0.103	0.153	0.145	19	−0.033	0.339	0.140	0.166	62
lab	41	6,583	0.79	0.10	7	0.85	0.08	36	0.77	0.11	14	0.109	0.111	0.212	0.137	57	0.037	0.387	0.189	0.235	133
Extraversion																					
nss	16	4,248	0.86	0.09	6	0.86	0.05	9	0.74	0.16	6	0.006	0.061	0.011	0.000	102	0.011	0.011	−0.019	0.041	−12
field	23	14,387	0.83	0.10	4	0.78	0.13	14	0.82	0.20	3	0.052	0.068	0.089	0.088	40	−0.024	0.202	0.072	0.105	18
lab	37	4,864	0.79	0.10	7	0.84	0.08	34	0.76	0.12	13	0.094	0.100	0.192	0.079	84	0.091	0.292	0.165	0.219	105
Openness to experience																					
nss	13	2,929	0.86	0.10	3	0.75	0.11	12	0.72	0.15	8	0.075	0.117	0.151	0.185	36	−0.085	0.388	0.116	0.187	27
field	15	6,809	0.86	0.16	2	0.76	0.10	13	0.82	0.20	3	0.125	0.065	0.206	0.046	80	0.147	0.266	0.184	0.229	47
lab	36	4,653	0.81	0.08	6	0.80	0.09	33	0.76	0.12	14	0.184	0.125	0.350	0.137	61	0.175	0.525	0.325	0.374	217

Abbreviations: CE, credibility interval; CI, confidence interval; field, field studies; lab, laboratory studies; lb, lower bound; N_{FS}, fail-safe *N*; nss, nonselection situations; r_{xx}, average reliability of predictor; r_{yy}, average reliability of criterion; ub, upper bound; U_x, average restriction in variance; VAR (%), part of variance that can be attributed to artifacts.

TABLE 5 Moderator analysis: Single-stimulus vs. forced-choice personality tests on level selection situations

	k	N	Artifact distribution									Bare-bones Meta-analysis		Meta-analysis with full correction for artifacts							
			r_{xx}	SD_{rxx}	n_{xx}	r_{yy}	SD_{ryy}	n_{yy}	U_x	SD_{Ux}	n_{Ux}	r	SD_r	ρ	σ	VAR (%)	80% CE-lb	80% CE-ub	95% CI-lb	95% CI-ub	N_{FS}
Agreeableness																					
FC	3	524	1.00	0.00	0	1.00	0.00	2	1.00	0.00	0	0.091	0.092	0.092	0.052	68	0.026	0.158	0.007	0.175	3
SS	66	27,822	0.81	0.010	10	0.79	0.07	55	0.78	0.13	16	0.054	0.083	0.103	0.123	38	−0.054	0.260	0.092	0.115	71
Conscientiousness																					
FC	7	1,375	1.00	0.00	1	0.62	0.10	4	0.68	0.05	3	0.174	0.098	0.310	0.114	53	0.165	0.456	0.263	0.356	37
SS	78	36,734	0.81	0.09	13	0.84	0.10	57	0.77	0.13	19	0.065	0.089	0.120	0.135	30	−0.054	0.293	0.110	0.130	109
Emotional stability																					
FC	4	660	1.00	0.00	0	0.66	0.07	2	1.00	0.00	1	0.074	0.100	0.091	0.076	61	−0.006	0.188	0.015	0.166	4
SS	67	28,125	0.81	0.10	10	0.84	0.08	51	0.78	0.13	16	0.099	0.105	0.180	0.158	28	−0.022	0.382	0.169	0.191	174
Extraversion																					
FC	4	660	1.00	0.00	0	0.75	0.17	2	1.00	0.00	1	0.061	0.090	0.071	0.052	75	0.004	0.138	−0.005	0.146	2
SS	65	19,716	0.80	0.10	11	0.83	0.10	47	0.77	0.14	15	0.060	0.080	0.115	0.098	57	−0.011	0.240	0.101	0.128	85
Openness to experience																					
FC	3	440	1.00	0.00	0	1.00	0.00	1	1.00	0.00	1	0.226	0.091	0.226	0.044	75	0.168	0.283	0.140	0.309	11
SS	55	13,604	0.82	0.09	8	0.80	0.09	46	0.78	0.13	16	0.137	0.092	0.251	0.100	62	0.123	0.379	0.235	0.266	221

Note: The high mean reliabilities and *u* values up to 1.00 and the associated low or nonexistent variance for forced-choice test are due to the very limited or nonexisting information on corresponding coefficients for the samples of these meta-analytical calculations.

Abbreviations: CE, credibility interval; CI, confidence interval; FC, forced-choice personality tests; lb, lower bound; N_{FS}, fail-safe *N*; r_{xx}, average reliability of predictor; r_{yy}, average reliability of criterion; U_x, average restriction in variance; SS, single-stimulus personality tests; ub, upper bound; VAR (%), part of variance that can be attributed to artifacts.

