Check for updates

Article

Organizational Research Methods
2024, Vol. 27(4) 681–712
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10944281231212570
journals.sagepub.com/home/orm
Sage

Confounding Effects of Insufficient Effort Responding Across Survey Sources: The Case of Personality Predicting Performance

Jason L. Huang[1], Nathan A. Bowling[2], Benjamin D. McLarty[3], Donald H. Kluemper[4], and Zhonghao Wang[5]

Abstract

Insufficient effort responding (IER) to surveys, which occurs when participants provide responses in a haphazard, careless, or random fashion, has been identified as a threat to data quality in survey research because it can inflate observed relationships between *self-reported* measures. Building on this discovery, we propose two mechanisms that lead to IER exerting an unexpected confounding effect between *self-reported* and *informant-rated* measures. First, IER can contaminate self-report measures when the means of attentive and inattentive responses differ. Second, IER may share variance with some informant-rated measures, particularly supervisor ratings of participants' job performance. These two mechanisms operating in tandem would suggest that IER can act as a "third variable" that inflates observed relationships between self-reported predictor scores and informant-rated criteria. We tested this possibility using a multisource dataset ($N = 398$) that included incumbent self-reports of five-factor model personality traits and supervisor-ratings of three job performance dimensions—task performance, organizational citizenship behavior (OCB), and counterproductive work behavior (CWB). We observed that the strength of the relationships between self-reported personality traits and supervisor-rated performance significantly decreased after IER was controlled: Across the five personality traits, the average reduction of magnitude from the zero-order to partial correlations was |.08| for task performance, |.07| for OCB, and |.14| for CWB. Because organizational practices are often driven by research linking incumbent-reported predictors to supervisor-rated criteria (e.g., validation of predictors used in various organizational contexts), our findings have important implications for research and practice.

[1]School of Human Resources & Labor Relations, Michigan State University, East Lansing, MI, USA
[2]Department of Psychology, University of Central Florida, Orlando, FL, USA
[3]School of Innovation & Entrepreneurship, William G. Rohrer College of Business, Rowan University, Glassboro, NJ, USA
[4]Rawls College of Business, Texas Tech University, Lubbock, TX, USA
[5]Marilyn Davies College of Business, University of Houston-Downtown, Houston, TX, USA

Corresponding Author:
Jason L. Huang, School of Human Resources & Labor Relations, Michigan State University, 368 Farm Lane, 437 SKH, East Lansing, MI 48824-5602, USA.
Email: huangjl@msu.edu

Keywords
insufficient effort responding, response effort, careless responding, personality, job performance

Research in management and organizational science often uses self-reports where individuals respond to survey measures assessing various aspects of themselves, such as beliefs, attitudes, cognitions, traits, and behaviors. Researchers have long recognized that basing one's conclusions solely on self-reports can be problematic because common method bias can inflate the observed associations (Podsakoff et al., 2003, 2012). By contrast, concerns over inflated associations are generally assumed to be less pronounced when self-reports are correlated with other sources of data such as informant-reports (Aguinis & Vandenberg, 2014).

As a case in point, five-factor model (FFM) personality traits have often been examined as antecedents to employee behaviors in management research (Sackett et al., 2017). Personality traits not only offer a lens for management scholars to understand employee behaviors, but also provide practical tools for organizations to select high-performing workers (Hough & Dilchert, 2017). In particular, conscientiousness and emotional stability are consistently related to various job performance dimensions, such as task performance, organizational citizenship behavior (OCB), and counterproductive work behavior (CWB; see Barrick et al., 2001; Barrick & Mount, 1991; Berry et al., 2007; Chiaburu et al., 2011; Dalal, 2005; Hurtz & Donovan, 2000). When assessing personality traits as predictors of job performance, the predominant approach is to ask employees to respond to a personality questionnaire about themselves and to use informants (e.g., supervisors) to provide ratings of respondents' performance. Indeed, it has become a typical practice for meta-analytic investigations of personality's influence on performance to exclude studies relying on self-reported performance (e.g., Chiaburu et al., 2011; van Aarde et al., 2017). An implicit assumption is that the use of different sources for personality and performance ratings mitigates the potential threat of inflated relationships due to common-method bias (Aguinis & Vandenberg, 2014; Podsakoff et al., 2012).

Such an assumption, however, can be faulty when some research participants engage in insufficient effort responding (IER) behavior—failing to "comply with survey instructions, correctly interpret item content, and provide accurate responses" (Huang et al., 2012, p. 100).[1] On the one hand, self-reports, such as personality ratings, can contain systematic errors unrelated to the variable being assessed (AERA, APA & NCME, 2014) due to IER (Credé, 2010; Huang et al., 2015; McGonagle et al., 2016). This is because as a group, random or careless responding tends to result in scores that average closer to the midpoint of a response scale when attentive respondents score on average above (or below) the midpoint. The mean difference between the two groups of respondents will translate into a negative (or positive) association between IER behavior and observed personality scores (Huang et al., 2015). As such, IER introduces unwanted systematic variance in self-report personality ratings. On the other hand, IER behavior reflects substantive behavioral tendencies (Bowling et al., 2016), which can influence behaviors that require effort and persistence, such as job performance. When these two mechanisms occur in tandem, IER behavior may be associated with both the antecedents and outcomes in a study. With IER serving as a "third variable" (Shadish et al., p. 7; also see Meade et al., 2009) between two substantive variables, the observed association between the self-report predictor and the outcome will be inflated. This inflation occurs even though the predictor and the outcome are *assessed via different sources*.

Examining these two mechanisms simultaneously will allow us to challenge the implicit belief that IER behavior only inflates Type I error rate when focal survey respondents (typically employees) are the only source of data, thus shedding new light on the problematic biasing effect introduced by IER behavior across sources. In the present paper, we identify and examine the mechanisms that

result in IER artificially inflating the relationship between self-reports and informant-reports. Specifically, we examine self-report personality and three supervisor-rated performance dimensions: (a) task performance, (b) OCB, and (c) CWB (for a discussion of this three-factor job performance model, see Rotundo & Sackett, 2002).

To accurately assess the role of IER, we simultaneously consider three types of effects (see Figure 1), integrating recent work that independently examined IER as either a source of systematic measurement error (Huang et al., 2015) or a substantive outcome of personality (Bowling et al., 2016). We situate the investigation in the context of using self-reported personality to predict supervisor-rated performance to provide a robust investigation of such abstract effects. First, path *a* (IER←→Scores) indicates the *observed* association between IER and self-report personality scores. Note that this path does not differentiate the cause of the association between the two variables; IER may be related to a self-reported trait for two reasons: (i) because IER introduces systematic measurement error (Huang et al., 2015) and (ii) the trait (e.g., conscientiousness) predisposes respondents to engage in IER (see Bowling et al., 2016). To account for the second possibility, path *b* (Personality→IER) captures the substantive effect of the underlying construct (e.g., conscientiousness) on IER.[2] Controlling for path *b* in IER while examining path *a* allows us to estimate the true degree to which IER behavior introduces systematic error in the measurement process.

Finally, path *c* (IER←→Performance) indicates the substantive association between IER and the outcome of interest (e.g., informant-reported job performance). In this case, whether individuals engage in IER behavior may be related to how they approach and perform their work. Assessing the three distinct effects simultaneously allows us to explicate how IER's associations with personality scores and performance may conspire to inflate the relationship between self-reported personality and informant-reported performance. We disentangle these effects using multi-source data from 398 employees, including personality ratings obtained through self-report and from family members as well as performance ratings from supervisors.

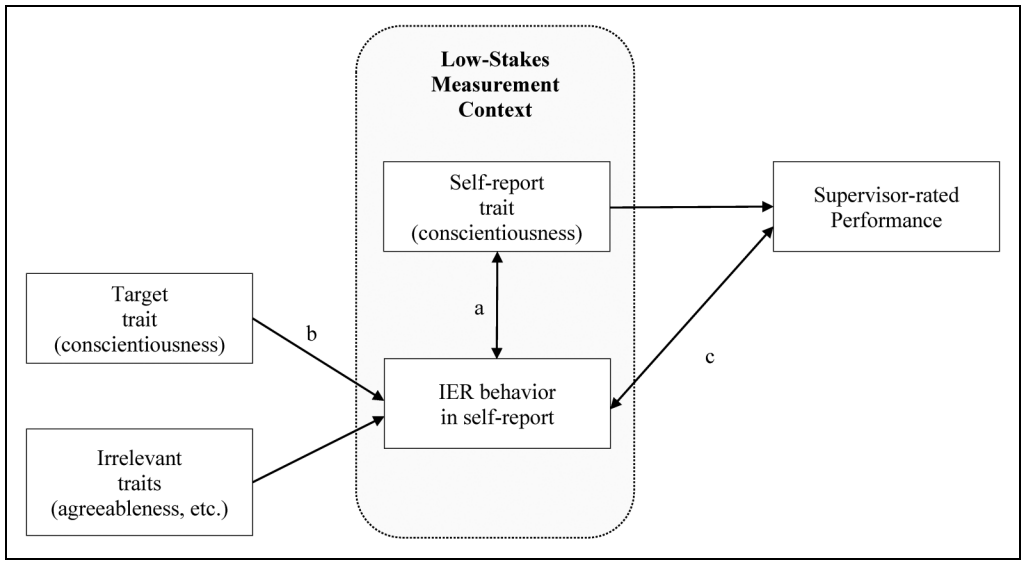

Figure 1. Insufficient effort responding (IER) as a potential confound in the personality-performance association.
Note: Path *a*: IER←→Scores; Path *b*: Personality→IER; Path c: IER←→Performance.

The present investigation has implications for both psychometric theory and measurement practice. From a theoretical perspective, disentangling methodological and substantive associations of IER broadens current thinking on how IER may affect observed associations between variables. While IER has been understood as a potential methodological confound within self-reports (e.g., Huang et al., 2015), uncovering the role of IER across self-reports and informant-reports can highlight the need to consider IER when outcome measures require consistent effort and attention. Practically speaking, if IER can inflate the associations between self-reports and informant-reports, researchers in the fields of management and organizational science should consider more standardized, rigorous, and deliberate screening for IER in data collection situations previously assumed to be practically invulnerable.

IER as a Methodological Confound in Observed Scores (IER → Scores)

Researchers have long assumed that IER produces random measurement error, thus attenuating observed relationships between self-report measures (see McGrath et al., 2010). Although IER can sometimes produce such attenuation effects (see Hough et al., 1990; Huang et al., 2012; Maniaci & Rogge, 2014), recent work has challenged this long-held assumption by showing that IER can introduce systematic error (Credé, 2010; Huang et al., 2015; McGonagle et al., 2016). This occurs when attentive participants' mean score ($M_{\text{attentive}}$) on a given measure is higher (or lower) than the mean score of IER participants (M_{IER}). More specifically, when $M_{\text{attentive}}$ is away (either higher or lower) from the midpoint of a response scale, M_{IER} is likely to shift closer to the midpoint, yielding consistent differences (either lower or higher) from the attentive respondents (Huang et al., 2015). We provide an example below to illustrate this point.

Imagine that participants responded to a set of conscientiousness items using a 7-point scale (1 = s*trongly disagree*; 7 = *strongly agree*; with a scale midpoint of 4). While the majority responded attentively and rated themselves, on average, higher than the scale's midpoint (e.g., $M_{\text{attentive}} = 5.00$), a small proportion of participants engaged in IER. We would expect the mean conscientiousness score for these IER respondents (M_{IER}) to be closer to the scale's midpoint for several reasons. First, when occasionally careless errors occur, there is a higher chance for scores to be lowered than increased. For instance, for someone with a score of 5, there are four response options for his/her score to be lowered, compared to two response options for his/her score to be increased. Second, when both positively and negatively keyed items are present—typical in personality assessments—consecutive endorsement of the same response options, also known as longstring or straightline responding (DeSimone et al., 2018), can lead to M_{IER} closer to the scale midpoint because positive and negative items cancel each other out. Third, deliberate attempts at random responding will lead to M_{IER} at the scale midpoint (see Huang et al., 2015 Study 2 for simulation results). Although individuals may not adopt a single approach when engaging in IER, these mechanisms above all point to the likelihood that M_{IER} is closer to the scale midpoint than $M_{\text{attentive}}$.