TABLE 6 Moderator analysis: Mixed vs. nonverbal vs. verbal cognitive ability tests

			Artifact distribution									Bare-bones meta-analysis		Meta-analysis with full correction for artifacts							
	k	N	r_{xx}	SD_{rxx}	n_{xx}	r_{yy}	SD_{ryy}	n_{yy}	U_x	SD_{Ux}	n_{Ux}	r	SD_r	ρ	σ	VAR (%)	80% CE-lb	80% CE-ub	95% CI-lb	95% CI-ub	N_{FS}
Agreeableness																					
mixed	50	18,686	0.79	0.09	9	0.80	0.07	46	0.78	0.13	15	0.031	0.081	0.061	0.120	42	−0.092	0.215	0.047	0.076	12
non	17	13,722	0.70	0.21	2	0.75	0.07	8	1.00	0.00	0	0.049	0.087	0.068	0.109	17	−0.071	0.208	0.052	0.085	7
verbal	16	6,088	0.74	0.06	8	0.79	0.06	13	1.00	0.00	1	0.012	0.093	0.015	0.100	31	−0.112	0.143	−0.010	0.040	−11
Conscientiousness																					
mixed	66	28,475	0.80	0.08	13	0.84	0.11	52	0.76	0.12	21	0.044	0.089	0.086	0.142	31	−0.096	0.268	0.075	0.098	48
non	19	13,871	0.76	0.18	3	0.81	0.07	9	1.00	0.00	0	0.071	0.096	0.091	0.112	16	−0.052	0.235	0.075	0.108	16
verbal	16	6,081	0.74	0.06	7	0.79	0.09	13	1.00	0.00	1	−0.041	0.099	−0.054	0.110	27	−0.195	0.087	−0.079	−0.029	−33
Emotional Stability																					
mixed	52	19,125	0.79	0.09	9	0.83	0.09	43	0.78	0.13	16	0.057	0.094	0.108	0.142	34	−0.074	0.290	0.094	0.122	61
non	17	13,722	0.70	0.21	2	0.78	0.06	8	1.00	0.00	0	0.129	0.097	0.175	0.120	15	0.022	0.328	0.159	0.191	43
verbal	16	6,088	0.74	0.06	8	0.82	0.07	13	1.00	0.00	1	0.029	0.086	0.037	0.087	36	−0.074	0.148	0.012	0.062	−4
Extraversion																					
mixed	52	18,244	0.79	0.09	10	0.83	0.10	42	0.77	0.14	15	0.061	0.078	0.121	0.103	53	−0.012	0.253	0.106	0.135	74
non	16	6,275	0.76	0.18	3	0.78	0.06	6	1.00	0.00	0	0.028	0.057	0.037	0.035	78	−0.007	0.082	0.013	0.062	−4
verbal	17	6,172	0.74	0.06	8	0.82	0.09	14	1.00	0.00	1	0.031	0.084	0.040	0.083	39	−0.067	0.146	0.015	0.065	−3
Openness to experience																					
mixed	44	10,546	0.81	0.08	7	0.79	0.10	41	0.77	0.14	16	0.147	0.094	0.283	0.100	66	0.155	0.411	0.265	0.300	205
non	12	5,951	0.07	0.21	2	0.76	0.09	5	1.00	0.00	0	0.095	0.071	0.131	0.074	42	0.036	0.225	0.106	0.155	20
verbal	16	7,697	0.74	0.06	7	0.78	0.07	12	1.00	0.00	1	0.129	0.088	0.170	0.099	27	0.044	0.296	0.149	0.191	39

Abbreviations: CE, credibility interval; CI, confidence interval; lb, lower bound; mixed, mixed cognitive ability tests; N_{FS}, fail-safe N; non, nonverbal cognitive ability tests; r_{xx}, average reliability of predictor; r_{yy}, average reliability of criterion; ub, upper bound; U_x, average restriction in variance; verbal, verbal cognitive ability tests; VAR (%), part of variance that can be attributed to artifacts.

still did not meet the 75% criterion, which hints at further moderator effects.

4.5 | Moderator analysis 3: Type of cognitive ability test

Table 6 summarizes the meta-analytic results separately for the use of verbal, nonverbal, or mixed cognitive ability tests. With regard to Hypothesis 4, we did not find a higher effect of verbal than of nonverbal cognitive ability tests for any of the personality dimensions. However, Conscientiousness, Emotional Stability, and Agreeableness showed the reversed pattern. For Extraversion and Openness to Experience, there was no difference between the two types of ability tests. The results of the mixed category tended to lie between those for verbal and nonverbal cognitive ability tests. In summary, these results do not provide any evidence for the moderation hypothesis specified under Hypothesis 4, but rather suggest that the relationships between different types of cognitive ability and faking may be more complex than hitherto assumed. Like the preceding moderation analysis, this analysis was not able to explain the majority of variance in the corresponding effect sizes: Only one of the 15 separate meta-analytic calculations fulfilled the 75% criterion for variance reduction.

5 | DISCUSSION

Our meta-analysis showed, for the first time, that the relationship between cognitive ability and personality scores differs between selection situations and nonselection situations. The correlations for selection situations were significantly positive for all dimensions of the Five-Factor model of personality, and we found significantly higher meta-analytical correlations for selection samples (ρ = 0.105–0.262) than for nonselection samples (ρ = 0.001–0.158). In other words, personality test scores share more variance with cognitive ability when measured under selection conditions. Assuming that applicant faking is primarily responsible for this change at the construct level, our results provide evidence to support those faking theories which argue that cognitive ability is a determinant of the ability to fake (e.g., Marcus, 2009; Snell et al., 1999; Tett & Simonet, 2011).

This pattern becomes even clearer when the results are considered separately according to the study design. Our results revealed significantly higher correlations between cognitive ability and personality in laboratory studies than in field studies. The proportion of variance in personality that can be explained by cognitive ability is particularly high in laboratory studies. These results fit in with the arguments put forward by some authors (e.g., Ones et al., 1996) that the mental processes involved in answering personality tests in a real application situation or in a laboratory study are hardly comparable. At this point, it can be stated that even if the correlations between field studies and nonselection studies differed, the results from these two study designs showed more similarity with each other than with the results of laboratory studies.

Indeed, there may be major motivational differences between the laboratory versus field situations. According to most current faking models, the relationship between faking motivation and faking behavior is moderated by the ability aspect of faking (e.g., Ellingson & McFarland, 2011; Goffin & Boyd, 2009; McFarland & Ryan, 2006; Roulin et al., 2016). The individual faking motivation in real application situations varies greatly due to individual differences, concrete subjective considerations, and situational circumstances. In contrast, participants' motivation to draw an improved picture of themselves in a laboratory study results from a well-controlled indirect (or sometimes direct) instruction to fake. This may result in a more uniform faking motivation in laboratory studies than in field studies. In line with an assumed moderating effect of cognitive ability, these limited motivational differences between participants in laboratory studies may lead to the more pronounced link between cognitive ability and actual faking behavior. At the same time, such differences in motivation may also be a reason why differences between field and nonselection samples emerge solely regarding the personality dimensions of Conscientiousness, Emotional Stability, and Extraversion. These dimensions likely have particularly high face validity for the work context—applicants might consider them to be especially important for future employers (see Jansen et al., 2012). Accordingly, the motivation of most applicants to present themselves in a better light regarding these dimensions should be uniformly high, which, in turn, should increase the relevance of cognitive ability for successful faking behavior.

Our findings regarding different types of personality tests, in particular single-stimulus and forced-choice, were less clear, mainly due to the small number of studies that actually used forced-choice tests. However, it is noteworthy that the correlations between Conscientiousness and cognitive ability in samples utilizing forced-choice tests were among the highest of all meta-analytic calculations in this study (ρ = 0.310). This may also be attributable to the fact that applicants consider this dimension to be particularly important for a future employer (cf. Jansen et al., 2012). In forced-choice tests, applicants usually have to choose between several response options that belong to different personality dimensions. Applicants with high cognitive ability might excel in recognizing the importance of Conscientiousness for the world of work, and therefore be more likely to choose answers corresponding to this dimension than applicants with lower cognitive ability. As such, our findings support many authors' claims that forced-choice personality tests appear to be harder to fake than single-stimulus tests (e.g., Christiansen et al., 2005; Jackson et al., 2000), but for this reason, they may also lead to a bias in favor of applicants with higher cognitive ability (Rothstein & Goffin, 2006; Vasilopoulos et al., 2006).

With regard to the type of cognitive ability tests, our meta-analytic results contradicted the findings of previous research (Grieve & Mahar, 2010; MacCann, 2013). Our findings concerning this moderator analysis did not show a higher effect in the samples in which verbal cognitive ability was measured; rather, they indicated a stronger effect of nonverbal cognitive ability on faking in personality tests. Moreover, we even found a negative relationship between Conscientiousness and verbal cognitive ability in selection samples ($\rho = -0.054$). A possible explanation for this counterintuitive finding might be that merely understanding the items can be accomplished equally well by all applicants and is not the main hurdle for faking in personality tests. Instead, nonverbal abilities such as the ability to see patterns behind items (i.e., being able to detect the corresponding dimension) and to conclude the required characteristics for a job (e.g., Kleinmann et al., 2011; König et al., 2007) might be more important for successful faking.