When two variables simultaneously contain systematic errors due to IER, the observed association will be inflated because of this confound. Indeed, simulations have shown that when $M_{\text{attentive}}$ deviates from the scale's midpoint, even a small proportion of IER (e.g., 5% of random responses) can introduce sizable spurious associations (e.g., r around .17) in otherwise uncorrelated measures (Huang et al., 2015). More importantly, when consistent careless responding is simulated from actual participant data instead of complete random responses, a similar confounding effect is observed (Huang & DeSimone, 2021). Further, partialling out the effects of IER in field surveys produces weakened relationships between self-reported variables (Huang et al., 2015; McGonagle et al., 2016). In other words, the mean difference between $M_{\text{attentive}}$ and M_{IER} results in IER acting as a confound within self-report measures, thus implying that it has a methodological effect on observed personality scores. Such effects are likely common

because self-report measures generally have $M_{\text{attentive}}$ values[3] that are away from a scale's midpoint (Huang et al., 2015).

Substantive Effects of Personality Traits on IER (Personality → IER)

As noted earlier, the observed association between IER and substantive measures (i.e., IER←→Scores) cannot be automatically attributed to the systematic measurement error introduced by IER. In the context of personality measurement, the substantive constructs of interest, personality traits, can influence IER behavior, thus partially accounting for the observed IER←→Scores effect. There are several reasons to expect Personality→IER (see Bowling et al., 2016; McKay et al., 2018; Ward et al., 2017). To understand the potential effects of personality on IER, it is necessary to first consider the substantive nature of IER. IER reflects one or more of the following characteristics: (a) an unwillingness or inability to expend and sustain the effort necessary to respond to questionnaire content; (b) an unwillingness or inability to comply with questionnaire instructions; and (c) an unwillingness to support the researcher's attempt to collect high-quality data. Given that researchers are expected to consider whether the respondents can comprehend and process the questions and instructions when designing a survey questionnaire (e.g., Goddard & Villanova, 2006; Lucas & Baird, 2006; Oppenheim, 1992), we expect the ability to play a trivial role and thus turn our focus to willingness. As such, a number of traits may substantively impact IER. For example, conscientiousness is likely to influence the first and second characteristics, while agreeableness is likely to influence the second and third characteristics (see Bowling et al., 2016).

Consistent with this theorizing, research has found that personality traits may influence the extent to which a respondent habitually engages in IER (see Bowling et al., 2016; McKay et al., 2018; Ward et al., 2017). Bowling et al. (2016), for instance, found that the extent to which participants engaged in IER displayed rank-order consistency across time (Study 1) and across different situations (Studies 2 and 3); and that four informant-reported FFM traits were significantly correlated with IER (rs = −.22 for conscientiousness, −.18 for agreeableness, −.20 for extraversion, and −.29 for emotional stability in Study 4). As a result, we expect to observe Personality→IER effects.

Disentangling Methodological and Substantive Effects

As we argued above, two mechanisms could contribute to an observed IER←→Scores association: (a) the methodological effect of IER behavior manifesting in artifactually higher/lower personality scores (IER→Scores); and (b) personality substantively drives IER behavior (Personality→IER). At first blush, the presence of a Personality→IER effect appears to alleviate the concern that observed personality scores may be confounded by IER. After all, if personality is driving variance in IER behavior, which in turn introduces variance in a self-report personality measure, then perhaps the resultant personality scores could still be valid. However, a closer inspection of the Personality→IER effect can help refute this intuition. Since the trait that is being measured (i.e., *target trait*) is unlikely the sole driver for IER behavior, and other *irrelevant traits* (i.e., constructs that are not being measured; see Messick, 1989) can account for the Personality→IER effect (e.g., implicit aggression; DeSimone et al., 2020). Thus, the systematic error introduced by IER traces back to these other irrelevant traits. As a result, we expect a methodological effect of IER in observed self-reported personality ratings notwithstanding the potential substantive effect due to the target trait.

We illustrate our expectations in Figure 2. Without considering the Personality→IER effect (upper panel of Figure 2), one might assume the shared variance between IER and self-reported personality scores—shaded areas A and B—to be methodological in nature. However, considering the target trait domain may exert influence on IER behavior (lower panel of Figure 2), we should first remove the

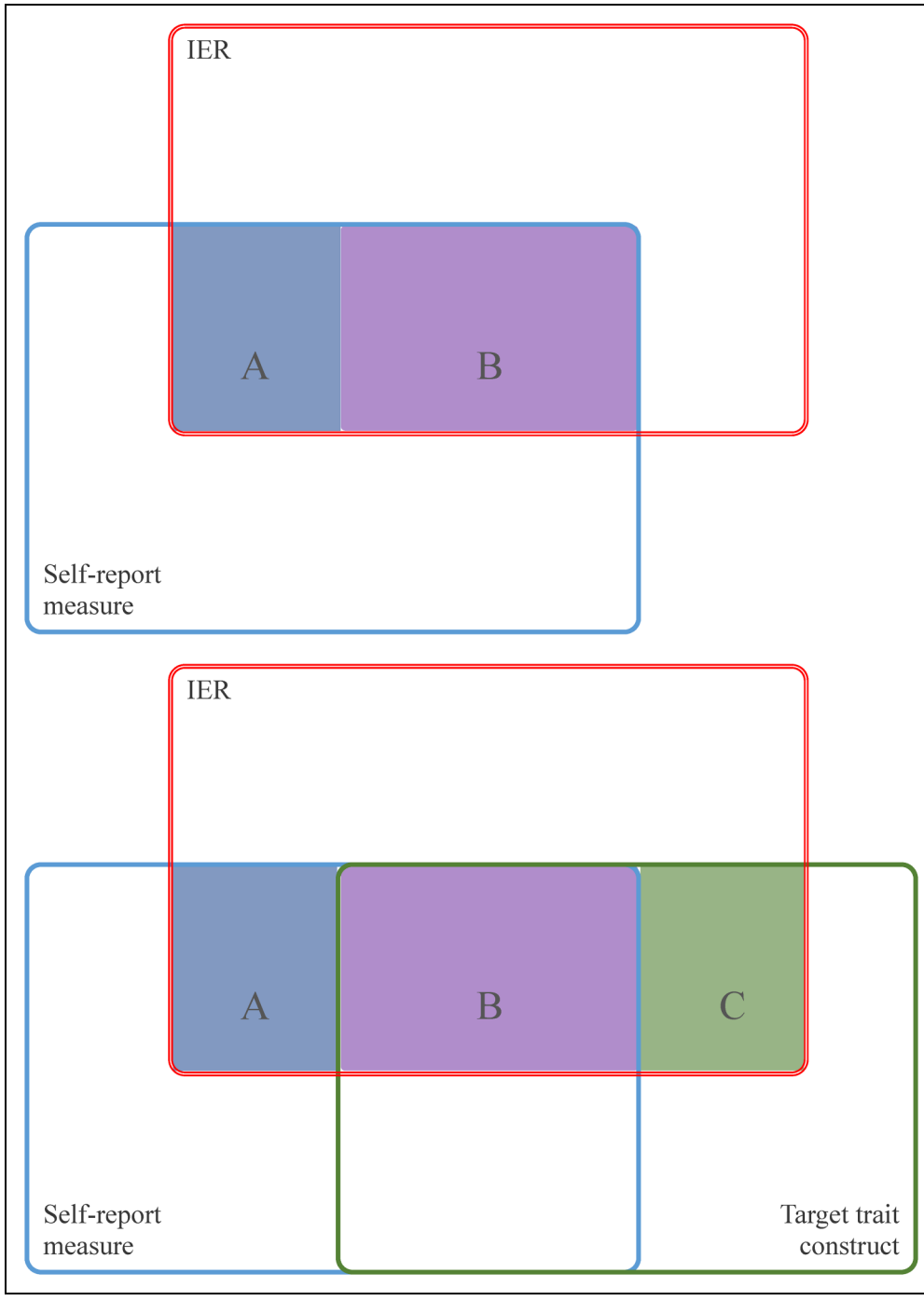

Figure 2. Illustration of shared variance between insufficient effort responding (IER) and personality.

Personality→IER effect (areas B and C) before examining how IER can contaminate personality scores (area A).

To disentangle IER→Scores from Personality→IER in the present study, we obtained participants' self-reports of personality and informant (family members) reports of participants' personality. Using self-reports of personality, we can first obtain an overall estimate of the association between IER and self-reported personality scores, that is, IER←→Scores (path *a* in Figure 1; areas A and B in Figure 2). In addition, we utilize the informant report of personality to capture the substantive Personality→IER effect (path *b* in Figure 1; areas B and C in Figure 2). In the present study, we rely on family members as informants for three reasons. First, with a high level of acquaintance, family members are in a good position to evaluate focal employees' global personality (Biesanz et al., 2007; Funder & Colvin, 1988; Kolar et al., 1996; Vazire & Carlson, 2010; Watson et al., 2000). Second, family members' judgment provides an independent and reasonably accurate account of respondents' personality traits and they are arguably the most accurate informants (Connelly & Ones, 2010; Connolly et al., 2007; Hofstee, 1994; Vazire & Mehl, 2008). Third, family members often have a better sample of the participant's behavior (in terms of time spent with the participant and the variety of different situations) than do other types of informants. Informants from the work setting such as coworkers and supervisors would essentially provide ratings of personality at work (Pathki et al., 2022); therefore, it is questionable if they might accurately report a participant's more general personality. With this approach, we can attempt to remove the effect of general personality on IER from an overall association between IER and self-report personality scores by controlling for informant-reported personality.[4] Thus, any remaining relationship between IER and self-reported personality would be attributable to the methodological IER→Scores effect (i.e., area A). This methodological effect is shown as a significant correlation between IER and a personality trait, after partialling out substantive traits from IER.

Hypothesis 1: Controlling for substantive effects of personality, IER is associated with self-reported personality ratings.

Substantive Effects of IER With Supervisor-Rated Job Performance

In this section, we consider IER's relationship with job performance. Specifically, we discuss three distinct job performance dimensions: (a) task performance, (b) OCB, and (c) CWB (see Rotundo & Sackett, 2002). Previously, we characterized IER as reflecting (a) an unwillingness to expend and sustain the effort necessary to respond to questionnaire content; (b) an unwillingness to comply with questionnaire instructions; and (c) an unwillingness to support the researcher in providing high-quality data. We expect IER will have substantive associations with each of these job performance dimensions (IER←→Performance) because of the conceptual overlap between survey IER and undesirable job performance behaviors. That is, several behavioral tendencies detrimental to performance at work can manifest in the survey response process as IER behavior. Our specific expectations below are grounded in the general observation that people display a moderate degree of behavioral consistencies across various situations (Endler & Magnusson, 1976; Kenrick & Funder, 1988). How a person behaves while completing a questionnaire, in other words, may resemble how he or she behaves at work.

First, demonstrating and sustaining effort at work is a key component to various conceptualizations of job performance. Campbell's (1990) performance taxonomy, for example, recognizes that workers need to commit to their job tasks with an intense and persistent effort to achieve high performance, and Sackett and colleagues (Sackett, 2007; Sackett et al., 1988) highlight how motivation contributes to the performance of daily tasks. Failing to expend or sustain effort on one's job tasks

can lead to low task performance. If such a tendency spills over to survey response, IER behavior ensues, likely in the form of occasional slips due to unwillingness to consistently display effort.

Second, OCB overlaps with attentive respondents' decision to support researchers with high-quality data. OCB is characterized by the willingness to perform voluntary behavior that is beneficial to the organization without being explicitly rewarded (Organ, 1988). Employees' OCB may be directed at individual coworkers in the form of helping and assistance or be directed at the organization in the form of upholding high standards for attendance and adhering to informal norms (Organ, 1997; Williams & Anderson, 1991). Habitually "going the extra mile" without direct personal benefits may be related to respondents' willingness to provide high-quality data. Conversely, absent such a commitment to data quality, IER is more likely to occur.