5.1 | Theoretical implications and future research directions

This study contributes to the theoretical understanding of faking in several main aspects. First, our results help to clarify the question of the role of cognitive ability in the process of faking in personality tests. We were able to show that personality tests share a higher proportion of their variance with cognitive ability in selection situations than in nonselection situations. In contrast to the basic assumptions regarding the psychological construct of personality (Allport & Odbert, 1936; McCrae & Costa, 1985), our findings suggested that cognitive ability does play a role in personality assessment in selection situations. This supports the idea already put forward by previous researchers (Klehe et al., 2012; Wrensen & Biderman, 2005) that filling out a personality test in a selection situation is driven by a slightly different underlying process than filling out such a test in a nonselection context.

Second, our findings support an additional explanation of the criterion validity of personality tests in personnel selection, which has also been discussed in previous faking research (e.g., Johnson & Hogan, 2006; Kleinmann et al., 2011). In general, cognitive ability is one of the best predictors of work performance, which may also explain at least some part of the criterion validity of personality tests through the relationship studied in this meta-analysis. Although the variance in personality tests is likely dominated by personality constructs, it also seems to be influenced by variance in cognitive ability, the questions arise to which extent this is the case.

Third, the discrepancies we found between different study designs also indicate that the construct captured in laboratory studies does not fully correspond to the construct captured in real selection situations. Although Ones et al. (1996) had already pointed out that the mental processes underlying the filling out of a personality test may differ between laboratory and field situations, our results even indicate that this discrepancy may be greater than that between selection and nonselection situations. This, in turn, raises the question of to what extent results from laboratory studies can be generalized to real selection situations, and whether recommendations for personnel selection should be derived from such results at all.

For further research, we would, therefore, like to encourage a stronger focus on field studies wherever possible. We also call for a stronger verification of the construct validity of the personality tests used in the selection context, and above all, we recommend that this psychometric property is evaluated in the actual selection context. Most importantly, in our opinion, faking research should focus more on the mental processes, strategies, and objectives of applicants in selection situations (cf. König et al., 2012; Ziegler, 2011). Only through a better understanding of what is going on in the mind of applicants when they fill out personality tests can we fully understand the phenomenon of faking.

Furthermore, we would like to encourage all researchers in the field of faking to publish more information in their papers to facilitate meta-analytical research. In this meta-analysis, we would also have liked to more directly analyze the relationship between cognitive ability and faking, but far too few studies reported the required correlations (e.g., between cognitive ability and the raw difference scores between honest and faking condition). Some of the primary studies included in our meta-analysis also lacked correlation tables for the study variables and we thus had to request this very basic static information from the authors. In our opinion, it is, therefore, essential for the aggregability but also for the replicability of faking research that all further studies report the following information: (a) a detailed description of the faking instruction, ideally in the original wording, (b) reliabilities, means, and standard deviations for all study variables, individually for all groups and conditions, and the corresponding correlation tables, and (c) for within-person studies the correlations of the raw as well as the regression adjusted difference faking scores (see Burns & Christiansen, 2011) with all study variables.

5.2 | Implications for personnel selection

In the real world of personnel selection, many organizations are concerned that applicants' faking behavior might seriously undermine the usefulness and validity of personality tests. Therefore, persons in charge of personnel selection may be greatly interested to know that cognitive ability plays a major role when applicants fill out a personality test and that more intelligent applicants also tend to have higher scores on such personality tests. In our opinion, these findings may inform the use of personality tests in personnel selection in at least two aspects. (a) Our findings raise the question of to which extent the intended personality constructs are being measured in selection situations and to which extent personality tests in assessments measure cognitive ability. At this point, a company may argue that as long as employees perform well, it does not matter whether they are doing so because they are truly conscientious or because they are conscientious and smart, but the answer to this question may affect organizations' internal justification and selection of personality tests as a personnel selection tool. (b) As a further practical implication of our findings, we recommend caution when using forced-choice tests to measure personality. Forced-choice tests are considered harder to fake (e.g., Christiansen et al., 2005; Jackson et al., 2000) but also showed a fairly large proportion of shared variance with cognitive ability. Especially for the dimension of Conscientiousness, which has the highest predictive validity for work performance (Barrick & Mount, 1991), we found high correlations with cognitive ability. Organizations should, therefore, be aware that forced-choice tests likely have the advantage of being less prone to faking and simultaneously the disadvantage of measuring the actual construct of personality to an even smaller extent than single-stimulus tests.

5.3 | Limitations

Three main limitations of the present meta-analysis need to be mentioned: First, we were unable to analyze the relationship between cognitive ability and faking in a direct manner—our approach only allowed us to compare correlations of personality and ability in selection and nonselection situations and to conclude the effect on faking from the corresponding discrepancies. The main reason for this limitation is that there was an insufficient number of primary studies that reported the correlations between cognitive ability and some direct measure of faking (e.g., the difference between an honest condition and an 'as applicant' condition). Hopefully, more researchers will report such information in the future, enabling such correlations to be summarized in future meta-analytic work. Second, many of our analyses did not fulfill the 75% rule for variance reduction, suggesting room for other moderators (Schmidt & Hunter, 2014), which should be explored by future research. Third, it must be pointed out that the correlations found between cognitive ability and personality in selection situations were significantly higher than those in nonselection samples, but were rather small in effect size (Cohen, 1992; Hemphill, 2003; Paterson et al., 2016). In general, cognitive ability plays a meaningful role in the assessment of personality in selection situations, but not the most influential role.

6 | CONCLUSION

Personality tests are considered to be a valid instrument for predicting work performance but are often criticized for their susceptibility to faking. In this context, the role played by applicants' cognitive ability in faking remains controversial. The results of this meta-analysis shed some light on this issue by revealing substantially higher correlations between cognitive ability and personality in selection situations than in nonselection situations. Thus, our findings suggest that other mental processes take place when filling out personality tests in selection situations and that accordingly, a somewhat different psychological construct might be captured compared to nonselection situations. Viewed as a whole, this also provides indirect evidence for a link between cognitive ability and faking. Moderator analyses showed that the correlations with cognitive ability are particularly high in laboratory studies, whereas the correlations in field studies differ from nonselection situations to a considerably lesser degree. These findings suggest that the response behavior of participants in laboratory studies may be less representative of applicants in real selection situations than expected. Accordingly, the results obtained in the laboratory should only be generalized with the utmost caution. To gain a more holistic view of faking, future research may also be well served by shifting the focus somewhat away from predictors of this phenomenon and moving toward mental processes, strategies, and objectives of applicants in selection situations.

ACKNOWLEDGMENT

We would like to thank all authors who responded to our e-mails and who pointed out further studies to us. In particular, we would like to thank the 22 authors who provided us with additional data and calculations for their studies, some of which had already been published 12 years earlier. Furthermore, we would like to thank Arthur Poropat for providing us with so far unpublished results of his meta-analysis regarding the relationship between personality and academic performance (2009). Open access funding enabled and organized by Projekt DEAL. [Correction added on 27 October 2020, after first online publication: Projekt Deal funding statement has been added.]

ORCID

Michael Schilling https://orcid.org/0000-0003-2338-5866
Cornelius J. König https://orcid.org/0000-0003-0477-8293

ENDNOTE

[1] We are very grateful that Arthur Poropat provided us with the previously unpublished results on the relationship between personality and cognitive ability, which were part of his meta-analysis regarding the relationship between personality and academic performance (Poropat, 2009).

REFERENCES

Ackerman, P. L., & Ellingsen, V. J. (2014). Vocabulary overclaiming—A complete approach: Ability, personality, self-concept correlates, and gender differences. *Intelligence*, *46*, 216–227. https://doi.org/10.1016/j.intell.2014.07.003

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*(1), i–171. https://doi.org/10.1037/h0093360.

Anglim, J., Bozic, S., Little, J., & Lievens, F. (2018). Response distortion on personality tests in applicants: Comparing high-stakes to low-stakes medical settings. *Advances in Health Sciences Education*, *23*, 311–321. https://doi.org/10.1007/s10459-017-9796-8

Arthur, W., Doverspike, D., Muñoz, G. J., Taylor, J. E., & Carr, A. E. (2014). The use of mobile devices in high-stakes remotely delivered assessments and testing. *International Journal of Selection and Assessment*, *22*, 113–123. https://doi.org/10.1111/ijsa.12062

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

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, *14*, 317–335. https://doi.org/10.1111/j.1468-2389.2006.00354.x

Boss, P., König, C. J., & Melchers, K. G. (2015). Faking good and faking bad among military conscripts. *Human Performance*, *28*, 26–39. https://doi.org/10.1080/08959285.2014.974758

Burns, G. N., & Christiansen, N. D. (2011). Methods of measuring faking behavior. *Human Performance*, *24*, 358–372. https://doi.org/10.1080/08959285.2011.597473

Cao, M., & Drasgow, F. (2019). Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations. *Journal of Applied Psychology*, *104*(11), 1347–1368. https://doi.org/10.1037/apl0000414

Christiansen, N. D., Burns, G. N., & Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance*, *18*, 267–307. https://doi.org/10.1207/s15327043hup1803_4

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. https://doi.org/10.1037/0033-2909.112.1.155

De Fruyt, F., Aluja, A., García, L. F., Rolland, J.-P., & Jung, S. C. (2006). Positive presentation management and intelligence and the personality differentiation by intelligence hypothesis in job applicants. *International Journal of Selection and Assessment*, *14*, 101–112. https://doi.org/10.1111/j.1468-2389.2006.00337.x

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Ellingson, J. E., & McFarland, L. A. (2011). Understanding faking behavior through the lens of motivation: An application of VIE theory. *Human Performance*, *24*, 322–337. https://doi.org/10.1080/08959285.2011.597477

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or *g*?: The relationship between the Scholastic Assessment Test and general cognitive ability. *Psychological Science*, *15*, 373–378. https://doi.org/10.1111/j.0956-7976.2004.00687.x

Furnham, A., Taylor, J., & Chamorro-Premuzic, T. (2008). Personality and intelligence correlates of assessment center exercises. *Individual Differences Research*, *6*, 181–192.