Finally, there is a conceptual overlap between CWB and IER. CWB involves "voluntary behavior that violates significant organizational norms" in which offending employees fail to behave in accordance with common expectations of behavior in the organizational context (Robinson & Bennett, 1995, p. 556). These behaviors are detrimental to organizational goals and negatively impact the organization or its stakeholders (Ones & Dilchert, 2013). CWBs (Sackett & DeVore, 2002) may include voluntary attempt to sabotage equipment, withdrawing from work activities, and social loafing. Employees who tend to engage in CWBs to the detriment of their work and organizations may be more likely to display a similar lack of concern for the research study and deliberately ignore or violate questionnaire instructions in the responding process.[5]

Hypothesis 2: IER is negatively related to both (a) supervisor-rated task performance and (b) supervisor-rated OCB, and positively related to (c) supervisor-rated CWB.

IER's Effect on the Observed Personality–Performance Relationship

Thus far, we proposed that IER may have (a) a methodological effect on respondents' self-reported personality scores (IER→Scores) despite (b) a substantive association with personality (Personality→IER). We also argued that IER may share (c) a substantive association with supervisor-rated job performance (IER←→Performance). Effects (a) and (c) are likely to occur simultaneously. To illustrate this, imagine that employees' self-reported conscientiousness[6] is being validated against supervisors' ratings of job performance. When a small proportion of employees carelessly respond to self-report personality measures, the resultant conscientiousness scores will contain systematic error due to IER, such that lower scores indicate not only lower standing on conscientiousness (in the majority of employees) but also random/careless selection of response options (in the small proportion of IER employees). At the same time, IER behavior and job performance are related because of their conceptual overlap: The small proportion of IER employees tend to also be those who perform worse than the attentive employees—that is, effect (c) IER←→Performance above. As a result, the observed validity of self-reported conscientiousness in predicting job performance is inflated.

Simply put, IER may serve as a third variable (Shadish et al., 2002) between self-reported personality scores and supervisor-rated job performance; when unaccounted for (Meade et al., 2009), it can inflate the observed personality-performance relationship. Although modeling IER may remove its effect, we need to return to our earlier argument that part of IER may stem from the target trait and thus may not be extraneous. We illustrate this phenomenon in Figure 3. When jointly considering personality, IER, and performance, the part of variance that IER shares with personality (A, B, and C) may overlap with performance, referred to as A_p, B_p, and C_p. Areas B_p and C_p trace back to the target trait, so the overlap can be attributable to the trait exerting influence on both IER and performance. In contrast, area A_p is extraneous to the target trait, and its presence is evidence that IER inflates the association between self-report personality scores and performance ratings. Hence, removing the

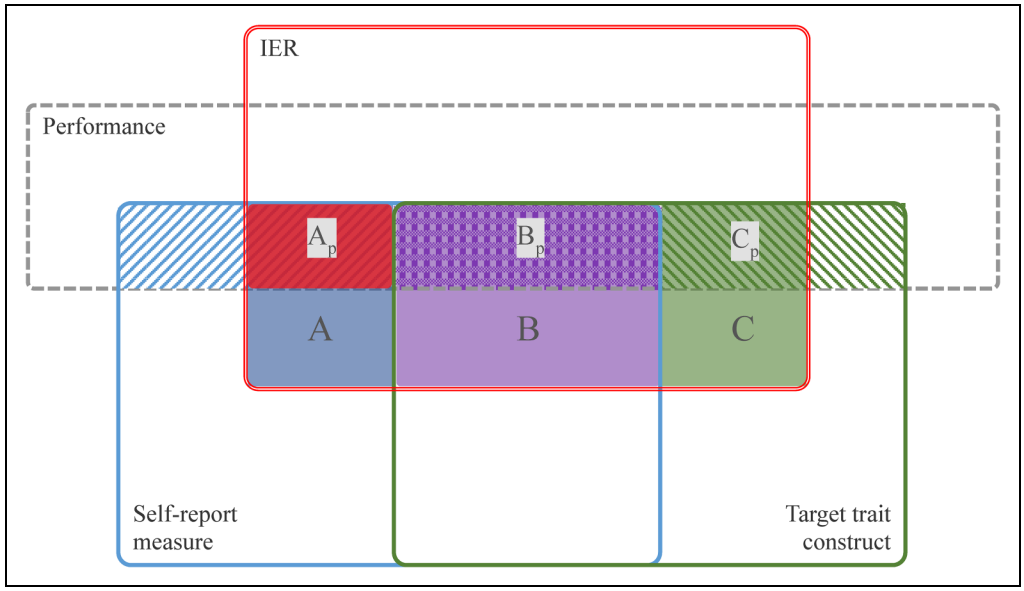

Figure 3. Illustration of shared variance among insufficient effort responding (IER), personality, and performance.

variance associated with A_p will uncover the true relationship between self-reported traits and supervisor-rated performance.

Hypothesis 3: Controlling for the methodological effect of IER will decrease the observed strength of associations between self-reported personality and supervisor-rated (a) task performance, (b) OCB, and (c) CWB.

Method

Participants and Procedure

We collected data using time-lagged online surveys from multiple sources (employees, their family members, and their supervisors). Employees were recruited from a wide range of organizations in the southern United States using similar approaches to other researchers (e.g., Kluemper et al., 2015) that seek to generalize findings to the overall workforce. Following Wheeler et al. (2014), we first recruited students enrolled in introductory management courses at two large public universities in the southern United States to nominate a full-time employee (i.e., someone working at least 30 hr a week), who could not be a fellow student, as a potential participant. Students were instructed to discuss the research project with potential nominees and obtain their permission before nominating them. The students received extra course credit in return, and those who could not recruit an employee were provided an alternative means to earn extra credit. The students gave the research team the names and email addresses of the employee nominees, which enabled direct contact with employee participants and removed the students from the process. Care was taken to ensure that students did not nominate themselves (i.e., putting their own names and email addresses) into the study.

At the first stage of data collection (Time 1), employees rated themselves on FFM traits, verified their full-time employment status, provided demographic data, and gave contact information for their direct supervisor and a family member. Approximately two weeks later (Time 2), a second survey

was sent to the employee's supervisor and family members directly via email with separate surveys asking them to assess the focal employee. Supervisor and family member email addresses were checked to ensure that they did not match those of the employee participants. Family members rated the focal employees on the FFM traits, whereas supervisors rated the employees' performance. Family members and supervisors also provided their own demographic information. We provided an incentive to employees, family members, and supervisors with the possibility of winning a raffle for one of ten $50 gift cards.

Our data collection resulted in 1,111 employee nominees. Of these, 812 (73%) responded to the employee survey. These employees nominated 772 family members and 723 supervisors to participate in the research. We received 503 complete surveys from family members and 517 complete surveys from supervisors. We retained a final sample of 398 participants who provided complete matched data (i.e., an employee survey, a family member survey, and a supervisor survey). Employee respondents were on average 42.6 years old ($SD = 12.9$). They averaged 23.9 years of work experience ($SD = 12.9$) and 10.5 years of tenure with their current organizations ($SD = 9.6$). The employee sample (69.8% female) consisted of 87.2% White, 6.8% Black, and 6.0% other racial/ethnic groups. Family member respondents had an average age of 44.9 years ($SD = 15.1$). The family member sample (61.0% female) consisted of 86.1% White, 6.5% Black, and 7.4% other racial/ethnic groups. Supervisors were on average 46.3 years old ($SD = 12.3$). They averaged 9.6 years of managerial experience ($SD = 23.6$) and 13.7 years of tenure with their organization ($SD = 10.3$). On average, they had supervised the focal employees for 7.2 years ($SD = 11.2$). The supervisor sample (50.1% female) consisted of 87.8% White, 7.6% Black, and 4.6% other racial/ethnic groups.

Measures

Self-rated FFM traits. We assessed the FFM personality traits from employee participants using the 50-item International Personality Item Pool scales (IPIP; Goldberg et al., 2006), with 10 items on each trait. Respondents indicated the extent to which each item accurately described them in general on a seven-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Sample items include "In general, I am always prepared." (Conscientiousness; $\alpha = .86$); "In general, I am relaxed most of the time." (Emotional Stability; $\alpha = .85$); "In general, I have a good word for everyone." (Agreeableness; $\alpha = .82$); "In general, I make friends easily." (Extraversion; $\alpha = .90$); and "In general, I can handle a lot of information." (Openness; $\alpha = .81$).

Family rated FFM traits. We modified the 50-item IPIP (Goldberg et al., 2006) to be rated by family members to assess employee's FFM traits. This was accomplished by altering the referent from "I" to "he/she" in each item. Cronbach's αs were .90 for Conscientiousness, .80 for Emotional Stability, .89 for Agreeableness, .88 for Extraversion, and .83 for Openness.

Task performance. Supervisors rated employees' task performance ($\alpha = .87$) using a five-item abridged scale (Kluemper et al., 2013) adapted from Williams and Anderson (1991). The response scales ranged from 1 (*never*) to 7 (*always*). Each item began with the stem "How often does this employee…?" Example items include "Adequately complete assigned duties" and "Engage in activities that will directly affect his/her performance evaluation."

Organizational citizenship behavior. Supervisors also assessed employees' OCB ($\alpha = .93$) using a 10-item measure developed by Kluemper et al. (2013). This scale was based on Lee and Allen's (2002) OCB measure. It used 7-point anchors (1 = *never*; 7 = *always*). Each OCB item began with the stem "How often does this employee…?" Example items include "Help others who have been absent" and "Willingly give his/her time to help others who have work-related problems."

Counterproductive workplace behavior. Supervisors rated employees' CWB ($\alpha = .91$) using another 10-item measure developed by Kluemper et al. (2013). This scale was based on Bennett and Robinson's (2000) CWB scale. It used 7-point anchors (1 = *never*; 7 = *always*). Each CWB item began with the stem "How often does this employee…?" Example items include "Say something hurtful to someone at work" and "Play a mean prank on someone at work."

Insufficient effort responding. We computed six individual IER indices from participants' self-report data: (a) an instructed-response index (e.g., "Please respond with 'agree' for this item"), (b) an individual-reliability index, (c) a psychometric antonyms index, (d) a psychometric synonyms index, (e) Mahalanobis distance index, and (f) intraindividual response variability (Dunn et al., 2018). Because these indices are each described in detail in previous papers (e.g., Curran, 2016; DeSimone et al., 2015; Meade & Craig, 2012), we describe them in detail in the Appendix.

We used two different approaches to operationalize IER across the six indices. First, we combined the IER indices as continuous variables. As shown in Table 1, the six IER indices displayed varying levels of convergence with each other (mean correlation between IER indices = .39), which is consistent with findings reported elsewhere (e.g., Huang et al., 2012; Maniaci & Rogge, 2014; Meade & Craig, 2012). Combining the individual IER indices into a single composite score (Bowling et al., 2023; Huang & DeSimone, 2021) confers two important advantages over individual IER indices: (a) composites may minimize the limitations unique to any individual IER index and (b) a single analysis using composite scores can replace multiple analyses using individual IER indices, thus simplifying the reporting of results (see Bowling et al., 2016 for a similar treatment).

Our second approach treated each IER index as a trichotomous variable,[7] classifying respondents into one of three categories: attentive (scored as 0), mild IER (scored as 1), and severe IER (scored as 2). This alternative approach focuses on differentiating IER respondents from attentive respondents. We describe the cutoff scores in detail in Supplemental Appendix A, acknowledging the challenge of finding the precise cutoff scores (DeSimone et al., 2015).

We conducted two exploratory factor analyses (EFA) to assess whether we could combine the six IER indices, either in continuous or trichotomous forms. For either approach, inspection of the screeplot, parallel analysis, and Velicer's minimum average partial (MAP) test (Velicer, 1976) all supported a single-factor solution. The respective first factor accounted for 50% of the variance in the six continuous indices and 43% of the variance in the six trichotomous indices. As shown in Table 1, the EFA found that the six continuous IER indices each loaded strongly onto the first factor (mean loading = .63). To compute the composite score for continuous IER, we averaged the standardized scores of the six IER indices ($\alpha = .79$). The EFA on the six trichotomous IER indices showed a similar pattern, with a mean loading on the first factor of .56. We averaged the scores

Table 1. Descriptive Statistics for IER Indices.

IER index	*M*	*SD*	First-factor loading[a]	1	2	3	4	5	6
1. IER composite	0.00	0.72	—	—					
2. Instructed response	0.03	0.13	.48	.61					
3. Reversed individual reliability	−0.81	0.17	.84	.83	.39				
4. Psychometric antonym	−0.79	0.24	.76	.80	.28	.60			
5. Reversed psychometric synonym	−0.89	0.14	.69	.74	.40	.61	.48		
6. Mahalanobis distance	−0.06	0.58	.43	.57	.20	.41	.52	.19	
7. Reversed intraindividual variability	−2.08	0.37	.58	.66	.32	.48	.49	.45	.07

Note: $N = 398$. When $|r| > .09$, $p < .05$; when $|r| > .12$, $p < .01$; when $|r| > .16$, $p < .001$.
Abbreviation: IER = insufficient effort responding.
[a]First-factor loading of each variable obtained from principal axis factor analysis.

of the six trichotomous IER indices into a composite score ($\alpha = .71$). Because the two composites were strongly correlated ($r = .87$) and the results did not differ in meaningful ways, we reported the results below using the IER composite based on continuous indices while presenting the parallel results based on trichotomous indices in the Supplemental Material.

Results

Table 2 reports the correlations between (a) the IER composite, (b) self-reported personality measures, (c) family reported personality measures, and (d) supervisor-rated job performance dimensions. The pattern of correlations we observed is consistent with that found in past research. Each self-reported personality measure, for instance, displayed moderate convergence (average convergence = .35) with the corresponding family reported personality measure (for similar findings, see Connelly & Ones, 2010, Study 2; Connolly et al., 2007). And self-reported personality measures —particularly measures of conscientiousness, emotional stability, and agreeableness—were consistently related to the three supervisor-rated work behaviors we examined (for similar findings, see Barrick et al., 2001; Berry et al., 2007; Chiaburu et al., 2011; Dalal, 2005). These findings thus provide evidence of the validity of the current measures. In addition, IER was moderately to strongly correlated with self-reported personality traits, with *r*s ranging from −.36 to −.62. These correlations are comparable to earlier reports of correlations between IER composite and FFM traits (e.g., *r*s ranging from −.26 to −.59, Huang & DeSimone, 2021), but they are much stronger than other reports (cf., Huang et al., 2015, Study 1; McKay et al., 2018; Ward et al., 2017). For instance, the correlations between individual reliability (reverse-scored) and FFM traits ranged from −.04 to −.22 in the study by Ward et al. (2017). Aside from the possibility that we used a comprehensive approach to measure IER more reliably, the stronger correlations in the present study might be attributed to attentive respondents' relatively higher scores on FFM traits, as the associations between IER and substantive traits depend in large part on $M_{\text{attentive}}$.

Hypothesis Tests

Hypothesis 1. Focusing on the methodological biasing effect of IER in personality scores, Hypothesis 1 stated that after partialling out substantive effects of personality on IER, the residualized IER would still be associated with self-reported personality. Prior to testing the hypothesis, we found evidence of substantive effects of personality on IER, as the IER composite was significantly (ps < .01) related to family reported conscientiousness ($r = -.34$), emotional stability ($r = -.21$), agreeableness ($r = -.21$), extraversion ($r = -.26$), and openness ($r = -.17$). With the exception of openness, these findings are similar to those reported by Bowling et al. (2016, Study 4). Furthermore, regression analyses showed that the five family reported personality traits explained 14% ($p < .001$) of the variance in the IER composite score (see Table 3). Thus, the observed associations between IER and personality are partially the product of a substantive Personality→IER effect.

To remove this substantive Personality→IER effect from the overall IER←→Scores effect, we saved the residual IER score from the regression analysis described above, assuming that family reported personality would reasonably capture the substantive Personality→IER effect. Correlating the residualized IER with self-reported personality showed support for Hypothesis 1: Residualized IER was significantly related to self-reported conscientiousness ($r = -.51$), emotional stability ($r = -.56$), agreeableness ($r = -.48$), extraversion ($r = -.37$), and openness ($r = -.36$), ps < .001. In other words, each of the five self-reported personality traits contained systematic error introduced by the presence of IER, and it is unlikely that the error can be attributed to the target traits being assessed.

Table 2. Descriptive Statistics and Correlations for Personality and Work Outcome Variables.

	M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. IER composite	0.00	0.70	(.79)													
Self-report[a]																
2. Conscientiousness	5.88	0.72	−.59	(.86)												
3. Emotional stability	5.63	0.88	−.62	.48	(.85)											
4. Agreeableness	5.69	0.76	−.52	.31	.43	(.82)										
5. Extraversion	5.26	1.00	−.45	.27	.37	.20	(.90)									
6. Openness	5.19	0.88	−.36	.22	.27	.21	.43	(.81)								
Family report[b]																
7. Conscientiousness	5.93	0.84	−.34	.36	.23	.19	.11	.00	(.90)							
8. Emotional stability	5.13	0.97	−.21	.12	.31	.09	.07	−.01	.43	(.80)						
9. Agreeableness	5.59	0.95	−.21	.05	.16	.32	.03	.01	.43	.51	(.89)					
10. Extraversion	5.49	0.92	−.26	.15	.21	.07	.44	.13	.32	.38	.26	(.88)				
11. Openness	5.06	0.87	−.17	.07	.14	.05	.15	.33	.36	.28	.31	.36	(.83)			
Supervisor-ratings[c]																
12. Task performance	6.48	0.72	−.29	.21	.14	.21	.05	.08	.30	.06	.13	.14	.12	(.87)		
13. OCB	5.92	0.89	−.27	.19	.16	.20	.12	.09	.27	.11	.17	.21	.08	.57	(.93)	
14. CWB	1.46	0.69	.42	−.23	−.20	−.29	−.08	−.09	−.32	−.10	−.21	−.20	−.15	−.48	−.44	(.91)

Note: $N = 398$. Cronbach's alphas are presented in parentheses along the diagonal. When $|r| > .09$, $p < .05$; when $|r| > .12$, $p < .01$; when $|r| > .16$, $p < .001$.

Abbreviations: IER = insufficient effort responding; OCB = organizational citizenship behavior; CWB = counterproductive work behavior.

[a]Variables 2–6 are the focal employee's self-reported personality.

[b]Variables 7–11 are the focal employee's personality reported by one of their family members.

[c]Variables 12–14 are supervisor ratings of the focal employee's work behavior.

Table 3. Predicting IER From Family Reported Personality.

Predictors	*b* (*SE*)	*p*	*β*
Family report conscientiousness	−.21 (.05)	<.001	−.26
Family report emotional stability	−.01 (.04)	.870	−.01
Family report agreeableness	−.04 (.04)	.387	−.05
Family report extraversion	−.13 (.04)	.002	−.16
Family report openness	−.00 (.04)	.933	−.00
R^2		.14***	

Note: $N = 398$.
Abbreviation: IER = insufficient effort responding.
*** $p < .001$.

Hypothesis 2. We predicted that IER would be negatively related to both supervisor-rated task performance and supervisor-rated OCB, and positively related to supervisor-rated CWB. The analyses reported in Table 2 support this hypothesis. The IER composite yielded negative correlations with supervisor ratings of both task performance ($r = -.29$, $p < .001$) and OCB ($r = -.27$, $p < .001$), and a positive correlation with supervisor ratings of CWB ($r = .42$, $p < .001$). Of note, the IER composite in many instances yielded stronger relationships with each of the job performance measures than did either the self-reported or family reported personality measures.

Hypothesis 3. Results for Hypothesis 1 indicating IER's methodological effect on personality scores (IER→Scores)—coupled with Hypothesis 2 finding that the IER composite was related to the three job performance dimensions (IER←→Performance)—lend credence to our prediction that IER acts as a third variable between self-reported personality measures and supervisor-reported performance measures. To the degree that such effects do exist, controlling for the methodological effect of IER would produce significant decreases in the strength of the relationships between self-reported personality and supervisor-rated job performance dimensions (Hypothesis 3).

Recall that the substantive Personality→IER effect can stem from the target trait as well as other irrelevant traits (see Figure 1), and the variance of IER associated with the target trait cannot act as a third variable. In preparing for this analysis, we first removed trait-relevant variance from IER by regressing IER on the family reported target trait and saved the residualized IER for each target trait ($IER_{\cdot \text{family reported target trait}}$). For the target trait of conscientiousness, for instance, IER was regressed on family reported conscientiousness, and the residual $IER_{\cdot \text{family reported conscientiousness}}$ was used for hypothesis testing. Next, for the association between each self-reported target trait and each performance dimension, we computed partial correlations by controlling for the $IER_{\cdot \text{family reported target trait}}$, as "the partial correlation is a valuable tool … for the statistical control of extraneous variables" (Bobko, 2001, p. 172). Results reported in Table 4 (top panel) supported Hypothesis 3. Removing variance associated with $IER_{\cdot \text{family reported target trait}}$ resulted in a significant decrease in the magnitude of association between self-report personality and performance dimensions in all 15 bivariate analyses ($ps < .05$). Of the five personality traits, the largest change came from emotional stability. Whereas self-reported emotional stability had significant zero-order correlations with task performance ($r = .14$, $p < .01$), OCB ($r = .16$, $p < .01$), and CWB ($r = -.20$, $p < .001$), its partial correlation became nonsignificant, $rs = -.03$, .02, and .05, respectively, $ps > .10$. In contrast, after controlling for IER, self-reported conscientiousness and agreeableness remained significantly associated with at least two of performance dimensions, despite the decrease in correlation magnitude. It is of note that the analyses for CWB provided the strongest support for Hypothesis 3. Across the five self-reported personality traits, the average reduction of magnitude from zero-order correlations to partial correlations was |.11| for task performance, |.09| for OCB, and |.16| for CWB.

Table 4. Partial Correlation Between Self-Report Personality and Supervisor-Rated Work Behavior, Controlling for IER.

	Task performance			OCB			CWB		
	r_{xy}	$r_{xy.IER}$	$r_{xy} - r_{xy.IER}$	r_{xy}	$r_{xy.IER}$	$r_{xy} - r_{xy.IER}$	r_{xy}	$r_{xy.IER}$	$r_{xy} - r_{xy.IER}$
Partialling out IER.$_{target\ trait}$									
Conscientiousness	.21***	.13*	.08**	.19***	.12*	.08**	−.23***	−.08	−.15***
Emotional stability	.14**	−.03	.17***	.16**	.02	.14***	−.20***	.05	−.25***
Agreeableness	.21***	.10*	.11***	.20***	.10*	.09***	−.29***	−.14**	−.15***
Extraversion	.05	−.05	.09***	.12*	.05	.07***	−.08	.06	−.14***
Openness	.08	.00	.08***	.09	.01	.08***	−.09	.03	−.13***
Partialling out IER.$_{FFM\ traits}$ (undercorrection)									
Conscientiousness	.21***	.13**	.08**	.19***	.13*	.06*	−.23***	−.08	−.15***
Emotional stability	.14**	.03	.10**	.16**	.08	.08*	−.20***	−.02	−.17***
Agreeableness	.21***	.13**	.08**	.20***	.14**	.06*	−.29***	−.17***	−.12***
Extraversion	.05	−.03	.08***	.12*	.06	.06**	−.08	.04	−.12***
Openness	.08	.01	.07***	.09	.03	.06**	−.09	.02	−.12***
Partialling out IER composite (overcorrection)									
Conscientiousness	.21***	.05	.16***	.19***	.05	.15***	−.23***	.03	−.26***
Emotional stability	.14**	−.06	.20***	.16**	−.01	.17***	−.20***	.10	−.29***
Agreeableness	.21***	.07	.14***	.20***	.07	.12***	−.29***	−.10	−.20***
Extraversion	.05	−.10	.14***	.12*	.00	.12***	−.08	.13	−.22***
Openness	.08	−.02	.11***	.09	−.01	.10***	−.09	.07	−.16***

Note: $N = 398$. r_{xy}: zero-order correlation between a self-reported personality trait and a performance dimension (reproduced from Table 2). $r_{xy.IER}$: partial correlation, controlling for IER. $r_{xy} - r_{xy.IER}$: change from zero-order correlation to partial correlation; significance level based on Olkin and Finn (1995).