Galić, Z., Jerneić, Z., & Kovačić, M. P. (2012). Do applicants fake their personality questionnaire responses and how successful are their attempts? A case of military pilot cadet selection. *International Journal of Selection and Assessment*, *20*, 229–241. https://doi.org/10.1111/j.1468-2389.2012.00595.x

Goffin, R. D., & Boyd, A. C. (2009). Faking and personality assessment in personnel selection: Advancing models of faking. *Canadian Psychology*, *50*, 151–160. https://doi.org/10.1037/a0015946

Grieve, R., & Mahar, D. (2010). The role of fluid and emotional intelligence in malingering. *Australian Journal of Psychology*, *62*, 103–111. https://doi.org/10.1080/00049530903032836

Griffin, B., & Wilson, I. G. (2012). Faking good: Self-enhancement in medical school applicants. *Medical Education*, *46*, 485–490. https://doi.org/10.1111/j.1365-2923.2011.04208.x

Grubb III, W. L. & McDaniel, M. A. (2007). The fakability of Bar-On's emotional quotient inventory short form: Catch me if you can. *Human Performance*, *20*, 43–59. https://doi.org/10.1207/s15327043hup2001_3

Hale, R. L., & Padgett, R. J. (2014, May 22). *Take it but don't fake it: Using self-awareness to reduce fabrications on personality tests*. Paper presented at 26th Annual Convention of the Association for Psychological Science, San Francisco, CA.

Hausknecht, J. P. (2010). Candidate persistence and personality test practice effects: Implications for staffing system management. *Personnel Psychology*, *63*, 299–324. https://doi.org/10.1111/j.1744-6570.2010.01171.x

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, *3*, 486–504. https://doi.org/10.1037/1082-989X.3.4.486

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*, 78–79. https://doi.org/10.1037/0003-066X.58.1.78

Hooper, A. C. (2007). *Self-presentation on personality measures in lab and field settings: A meta-analysis* (Doctoral dissertation). University of Minnesota, ProQuest Dissertations and Theses database (Order No. 3273134).

Jackson, D. N., & Rushton, J. P. (2006). Males have greater *g*: Sex differences in general mental ability from 100,000 17-to 18-year-olds on the Scholastic Assessment Test. *Intelligence*, *34*, 479–486. https://doi.org/10.1016/j.intell.2006.03.005

Jackson, D. N., Wroblewski, V. R., & Ashton, M. C. (2000). The impact of faking on employment tests: Does forced choice offer a solution? *Human Performance*, *13*, 371–388. https://doi.org/10.1207/S15327043HUP1304_3

Jansen, A., König, C. J., Kleinmann, M., & Melchers, K. G. (2012). The interactive effect of impression motivation and cognitive schema on self-presentation in a personality inventory. *Journal of Applied Social Psychology*, *42*, 1932–1957. https://doi.org/10.1111/j.1559-1816.2012.00925.x

Johnson, J. A., & Hogan, R. (2006). A socioanalytic view of faking. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 209–231). Information Age Publishing.

Klehe, U.-C., Kleinmann, M., Hartstein, T., Melchers, K. G., König, C. J., Heslin, P. A., & Lievens, F. (2012). Responding to personality tests in a selection context: The role of the ability to identify criteria and the ideal-employee factor. *Human Performance*, *25*, 273–302. https://doi.org/10.1080/08959285.2012.703733

Kleinmann, M., Ingold, P. V., Lievens, F., Jansen, A., Melchers, K. G., & König, C. J. (2011). A different look at why selection procedures work: The role of candidates' ability to identify criteria. *Organizational Psychology Review*, *1*, 128–146. https://doi.org/10.1177/2041386610387000

Koenig, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence*, *36*, 153–160. https://doi.org/10.1016/j.intell.2007.03.005

Komarraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. *Learning and Individual Differences*, *24*, 103–109. https://doi.org/10.1016/j.lindif.2012.12.007

König, C. J., Hafsteinsson, L. G., Jansen, A., & Stadelmann, E. H. (2011). Applicants' self-presentational behavior across cultures: Less self-presentation in Switzerland and Iceland than in the United States. *International Journal of Selection and Assessment*, *19*, 331–339. https://doi.org/10.1111/j.1468-2389.2011.00562.x

König, C. J., Melchers, K. G., Kleinmann, M., Richter, G. M., & Klehe, U.-C. (2006). The relationship between the ability to identify evaluation criteria and integrity test scores. *Psychological Science*, *48*, 369–377.