Abbreviations: IER = insufficient effort responding; FFM = five-factor model; OCB = organizational citizenship behavior; CWB = counterproductive work behavior.

* $p < .05$. ** $p < .01$. *** $p < .001$.

We acknowledge a potential limitation that family members may not share the unique insight individuals have about their own behavioral tendencies in a given domain, and thus family members' reports and self-reports of a given trait such as conscientiousness may not be interchangeable (Hogan & Holland, 2003). To address this potential concern, we also adopted a more conservative approach to remove family members' holistic view of the employees' personality. Regardless of the trait under examination, we removed variance in IER associated with all five family reported personality traits simultaneously, as in the test for Hypothesis 1 above. The residualized IER ($IER_{\cdot \text{family reported FFM traits}}$) was then partialled out from bivariate associations between self-report personality traits and performance dimensions (Table 4, middle panel). This treatment likely underestimated the biasing effect of IER because it ignored the substantive effect stemming from irrelevant traits. Even these more conservative estimates based on $IER_{\cdot \text{family reported FFM traits}}$ still provided support for Hypothesis 3. The average reduction of magnitude from zero-order correlations to partial correlations was |.08| for task performance, |.06| for OCB, and |.14| for CWB, with all ps < .05.

Finally, our third set of analyses focused on the IER composite, which combined both methodological IER→Scores and substantive Personality→IER, without removing any trait-related variance from IER (Table 4, bottom panel). This is consistent with existing research that sought to partial out IER from observed correlations (e.g., Huang et al., 2015). Because the substantive Personality→IER effect may stem from both the target trait and other irrelevant traits, this analysis may have removed a portion of the substantive variance in target trait measurement, resulting in overcorrection.

Supplemental Analyses Assessing the Removal of IER Cases

Now that we established that the presence of IER in self-report personality data can significantly inflate the observed personality-performance associations, we proceed to explore whether removing suspect IER cases can reduce this inflation. We created two cutoff scores for the composite IER variable that approximate the upper limits of 95% and 80% confidence intervals of the mean. The more conservative cutoff resulted in the removal of 16 individuals (5%), whereas the more stringent cutoff resulted in the removal of 33 individuals (8%). As shown in Table 5, the removal of these IER cases resulted in decreased associations between self-reported personality and supervisor-rated performance.

Table 5. Reducing the Inflation by Removing IER Cases.

	Full sample $N = 398$			Trimmed Sample A $N = 382$			Trimmed Sample B $N = 365$		
	TP	OCB	CWB	TP	OCB	CWB	TP	OCB	CWB
1. Conscientiousness	.21***	.19***	−.23***	.13*	.13**	−.09	.13*	.12*	−.10
2. Emotional stability	.14**	.16**	−.20***	.07	.11*	−.12*	.03	.06	−.10
3. Agreeableness	.21***	.20***	−.29***	.11*	.13**	−.17***	.09	.12*	−.15**
4. Extraversion	.05	.12*	−.08	.01	.09	−.02	−.01	.07	−.01
5. Openness	.08	.09	−.09	.02	.05	−.01	−.01	.03	.01

Note: Trimmed samples A and B were created with cutoff scores that approximate the upper limits of 95% and 80% confidence intervals of the mean on the composite IER variable.

Abbreviations: IER = insufficient effort responding; TP = task performance; OCB = organizational citizenship behavior; CWB = counterproductive work behavior.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Supplemental Analyses Examining the Effects of IER Within a Practical Context

To demonstrate the potential effect of IER within practical settings (e.g., in an employee selection context), we conducted supplemental analyses on conscientiousness, emotional stability, and agreeableness. We focused on these three personality traits because they are most consistently associated with various job performance dimensions (see Barrick et al., 2001; Berry et al., 2007; Chiaburu et al., 2011). Our demonstration is consistent with cumulated research linking conscientiousness and emotional stability to task performance, OCB, and CWB while linking agreeableness to OCB and CWB. We started with the *naïve model*, where each personality trait domain, including family reported personality and self-reported personality, predicted job performance dimensions. The term "naïve model" reflects a situation in which the researcher is unaware of the presence of IER. Results from these naïve models indicate that self-report conscientiousness (Table 6), emotional stability (Table 7), and agreeableness (Table 8) would uniquely contribute to the prediction of relevant job performance outcomes, controlling for their family report counterparts. As there are multiple results to unpack, we turn to the widely established association between self-reported conscientiousness and task performance (Barrick et al., 2001; Barrick & Mount, 1991) as an illustrative example. The naïve model in Table 6 shows that self-reported conscientiousness uniquely predicted task performance ($b = .12$, $\beta = .12$, $p < .05$) in the presence of other-rated conscientiousness.

Next, we examined three hierarchical regression models for each of the personality-performance relationships. First, Model 1 included family reported personality as the sole predictor. Given a family member's close knowledge of the employee's behavioral tendencies (e.g., Connelly & Ones, 2010), Model 1 represents a reasonable assessment of the association between the trait domain and the job performance dimension of interest. Results indicate that the trait-performance relationships were significant in the expected directions, with the exception of the emotional stability-task performance association, which failed to reach significance. Model 1 also serves as a comparison to the naïve model in that the R^2 difference between the two models reflects the unique contribution of self-report personality. Returning to the illustrative example of using conscientiousness to predict task performance (Table 6), we obtained R^2 of .09 for Model 1 and thus concluded that the trait domain of conscientiousness, as measured using family reports, can predict task performance. Further, comparing the R^2 of Model 1 (.09) and the naïve model (.10), we can conclude that self-reported conscientiousness explained 1% (i.e., .01) unique variance in task performance.

Next, Model 2 included both family reported personality and IER composite as predictors. Although the IER composite contained both methodological and substantive components, it is important to note that the additional variance (ΔR^2) explained in this model, compared to Model 1, is not relevant to the trait of interest, which has been partialled out in Model 1. Model 2 serves primarily as a basis for evaluating Model 3.

Model 3 included family reported personality, the IER composite, and self-reported personality. This model allows us to assess whether self-report personality adds a unique contribution to the prediction of job performance outcomes beyond family report, after controlling for methodological and trait-irrelevant variance associated with IER. In the case of conscientiousness as a predictor of task performance, Model 3 indicates that self-reported conscientiousness failed to predict task performance ($b = .00$, $\beta = .00$, $p = .96$) after removing the effects of family report conscientiousness and IER. Recall that self-reported conscientiousness appeared to account for 1% unique variance in task performance earlier in the naïve model. This 1% should now be correctly interpreted as the amount of bias due to IER.

Taken together, Model 3 results provide an interesting contrast to the naïve model. In seven out of eight analyses, self-report personality no longer predicted ($\Delta R^2 = .00$) job performance dimensions after accounting for family report personality and IER—the only exception was that self-reported emotional stability added 1% of the variance in predicting CWB. Thus, we can conclude that the

Table 6. Hierarchical Regression Using Conscientiousness and IER as Predictors of Job Performance Dimensions.

	Naïve model			Model 1			Model 2			Model 3		
Predictors	*b* (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*
Outcome = task performance												
Intercept	4.49 (.32)	.000		4.98 (.25)	.000		5.34 (.26)	.000		5.33 (.39)	.000	
Family report conscientiousness	.22 (.04)	.000	.25	.25 (.04)	.000	.30	.19 (.04)	.000	.23	.19 (.04)	.000	.22
IER							−.22 (.05)	.000	−.21	−.22 (.06)	.000	−.21
Self-report conscientiousness	.12 (.05)	.022	.12							.00 (.06)	.955	.00
R^2 (ΔR^2)	.10*** (—)			.09*** (—)			.13*** (.04***)			.13*** (.00)		
Outcome = OCB												
Intercept	3.69 (.40)	.000		4.25 (.31)	.000		4.67 (.32)	.000		4.64 (.49)	.000	
Family report conscientiousness	.24 (.06)	.000	.23	.28 (.05)	.000	.27	.21 (.05)	.000	.20	.21 (.06)	.000	.20
IER							−.25 (.06)	.000	−.20	−.25 (.08)	.001	−.20
Self-report conscientiousness	.14 (.06)	.033	.11							.01 (.08)	.936	.01
R^2 (ΔR^2)	.08*** (—)			.07*** (—)			.11*** (.04***)			.11*** (.00)		
Outcome = CWB												
Intercept	3.52 (.31)	.000		3.00 (.24)	.000		2.42 (.23)	.000		2.01 (.36)	.000	
Family report conscientiousness	−.22 (.04)	.000	−.27	−.26 (.04)	.000	−.32	−.16 (.04)	.000	−.20	−.17 (.04)	.000	−.21
IER							.35 (.05)	.000	.36	.40 (.06)	.000	.40
Self-report conscientiousness	−.13 (.05)	.010	−.13							.08 (.05)	.135	.09
R^2 (ΔR^2)	.12*** (—)			.10*** (—)			.21*** (.11***)			.22*** (.00)		

Note: $N = 398$.

Abbreviations: IER = insufficient effort responding; OCB = organizational citizenship behavior; CWB = counterproductive work behavior.

Naïve model: conscientiousness, measured as family report and self-report, predicting job performance dimensions; Model 1: family report conscientiousness as a predictor; Model 2: association between IER and job performance dimensions, after controlling for family report of conscientiousness; Model 3: unique contribution of self-report conscientiousness, after controlling for family report conscientiousness and IER.

*** $p < .001$.

Table 7. Hierarchical Regression Using Emotional Stability and IER as Predictors of Job Performance Dimensions.

	Naïve model			Model 1			Model 2			Model 3		
Predictors	*b* (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*
Outcome = task performance												
Intercept	5.79 (.27)	.000		6.25 (.20)	.000		6.48 (.19)	.000		6.77 (.31)	.000	
Family report emotional stability	.02 (.04)	.682	.02	.05 (.04)	.220	.06	.00 (.04)	.976	.00	.01 (.04)	.766	.02
IER							−.30 (.05)	.000	−.29	−.34 (.06)	.000	−.33
Self-report emotional stability	.11 (.04)	.012	.13							−.06 (.05)	.242	−.07
R^2 (ΔR^2)	.02* (—)			.00 (—)			.08*** (.08***)			.09*** (.00)		
Outcome = OCB												
Intercept	4.81 (.33)	.000		5.39 (.24)	.000		5.64 (.24)	.000		5.81 (.39)	.000	
Family report emotional stability	.07 (.05)	.178	.07	.10 (.05)	.026	.11	.05 (.05)	.237	.06	.06 (.05)	.203	.07
IER							−.32 (.06)	.000	−.25	−.35 (.08)	.000	−.27
Self-report emotional stability	.14 (.05)	.010	.14							−.04 (.07)	.592	−.03
R^2 (ΔR^2)	.03**(—)			.01* (—)			.07*** (.06***)			.07*** (.00)		
Outcome = CWB												
Intercept	2.44 (.25)	.000		1.83 (.19)	.000		1.51 (.17)	.000		1.05 (.29)	.000	
Family report emotional stability	−.03 (.04)	.386	−.05	−.07 (.04)	.045	−.10	−.01 (.03)	.775	−.01	−.03 (.03)	.457	−.04
IER							.41 (.05)	.000	.42	.48 (.05)	.000	.49
Self-report emotional stability	−.14 (.04)	.000	−.18							.10 (.05)	.043	.12
R^2 (ΔR^2)	.04*** (—)			.01* (—)			.18*** (.17***)			.19*** (.01*)		

Note: $N = 398$.

Abbreviations: IER = insufficient effort responding; OCB = organizational citizenship behavior; CWB = counterproductive work behavior.

Naïve model: emotional stability, measured as family report and self-report, predicting job performance dimensions; Model 1: family report emotional stability as a predictor; Model 2: association between IER and job performance dimensions, after controlling for family report of emotional stability; Model 3: unique contribution of self-report emotional stability, after controlling for family report emotional stability and IER.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 8. Hierarchical Regression Using Agreeableness and IER as Predictors of Job Performance Dimensions.