König, C. J., Melchers, K. G., Kleinmann, M., Richter, G. M., & Klehe, U.-C. (2007). Candidates' ability to identify criteria in nontransparent selection procedures: Evidence from an assessment center and a structured interview. *International Journal of Selection and Assessment*, *15*, 283–292. https://doi.org/10.1111/j.1468-2389.2007.00388.x

König, C. J., Merz, A.-S., & Trauffer, N. (2012). What is in applicants' minds when they fill out a personality test? Insights from a qualitative study. *International Journal of Selection and Assessment*, *20*, 442–452. https://doi.org/10.1111/ijsa.12007

Lange, S. (2013). *Metaanalysen zum Zusammenhang von der Allgemeinen, Fluiden und Kristallisierten Intelligenz mit dem Fünf-Faktoren-Modell der Persönlichkeit unter besonderer Berücksichtigung des Publikationsbias* [Meta-analysis on the relationship between the general, fluid and crystallized intelligence with the Five-Factor model of personality, taking publication bias into account] (Doctoral dissertation). Rheinische Friedrich-Wilhelms-Universität zu Bonn. http://hss.ulb.uni-bonn.de/2013/3286/3286.pdf

Levashina, J., Morgeson, F. P., & Campion, M. A. (2009). They don't do it often, but they do it well: Exploring the relationship between applicant mental abilities and faking. *International*

Journal of Selection and Assessment, *17*, 271–281. https://doi.org/10.1111/j.1468-2389.2009.00469.x

Levashina, J., Weekley, J. A., Roulin, N., & Hauck, E. (2014). Using blatant extreme responding for detecting faking in high-stakes selection: Construct validity, relationship with general mental ability, and subgroup differences. *International Journal of Selection and Assessment*, *22*, 371–383. https://doi.org/10.1111/ijsa.12084

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Harvard University Press.

MacCann, C. (2013). Instructed faking of the HEXACO reduces facet reliability and involves more Gc than Gf. *Personality and Individual Differences*, *55*, 828–833. https://doi.org/10.1016/j.paid.2013.07.007

MacKenzie, W. I., Ployhart, R. E., Weekley, J. A., & Ehlers, C. (2010). Contextual effects on SJT responses: An examination of construct validity and mean differences across applicant and incumbent contexts. *Human Performance*, *23*, 1–21. https://doi.org/10.1080/08959280903400143

Marcus, B. (2009). 'Faking' from the applicant's perspective: A theory of self-presentation in personnel selection settings. *International Journal of Selection & Assessment*, *17*, 417–430. https://doi.org/10.1111/j.1468-2389.2009.00483.x

Martin, B. A., Bowen, C.-C., & Hunt, S. T. (2002). How effective are people at faking on personality questionnaires? *Personality and Individual Differences*, *32*, 247–256. https://doi.org/10.1016/S0191-8869(01)00021-6

McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequacy 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

McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, *85*, 812–821. https://doi.org/10.1037/0021-9010.85.5.812

McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology*, *36*, 979–1016. https://doi.org/10.1111/j.0021-9029.2006.00052.x

McNatt, D. B. (2000). Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology*, *85*, 314–322. https://doi.org/10.1037//0021-9010.85.2.314

Merkulova, N., Melchers, K. G., Kleinmann, M., Annen, H., & Tresch, T. S. (2014). Effects of individual differences on applicant perceptions of an operational assessment center. *International Journal of Selection and Assessment*, *22*, 355–370. https://doi.org/10.1111/ijsa.12083

Mersman, J. L., & Shultz, K. S. (1998). Individual differences in the ability to fake on personality measures. *Personality and Individual Differences*, *24*, 217–227. https://doi.org/10.1016/S0191-8869(97)00160-8

Mudgett, B. O. (2000). *Individual and situational influences on distortion of personality tests* (Doctoral dissertation). Rice University, ProQuest Dissertations and Theses database (Order No. 9969296).

Mueller-Hanson, R. A., Heggestad, E. D., Thornton III, G. C. (2006). Individual differences in impression management: An exploration of the psychological processes underlying faking. *Psychology Science*, *48*, 288–312.