Predictors	Naïve model			Model 1			Model 2			Model 3		
	b (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*	*b* (*SE*)	*p*	*β*
Outcome = OCB												
Intercept	4.24 (.37)	.000		5.03 (.26)	.000		5.29 (.26)	.000		5.02 (.41)	.000	
Family report agreeableness	.11 (.05)	.022	.12	.16 (.05)	.001	.17	.11 (.05)	.015	.12	.10 (.05)	.033	.11
IER							−.31 (.06)	.000	−.24	−.28 (.07)	.000	−.22
Self-report agreeableness	.19 (.06)	.002	.16							.06 (.07)	.403	.05
R^2 (ΔR^2)	.05*** (—)			.03*** (—)			.08*** (.06***)			.09*** (.00)		
Outcome = CWB												
Intercept	3.30 (.28)	.000		2.31 (.20)	.000		1.98 (.19)	.000		2.29 (.30)	.000	
Family report agreeableness	−.09 (.04)	.011	−.13	−.15 (.04)	.000	−.21	−.09 (.03)	.006	−.13	−.08 (.04)	.021	−.11
IER							.39 (.05)	.000	.40	.36 (.05)	.000	.36
Self-report agreeableness	−.23 (.05)	.000	−.25							−.07 (.05)	.191	−.07
R^2 (ΔR^2)	.10*** (—)			.04*** (—)			.19*** (.15***)			.20*** (.00)		

Note: $N = 398$.

Abbreviations: IER = insufficient effort responding; OCB = organizational citizenship behavior; CWB = counterproductive work behavior.

Naïve model: agreeableness, measured as family report and self-report, predicting job performance dimensions; Model 1: family report agreeableness as a predictor; Model 2: association between IER and job performance dimensions, after controlling for family report of agreeableness; Model 3: unique contribution of self-report agreeableness, after controlling for family report agreeableness and IER.

***$p < .001$.

significant contribution of self-report personality in the naïve models is largely an artifact introduced by the presence of IER.

Discussion

Recent research has established that self-report measures commonly used in organizational research can be contaminated by IER (e.g., Credé, 2010; Huang et al., 2015; McGonagle et al., 2016). Building on the finding that IER can reflect individual behavioral tendencies (Bowling et al., 2016), we examined the presence of IER as a confounding factor between self-reported personality and supervisor-rated performance, demonstrating IER as a characteristic associated with the respondent as opposed to the method of data collection. The current study provided evidence of the hypothesized effect of IER across three dimensions of job performance—task performance, OCB, and CWB. Specifically, we found evidence that (a) IER had a methodological effect on self-reported personality scores, after removing the substantive effect of target personality on IER; (b) IER had substantive associations with task performance, OCB, and CWB; and (c) personality-task performance, personality-OCB, and personality-CWB relationships were each significantly weakened after the methodological effects of IER were controlled.

Implications for Research

The present study provides valuable knowledge for research on the association between self-reported and other-reported variables in general, and more specifically, personality and performance at work. Cumulative research—built on a preponderance of studies using self-reported personality—suggests that conscientiousness and emotional stability are valid predictors of performance across all jobs (see Barrick et al., 2001; Barrick & Mount, 1991; Tett et al., 1991). Our findings suggest that researchers may need to revisit this association: controlling for the methodological effects of IER significantly weakened the relationships between self-reported personality traits and supervisor-ratings of task performance, OCB, and CWB. Our findings thus provide new insights into why self-reported variables such as personality measures often yield zero-order relationships with job performance. Although the predominant view is that self-reported personality measures predict job performance because they provide accurate assessments of respondents' underlying personality traits, our results offer an alternative possibility. That is, self-reported constructs including personality traits may appear related to job performance when they are contaminated by IER, and a wide array of established relationships found in the literature are likely inflated because of the presence of IER. An interesting note is the contrast in the present study between conscientiousness and emotional stability. We found that self-reports of conscientiousness, but not emotional stability, continued to share significant variance with job performance dimensions after partialling out the methodological effect of IER. In other words, the current findings suggest that, while self-reported conscientiousness remains a valid predictor, the association between self-reported emotional stability and supervisor-reported job performance appears largely spurious.

With the findings of the current study in mind, we call for additional research comparing the validities of self-report measures between predictive and concurrent validation designs. When using comparable designs for predictive and concurrent studies, past research tends to show slightly higher validities in concurrent samples than in predictive samples (e.g., Gupta et al., 2013; Hough, 1998; Weekley et al., 2004). van Iddekinge and Ployhart (2008, p. 890), for instance, made the following observation on Hough's (1998) report: "Across criteria, observed correlations were between .04 and .15 smaller for predictive designs than for concurrent designs, with an average difference of .07," while Weekley et al. (2004, p. 447) concluded that even after correcting for direct range restriction in the applicant sample, "the validities … for the personality measures tended to be slightly lower in

the applicant sample than in the incumbent sample." Although researchers have readily acknowledged lower test-taking motivation and increased carelessness among incumbents relative to applicants (e.g., van Iddekinge & Ployhart, 2008, p. 894), such a motivational difference has only been linked to *observed score difference* between applicants and incumbents (e.g., Weekley et al., 2004). Adding a missing piece of the puzzle, our study shows that low test-taking motivation (and more specifically IER) in incumbents can result in inflated observed validities. In contrast, applicants are much less likely to engage in IER, making IER unlikely to act as a third variable between applicants' self-reported personality and their performance. With this new knowledge, researchers can now examine IER as one of the many factors (see Gupta et al., 2013) that lead to higher validities in concurrent designs.

Our study also provides new insight into the relative validity of self- versus other-reports of personality in predicting job performance. Based on meta-analytic evidence, Oh et al. (2011) concluded that observer-reports of personality had on average .10 higher validity than self-reports for predicting overall job performance. The zero-order correlations in our study showed a similar pattern in four of the five trait domains, with emotional stability being the sole exception. Indeed, it may be no coincidence that self-reported emotional stability appeared to be the most contaminated by IER among the FFM in the present study. Considering the possibility that some previous validation studies may have observed an inflated association between self-reports of personality and job performance due to the confounding effect of IER, it is likely that the .10 difference in validity in Oh et al. (2011) is a lower bound estimate. Furthermore, compared to task performance and OCB, IER had a stronger association with CWB, and consequently, controlling for IER resulted in an especially pronounced change in the association between self-reported personality and CWB. This may have occurred because IER and CWB share considerable conceptual overlap: both variables may reflect the extent to which a person is uncooperative and deliberately deviant, albeit in different contexts. Removing the methodological confounding effect of IER in personality-CWB associations leads to the conclusion that agreeableness and conscientiousness are both negatively associated with supervisor-reported CWB, echoing Berry et al.'s (2012) meta-analytic conclusion.

As noted above, we focused on the confounding effect of IER in the context of using self-reported personality to predict supervisor-rated performance, but the current findings have implications for broader research in organizations where self-report surveys are utilized. For instance, employees' experience of fair treatment at work, typically captured using self-reports, has been linked to their task performance and OCB (Colquitt et al., 2013). When designing a questionnaire study to examine the mechanisms linking employees' justice perceptions to their performance, a researcher may well recognize the issue of common method bias and obtain supervisors' ratings of performance. However, the undetected presence of IER in employees' self-report justice perceptions may nevertheless inflate the observed association between justice perceptions and supervisor-rated performance. Considering the potential impact of IER, we encourage researchers to adhere to best practices in safeguarding the quality of survey data (Curran, 2016; DeSimone et al., 2015). Furthermore, although the methodological effect of IER depends in large part on attentive respondents' standing on a given substantive measure, researchers may begin documenting organizational survey variables that tend to have means away from the scale midpoint and thus are more susceptible to IER's biasing effect.

Implications for Practice

The current study has direct implications for the validation of self-report personality in employee selection. As noted above, validity coefficients may be inflated in concurrent designs compared to predictive designs due, in part, to IER. Organizations interested in validating selection instruments may react by choosing a predictive design as a preferred option. Given the ease of conducting

validation studies with current employees, organizations may still opt to use a concurrent design, and when doing so, they should take precautionary steps to circumvent IER as a confounding variable between a self-reported predictor and the criterion. First, care should be taken to encourage employees participating in a validation study to not only respond to the questionnaire but to also provide high-quality data. One way to achieve this is to raise the perceived stakes for the validation study, such as making the survey responding part of the employees' job. Second, although nonresponse bias remains a concern (e.g., Rogelberg et al., 2000), it may be a viable strategy to make it easy for employees to opt out of a validation study without any perceived repercussions, as opposed to providing IER data under the perceived pressure to respond. Finally, concurrent validation efforts should include a combination of detection methods to screen for IER (see Curran, 2016; DeSimone et al., 2015; Meade & Craig, 2012).

Beyond the selection context, our findings can inform the design of a wide range of studies using self-report measures in conjunction with non-self-report measures. Similar to the logic we presented earlier linking IER to the performance dimensions, IER may be related to a wide range of independent variables beyond personality as well as numerous outcome variables such as tasks that require effort, attention, and persistence, and thus can introduce spurious associations with self-reported measures. For instance, consultants and practitioners alike assess the extent to which knowledge, skills, and attitudes of employees gained in the training context translates to improved trainee job performance (Baldwin & Ford, 1988). When such training transfer validation efforts consist of self-reported and non-self-reported measures, IER may be an issue. As another example, if an organizational survey reveals a positive association between job engagement and employee safety compliance, the association might be spurious due to the confounding effect of IER—employees who respond carelessly on the job engagement survey also tend to fail to adhere to safety guidelines. Thus, organizational survey results should be interpreted in light of the potential threat of IER in inflating Type I error rates across survey sources. As such, screening for IER is a reasonable starting point in analyzing organizational surveys.

Limitations and Future Research

Three limitations of the current study should be noted. First, we used family reported personality to control for trait-relevant variance in IER, assuming the remaining relationship between IER and self-reported personality represented a methodological effect of IER on self-reported personality (IER→Personality). Our assumption is based on research indicating family reports of personality tend to be accurate (for meta-analytic evidence, see Connelly & Ones, 2010; Connolly et al., 2007). However, one may argue that self-report may capture unique insight about oneself that others cannot easily capture. Indeed, socioanalytic theory (Hogan & Holland, 2003) distinguishes identity from reputation, with the former reflecting internal dynamics that are typically associated with self-report whereas the latter representing summaries of behavioral patterns accessible to observers. Although we bolstered our conclusion with additional analyses that controlled for all five family reported traits—thus removing family members' holistic view of the focal employees' personality from a single self-reported trait under examination—it remains a question whether doing so is sufficient to remove all trait-relevant variance. We hasten to note that potential alternatives to the current design would be limited in other regards. For instance, after obtaining self-reported personality data under the current survey instructions, one could administer the same measures with a warning message against IER. However, given the general weak effect of warning on IER (Huang et al., 2012; Meade & Craig, 2012), the second administration may still be somewhat confounded by IER. This would be especially problematic since IER behavior displays rank-order consistency over time (Bowling et al., 2016). Another possibility is to assess the same employees again in a high-stakes environment to obtain an accurate assessment of their identity. However,

raising the stakes in assessment would introduce socially desirable responses at the same time. We call for future research to address this design challenge.

Second, our study differs from organization-sponsored studies in that the participating employees understood that their data would be used for this research study only and would not bear any personal consequences. Participants in a validation study sponsored by the organization may be more motivated to respond carefully for different reasons. For instance, employees may be interested in providing input that may influence the organization's selection practice, or they may refrain from IER behavior for fear that poor responses may be identified. The current participants, in other words, may have had less incentive to respond carefully than participants for a validation study.

Third, the self-report personality measures in our study contained many negatively keyed items, and when most or all of the items are scored in the same direction, IER may exert a somewhat different impact. For instance, with limited negatively keyed items, longstring responding or straightlining (DeSimone et al., 2018) may lead to M_{IER} being higher than the midpoint of the response scale because of the general tendency for longstring respondents to indicate agreement than disagreement.

Our study lays the groundwork for future research. Future research may examine IER as a source of difference between applicants' and incumbents' responses to personality measures. Because applicants respond to personality measures in a high-stakes context, they are less likely to engage in IER behavior compared to incumbents. Indeed, a body of research exists on socially desirable responding or faking (e.g., Ones et al., 1996; Schmit & Ryan, 1993) in the selection setting. In sharp contrast, incumbents typically respond to personality measures in a low-stakes context, giving rise to potential IER behavior that can lead to inflated validity coefficients. Future studies should simultaneously consider the presence of socially desirable responding and IER in the respective samples to better understand the validities of personality measures.

The conceptual overlap between IER and deviant behavior provides another input for future research. Researchers may be especially concerned about workers who tend to engage in IER and exhibit deviant behavior (at work, in life, or in a laboratory setting), and it remains unknown whether these workers can be retained in a study without compromising study findings. Future studies are needed to develop methods to identify these respondents and prevent their IER behavior.

Future studies should further examine the potential causes of IER. Such research may find the main effects for both personality traits (e.g., conscientiousness; see Bowling et al., 2016) and situational variables (e.g., warning participants that they will be punished for engaging in IER; see Huang et al., 2012), as well as Personality × Situation interactions. Situational strength could provide a lens for understanding such interaction effects. According to this perspective (see Meyer et al., 2010 and Withey et al., 2005 for general reviews on situational strength), situations vary from each other in their strength: Weak situations provide few clues about how one should behave, and there are few consequences to behaving incorrectly; strong situations provide obvious cues about how one should behave, oftentimes with serious consequences for behaving incorrectly. Simply put, weak situations provide people with the freedom to behave according to their personalities. As a result, personality traits are more likely to predict behavior—including IER—within weak, rather than strong, situations.

To illustrate such an interaction, consider a potential respondent conscientiousness × warning manipulation effect on IER. In a weak situation, such as a low-stakes survey setting, survey respondents have the latitude to engage in an amount of IER that corresponds with their level of conscientiousness. Thus, conscientiousness will be negatively related to IER. In contrast, in a strong situation, such as when a stern warning is present, most respondents will be deterred from IER behavior, regardless of their level of conscientiousness. Future IER research should examine factors that produce either strong (e.g., the presence of a warning, in-person data collection) or weak (e.g., the absence of a warning, remote data collection) situations (see Bowling et al., 2021). Such research would not only provide insights into the fundamental nature of IER, but it could also be used to inform IER prevention strategies.

Finally, an anonymous reviewer alerted us to the possibility that IER behavior may be partially a result of relationships within a family. Considering research that shows low to moderate personality similarity between parents and their children (Bratko & Marušić, 1997), between siblings (Vernon & Jang, 1993), and between spouses (Caspi et al., 1992), we suspect a person's tendency to engage in IER could share some similarity with their family members. Because of the brevity of the family reported personality inventory included in our study, we did not embed attention checks and could not compute posthoc IER indices in the family reported personality. We call for future research to investigate whether IER "runs in the family."

Conclusion

Although recent research has raised awareness about the potential confounding role of IER between self-report measures, it was previously unknown how IER might affect the association between self-report predictors and informant-rated criteria. We described the mechanisms by which IER may act as a third variable between self-reported personality and supervisor-reported performance. Our study based on multisource data from 398 employees demonstrates that IER can inflate the observed association between self-reports and supervisor-reports. The present findings add to the understanding of the potential threat to research conclusions posed by IER.

Appendix

Description of IER Indices

Employees' self-report data allowed for the computation of six IER indices: (a) an instructed-response index, (b) an individual reliability index, (c) a psychometric antonyms index, (d) a psychometric synonyms index, (e) Mahalanobis distance index, and (f) intraindividual response variability index. We describe these indices below.

Instructed response index: We followed Meade and Craig's (2012) instructed response approach and scattered in the survey three instructed response items (Meade & Craig, 2012) that asked participants to respond in a particular manner. For example, one of the items asked respondents to "Please respond with 'disagree' for this item. We coded correct responses as 0 and incorrect responses as 1. Scores across the three items were averaged into the instructed response index.

Individual reliability index: Individual reliability (Jackson, 1976) relies on the assumption that attentive individuals' responses to two halves of the same substantive measure should be highly correlated. Recent research has shown that individual reliability (also called even-odd consistency) can effectively detect IER (Huang et al., 2012; Meade & Craig, 2012). In the present study, we first separated odd- and even-numbered items on each of the 15 trait measures and calculated half-scale scores (Huang et al., 2012; Meade & Craig, 2012). A within-person correlation between odd- and even-numbered half scales was computed for each respondent, with a higher correlation indicating consistent responding (and a lower likelihood of IER). We multiplied the within-person correlations by -1 to obtain the individual reliability index such that high scores would indicate IER.

Psychometric antonym index: The psychometric antonym approach (Huang et al., 2012; Meade & Craig, 2012) follows a simple logic: Items that are strongly and negatively correlated are psychometrically opposite to each other. An attentive participant's answers to two psychometric antonyms are expected to be dissimilar to each other. Following this logic, we first identified 30 item pairs that shared the strongest negative associations with each other. We then computed a within-person correlation for each respondent across the 30 item pairs. As attentive respondents were expected to display low (i.e., negative) within-person correlations, we adopted the within-person correlations as the psychometric antonym index.

Psychometric synonym index: The psychometric synonym index (Meade & Craig, 2012) follows a logic similar to the psychometric antonym index, with the exception that 30 items that shared the largest positive correlations were identified. As the within-person correlations across the 30 psychometric synonyms should be positive and strong for attentive respondents, we reversed the within-person correlations by multiplying them with −1 such that higher scores would indicate stronger IER behavior.

Mahalanobis distance index: Following Meade and Craig (2012), we computed Mahalanobis distance to detect multivariate outliers (i.e., improbable response patterns) on each substantive measure's items. The Mahalanobis distance across all trait measures was averaged into the Mahalanobis distance index, with higher scores indicating a greater likelihood of IER.

Intraindividual response variability index: Following Dunn et al. (2018), we calculated intraindividual response variability as the standard deviation of response options endorsed throughout the entire survey. We reverse-scored intraindividual response variability such that higher scores represent low variability across items and thus higher likelihood of IER behavior.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Jason L. Huang https://orcid.org/0000-0001-9548-0542
Nathan A. Bowling https://orcid.org/0000-0003-3768-4939
Benjamin D. McLarty https://orcid.org/0000-0003-0247-558X
Donald H. Kluemper https://orcid.org/0000-0002-0499-3025
Zhonghao Wang https://orcid.org/0000-0002-4742-042X

Supplemental Material

Supplemental material for this article is available online.

References

Aguinis, H., & Vandenberg, R. J. (2014). An ounce of prevention is work a pound of cure: Improving research quality before data collection. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 69-95. https://doi.org/10.1146/annurev-orgpsych-031413-091231

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. American Psychological Association.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, *41*(1), 63-105. https://doi.org/10.1111/j.1744-6570.1988.tb00632.x

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1-26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*(1-2), 9-30. https://doi.org/10.1111/1468-2389.00160

Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology*, *85*(3), 349-360. https://doi.org/10.1037/0021-9010.85.3.349

Berry, C. M., Carpenter, N. C., & Barratt, C. L. (2012). Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *Journal of Applied Psychology*, *97*(3), 613-636. https://doi.org/10.1037/a0026739

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. *Journal of Applied Psychology*, *92*(2), 410-424. https://doi.org/10.1037/0021-9010.92.2.410

Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self-other agreement in judgments of personality. *Journal of Personality and Social Psychology*, *92*(1), 119-135. https://doi.org/10.1037/0022-3514.92.1.119

Bobko, P. (2001). *Correlation and regression: Applications for industrial organizational psychology and management*. Sage.

Bowling, N. A., Gibson, A. M., Houpt, J. W., & Brower, C. K. (2021). Will the questions ever end? Person-level increases in careless responding during questionnaire completion. *Organizational Research Methods*, *24(4)*, 718-738. https://doi.org/10.1177/1094428120947794

Bowling, N. A., Huang, J. L., Bragg, C. B., Khazon, S., Liu, M., & Blackmore, C. E. (2016). Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality. *Journal of Personality and Social Psychology*, *111*(2), 218-229. https://doi.org/10.1037/pspp0000085

Bowling, N. A., Huang, J. L., Brower, C. K., & Bragg, C. B. (2023). The quick and the careless: The construct validity of page time as a measure of insufficient effort responding to surveys. *Organizational Research Methods*, *26*(2), 323–352. https://doi.org/10.1177/10944281211056520

Bratko, D., & Marušić, I. (1997). Family study of the big five personality dimensions. *Personality and Individual Differences*, *23*(3), 365–369. https://doi.org/10.1016/S0191-8869(97)80001-3

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 687-732). Consulting Psychologists Press.

Caspi, A., Herbener, E. S., & Ozer, D. J. (1992). Shared experiences and the similarity of personalities: A longitudinal study of married couples. *Journal of Personality and Social Psychology*, *62*(2), 281–291. https://doi.org/10.1037/0022-3514.62.2.281

Chiaburu, D. S., Oh, I.-S., Berry, C. M., Li, N., & Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: A meta-analysis. *Journal of Applied Psychology*, *96*(6), 1140-1166. https://doi.org/10.1037/a0024004

Colquitt, J. A., Scott, B. A., Rodell, J. B., Long, D. M., Zapata, C. P., Conlon, D. E., & Wesson, M. J. (2013). Justice at the millennium, a decade later: A meta-analytic test of social exchange and affect-based perspectives. *Journal of Applied Psychology*, *98*(2), 199-236. https://doi.org/10.1037/a0031757

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*(6), 1092-1122. https://doi.org/10.1037/a0021212

Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, *15*(1), 110-117. https://doi.org/10.1111/j.1468-2389.2007.00371.x

Credé, M. (2010). Random responding as a threat to the validity of effect size estimates in correlational research. *Educational and Psychological Measurement*, *70*(4), 596-612. https://doi.org/10.1177/0013164410366686

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4-19. https://doi.org/10.1016/j.jesp.2015.07.006

Dalal, R. S. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology*, *90*(6), 1241-1255. https://doi.org/10.1037/0021-9010.90.6.1241

DeSimone, J. A., Davison, H. K., Schoen, J. L., & Bing, M. N. (2020). Insufficient effort responding as a partial function of implicit aggression. *Organizational Research Methods*, *23*(1), 154-180. https://doi.org/10.1177/109442811879948

DeSimone, J. A., DeSimone, A. J., Harms, P. D., & Wood, D. (2018). The differential impacts of two forms of insufficient effort responding. *Applied Psychology: An International Review*, *67*(2), 309-338. https://doi.org/10.1111/apps.12117

DeSimone, J. A., Harms, P. D., & DeSimone, A. J. (2015). Best practice recommendations for data screening. *Journal of Organizational Behavior*, *36*(2), 171-181. https://doi.org/10.1002/job.1962

Dunn, A. M., Heggestad, E. D., Shanock, L. R., & Theilgard, N. (2018). Intra-individual response variability as an indicator of insufficient effort responding: Comparison to other indicators and relationships with individual differences. *Journal of Business and Psychology*, *33*(1), 105-121. https://doi.org/10.1007/s10869-016-9479-0

Endler, N. S., & Magnusson, D. (1976). Toward an interactional psychology of personality. *Psychological Bulletin*, *83*(5), 956-974. https://doi.org/10.1037/0033-2909.83.5.956

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology*, *55*(1), 149-158. https://doi.org/10.1037/0022-3514.55.1.149

Goddard, R. D., & Villanova, P. (2006). Designing surveys and questionnaires for research. In F. T. L. Leong, & J. T. Austin (Eds.), *The psychology research handbook: A guide for graduate students and research assistants* (pp. 114-125). Sage.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84-96. https://doi.org/10.1016/j.jrp.2005.08.007

Gupta, N., Ganster, D. C., & Kepes, S. (2013). Assessing the validity of sales self-efficacy, a cautionary tale. *Journal of Applied Psychology*, *98*(4), 690-700. https://doi.org/10.1037/a0032232

Hofstee, W. K. (1994). Who should own the definition of personality? *European Journal of Personality*, *8*(3), 149-162. https://doi.org/10.1002/per.2410080302

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, *88*(1), 100-112. https://doi.org/10.1037/0021-9010.88.1.100

Hough, L., & Dilchert, S. (2017). Personality: Its measurement and validity for employee selection. In J. L. Farr, & N. T. Tippins (Eds.), *Handbook of employee selection* (pp. 298-325). Routledge.