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, *81*, 660–679. https://doi.org/10.1037/0021-9010.81.6.660

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics*, *8*, 157–159. https://doi.org/10.2307/1164923

Paterson, T. A., Harms, P. D., Steel, P., & Credé, M. (2016). An assessment of the magnitude of effect sizes: Evidence from 30 years of meta-analysis in management. *Journal of Leadership & Organizational Studies*, *23*, 66–81. https://doi.org/10.1177/1548051815614321

Pauls, C. A., & Crost, N. W. (2005). Cognitive ability and self-reported efficacy of self-presentation predict faking on personality measures. *Journal of Individual Differences*, *26*, 194–206. https://doi.org/10.1027/1614-0001.26.4.194

Peeters, H., & Lievens, F. (2005). Situational judgment tests and their predictiveness of college students' success: The influence of faking. *Educational and Psychological Measurement*, *65*, 70–89. https://doi.org/10.1177/0013164404268672

Peterson, M. H., Griffith, R. L., & Converse, P. D. (2009). Examining the role of applicant faking in hiring decisions: Percentage of fakers hired and hiring discrepancies in single- and multiple-predictor selection. *Journal of Business and Psychology*, *24*, 373–386. https://doi.org/10.1007/s10869-009-9121-5

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

R Core Team. (2017). *R: A language and environment for statistical computing (Version 3.3.3)*. R Foundation for Statistical Computing. http://www.R-project.org/

Raven, J. C. (1938). *Raven's progressive matrices*. Western Psychological Services.

Raymark, P. H., & Tafero, T. L. (2009). Individual differences in the ability to fake on personality measures. *Human Performance*, *22*, 86–103. https://doi.org/10.1080/08959280802541039

Robie, C., Komar, S., & Brown, D. J. (2010). The effects of coaching and speeding on Big Five and impression management scale scores. *Human Performance*, *23*, 446–467. https://doi.org/10.1080/08959285.2010.515278

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review*, *16*, 155–180. https://doi.org/10.1016/j.hrmr.2006.03.004

Roulin, N., Krings, F., & Binggeli, S. (2016). A dynamic model of applicant faking. *Organizational Psychology Review*, *6*, 145–170. https://doi.org/10.1177/2041386615580875

Ryan, A. M., Ployhart, R. E., & Friedel, L. A. (1998). Using personality testing to reduce adverse impact: A cautionary note. *Journal of Applied Psychology*, *83*, 298–307. https://doi.org/10.1037/0021-9010.83.2.298

Salgado, J. F., & Táuriz, G. (2014). The Five-Factor Model, forced-choice personality inventories and performance: A comprehensive meta-analysis of academic and occupational validity studies. *European Journal of Work and Organizational Psychology*, *23*, 3–30. https://doi.org/10.1080/1359432X.2012.716198

Schilling, M., Sparfeldt, J. R., Becker, N., Engel, M., Levacher, J., Sebastian, T. F. P., Schäfer, J., Schwabe, S., & König, C. J. (2020). Is it enough to be willing to win or do you have to be smart? The relationship between competitive worldviews, cognitive abilities, and applicant faking in personality tests. *International Journal of Selection and Assessment*, *28*, 264–282. https://doi.org/10.1111/ijsa.12296

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd. ed.). Sage.

Schmidt, F. L., & Le, H. (2005). *Hunter & Schmidt's meta-analysis program (Version 1.1)*. The University of Iowa.

Snell, A. F., Sydell, E. J., & Lueke, S. B. (1999). Towards a theory of applicant faking: Integrating studies of deception. *Human Resource Management Review*, *9*, 219–242. https://doi.org/10.1016/S1053-4822(99)00019-4

Tett, R. P., & Simonet, D. V. (2011). Faking in personality assessment: A "multisaturation" perspective on faking as performance. *Human Performance*, *24*, 302–321. https://doi.org/10.1080/08959285.2011.597472

van Hooft, E. A. J., & Born, M. P. (2012). Intentional response distortion on personality tests: Using eye-tracking to understand response processes when faking. *Journal of Applied Psychology*, *97*, 301–316. https://doi.org/10.1037/a0025711

Vasilopoulos, N. L., Cucina, J. M., Dyomina, N. V., Morewitz, C. L., & Reilly, R. R. (2006). Forced-choice personality tests: A measure of personality and cognitive ability? *Human Performance*, *19*, 175–199. https://doi.org/10.1207/s15327043hup1903_1

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wechsler, D. (2014). *WAIS-IV: Wechsler Adult Intelligence Scale—Fourth edition*. Pearson Assessment.

Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, *75*, 315–321. https://doi.org/10.1037/0021-9010.75.3.315

Wonderlic, E. F. (1996). *Wonderlic personnel test and scholastic level exam user's manual*. Wonderlic and Associates.

Wrensen, L. B., & Biderman, M. D. (2005, April). *Factors related to faking ability: A structural equation model application*. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Wright, N. A., Meade, A. W., & Gutierrez, S. L. (2014). Using invariance to examine cheating in unproctored ability tests. *International Journal of Selection and Assessment*, *22*, 12–22. https://doi.org/10.1111/ijsa.12053

Ziegler, M. (2011). Applicant faking: A look into the black box. *The Industrial-Organizational Psychologist*, *49*(1), 29–36.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Schilling M, Becker N, Grabenhorst MM, König CJ. The relationship between cognitive ability and personality scores in selection situations: A meta-analysis. *Int J Select Assess*. 2021;29:1–18. https://doi.org/10.1111/ijsa.12314