Hough, L. M. (1998). Personality at work: Issues and evidence. In M. D. Hakel (Ed.), *Beyond multiple choice: Evaluating alternatives to traditional testing for selection* (pp. 131-166). Erlbaum.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, *75*(5), 581-595. https://doi.org/10.1037/0021-9010.75.5.581

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, *27*, 99-114. https://doi.org/10.1007/s10869-011-9231-8

Huang, J. L., & DeSimone, J. A. (2021). Insufficient effort responding as a potential confound between survey measures and objective tests. *Journal of Business and Psychology*, *36*, 807-828. https://doi.org/10.1007/s10869-020-09707-2

Huang, J. L., Liu, M., & Bowling, N. A. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology*, *100*(3), 828-845. https://doi.org/10.1037/a0038510

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, *85*(6), 869-879. https://doi.org/10.1037/0021-9010.85.6.869

Jackson, D. N. (1976). *The appraisal of personal reliability* [Conference session]. The Annual Meeting of the Society of Multivariate Experimental Psychology, University Park, PA.

Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of Research in Personality*, *39*(1), 103-129. https://doi.org/10.1016/j.jrp.2004.09.009

Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist*, *43*(1), 23-34. https://doi.org/10.1037/0003-066X.43.1.23

Kluemper, D. H., DeGroot, T., & Choi, S. (2013). Emotion management ability: Predicting task performance, citizenship, and deviance. *Journal of Management*, *39*(4), 878-905. https://doi.org/10.1177/0149206311407326

Kluemper, D. H., McLarty, B. D., & Bing, M. N. (2015). Acquaintance ratings of the big five personality traits: Incremental validity beyond and interactive effects with self-reports in the prediction of workplace deviance. *Journal of Applied Psychology*, *100*(1), 237-248. https://doi.org/10.1037/a0037810

Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality*, *64*(2), 311-337. https://doi.org/10.1111/j.1467-6494.1996.tb00513.x

Lee, K., & Allen, N. J. (2002). Organizational citizenship behavior and workplace deviance: The role of affect and cognitions. *Journal of Applied Psychology*, *87*(1), 131-142. https://doi.org/10.1037/0021-9010.87.1.131

Lucas, R. E., & Baird, B. M. (2006). Global self-assessment. In M. Eid, & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 29-42). American Psychological Association. https://doi.org/10.1037/11383-003

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality*, *48*, 61-83. https://doi.org/10.1016/j.jrp.2013.09.008

McGonagle, A. K., Huang, J. L., & Walsh, B. M. (2016). Insufficient effort survey responding: An underappreciated problem in work and organizational health psychology research. *Applied Psychology: An International Review*, *65*(2), 287-321. https://doi.org/10.1111/apps.12058

McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, *136*(3), 450-470. https://doi.org/10.1037/a0019216

McKay, A. S., Garcia, D. M., Clapper, J. P., & Shultz, K. S. (2018). The attentive and the careless: Examining the relationship between benevolent and malevolent personality traits with careless responding in online surveys. *Computers in Human Behavior*, *84*, 295-303. https://doi.org/10.1016/j.chb.2018.03.007

Meade, A. W., Behrend, T. S., & Lance, C. E. (2009). Dr. StrangeLOVE, or: How I learned to stop worrying and love omitted variables. In C. E. Lance, & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 89-106). Routledge/Taylor & Francis Group.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*(3), 437-455. https://doi.org/10.1037/a0028085

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*(2), 5-11. https://doi.org/10.3102/0013189X018002005

Meyer, R. D., Dalal, R. S., & Hermida, R. (2010). A review and synthesis of situational strength in the organizational sciences. *Journal of Management*, *36*(1), 121-140. https://doi.org/10.1177/0149206309349309

Oh, I.-S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*, *96*(4), 762-773. https://doi.org/10.1037/a0021832

Olkin, I., & Finn, J. D. (1995). Correlations redux. *Psychological Bulletin*, *118*(1), 155-164. https://doi.org/10.1037/0033-2909.118.1.155

Ones, D. S., & Dilchert, S. (2013). Counterproductive work behaviors: Concepts, measurement, and nomological network. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA Handbook of testing and assessment in psychology: Test theory and testing and assessment in industrial and organizational psychology* (pp. 643-659). American Psychological Association. https://doi.org/10.1037/14047-035

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*(6), 660-679. https://doi.org/10.1037/0021-9010.81.6.660

Oppenheim, A. N. (1992). *Questionnaire design, interviewing and attitude measurement*. Pinter Publishers.

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington Books.

Organ, D. W. (1997). Organizational citizenship behavior: It's construct clean-up time. *Human Performance*, *10*(2), 85-97. https://doi.org/10.1207/s15327043hup1002_2

Pathki, C. S., Kluemper, D. H., Meuser, J. D., & McLarty, B. D. (2022). The Org-B5: Development of a short work frame-of-reference measure of the Big Five. *Journal of Management*, *48*(5), 1299-1337. https://doi.org/10.1177/01492063211002627

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903. https://doi.org/10.1037/0021-9010.88.5.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539-569. https://doi.org/10.1146/annurev-psych-120710-100452

Robinson, S., & Bennett, R. (1995). A typology of deviant workplace behaviors: A multi-dimensional scaling study. *Academy of Management Journal*, *38*(2), 555-572. https://doi.org/10.5465/256693

Rogelberg, S. G., Luong, A., Sederburg, M. E., & Cristol, D. S. (2000). Employee attitude surveys: Examining the attitudes of noncompliant employees. *Journal of Applied Psychology*, *85*(2), 284-293. https://doi.org/10.1037/0021-9010.85.2.284

Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. *Journal of Applied Psychology*, *87*(1), 66-80. https://doi.org/10.1037/0021-9010.87.1.66

Sackett, P. R. (2007). Revisiting the origins of the typical-maximum performance distinction. *Human Performance*, *20*(3), 179-185. https://doi.org/10.1080/08959280701332968

Sackett, P. R., & DeVore, C. J. (2002). Counterproductive behaviors at work. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology: Personnel psychology* (pp. 145-164). Sage.

Sackett, P. R., Lievens, F., Van Iddekinge, C. H., & Kuncel, N. R. (2017). Individual differences and their measurement: A review of 100 years of research. *Journal of Applied Psychology*, *102*(3), 254-273. https://doi.org/10.1037/apl0000151

Sackett, P. R., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximum job performance. *Journal of Applied Psychology*, *73*(3), 482-486. https://doi.org/10.1037/0021-9010.73.3.482

Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology*, *78*(6), 966-974. https://doi.org/10.1037/0021-9010.78.6.966

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*(4), 703-742. https://doi.org/10.1111/j.1744-6570.1991.tb00696.x

van Aarde, N., Meiring, D., & Wiernik, B. M. (2017). The validity of the Big Five personality traits for job performance: Meta-analyses of South African studies. *International Journal of Selection and Assessment*, *25*(3), 223-239. https://doi.org/10.1111/ijsa.12175

van Iddekinge, C. H., & Ployhart, R. E. (2008). Developments in the criterion-related validation of selection procedures: A critical review and recommendations for practice. *Personnel Psychology*, *61*(4), 871-925. https://doi.org/10.1111/j.1744-6570.2008.00133.x

Vazire, S., & Carlson, E. N. (2010). Self-knowledge of personality: Do people know themselves? *Social and Personality Psychology Compass*, *4*(8), 605-620. https://doi.org/10.1111/j.1751-9004.2010.00280.x

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*(5), 1202-1216. https://doi.org/10.1037/a0013314

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321-327. https://doi.org/10.1007/BF02293557

Vernon, P. A., & Jang, K. L. (1993). Self-rated vs "actual" personality similarity in monozygotic and dizygotic twins and non-twin siblings. *Personality and Individual Differences*, *15*(2), 219–220. https://doi.org/10.1016/0191-8869(93)90028-2

Ward, M. K., Meade, A. W., Allred, C. M., Pappalardo, G., & Stoughton, J. W. (2017). Careless response and attrition as sources of bias in online survey assessments of personality traits and performance. *Computers in Human Behavior*, *76*, 417-430. https://doi.org/10.1016/j.chb.2017.06.032

Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, *78*(3), 546-558. https://doi.org/10.1037/0022-3514.78.3.546

Weekley, J. A., Ployhart, R. E., & Harold, C. M. (2004). Personality and situational judgment tests across applicant and incumbent settings: An examination of validity, measurement, and subgroup differences. *Human Performance*, *17*(4), 433-461. https://doi.org/10.1207/s15327043hup1704_5

Wheeler, A. R., Shanine, K. K., Leon, M. R., & Whitman, M. V. (2014). Student-recruited samples in organizational research: A review, analysis, and guidelines for future research. *Journal of Occupational and Organizational Psychology*, *87*(1), 1-26. https://doi.org/10.1111/joop.12042

Williams, L. J., & Anderson, S. E. (1991). Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management*, *17*(3), 601-617. https://doi.org/10.1177/014920639101700305

Withey, M. J., Gellatly, I. R., & Annett, M. (2005). The moderating effect of situation strength on the relationship between personality and provision of effort. *Journal of Applied Social Psychology*, *35*(8), 1587-1606. https://doi.org/10.1111/j.1559-1816.2005.tb02186.x

Author Biographies

Jason L. Huang (PhD 2012, Michigan State University) is an associate professor and associate director for Graduate Programs in the School of Human Resources and Labor Relations at Michigan State University. His research focuses on individual differences, employee adaptability, training and transfer, and response effort. Jason's work has been published in top-tier journals such as the *Journal of Applied Psychology*, *Journal of Management*, *Personnel Psychology*, *Psychological Bulletin,* and *Organizational Research Methods*. Jason is a former Associate Editor for the *Journal of Management* and currently serves on the editorial boards for the *Journal of Applied Psychology*, *Journal of Management*, *Journal of Vocational Behavior*, and *Journal of Business Psychology*.

Nathan A. Bowling (PhD 2005, Central Michigan University) is an associate professor of psychology at the University of Central Florida, where he serves as the associate director of the Industrial and Organizational Psychology PhD program. Nathan is a fellow of both the Society for Industrial and Organizational Psychology (SIOP) and the Association for Psychological Science (APS). He has published over 65 peer-reviewed journal articles, many of which have appeared in top-tier journals, including the *Journal of Applied Psychology*, the *Journal of Personality and Social Psychology*, and *Organizational Research Methods*. As a PI, Nathan has been awarded over $1M in external funding since 2021. His most recent funding was awarded by the Office of Naval Research.

Benjamin D. McLarty (PhD 2015, Louisiana State University) is an associate professor of Entrepreneurship at Rowan University in the Willian G. Rohrer College of Business School of Innovation & Entrepreneurship. His research interests intersect at the crossroads of micro topics and entrepreneurship and include individual differences, personality, job performance, new venture formation, and family business. Ben has published in the *Journal of Applied Psychology, Journal of Management, Entrepreneurship, Theory and Practice, Family*

Business Review, Journal of Small Business Management, Human Relations, Journal of Business Research, and the *Journal of Business & Psychology*, among others. He serves as an Associate Editor for the *Journal of Small Business Management* and on editorial boards for *Family Business Review*, the *Journal of Business & Psychology* and *Group and Organization Management*. He enjoys traveling, meeting interesting people, and being an active member of the academic community.

Donald H. Kluemper is a professor, area coordinator, and John B. Walker Endowed Professor of Business for the Area of Management in the Rawls College of Business at Texas Tech University. Don's research has been published in outlets such as the *Journal of Applied Psychology, Journal of Management, Personnel Psychology, Organizational Behavior and Human Decision Processes, Journal of Organizational Behavior, and Leadership Quarterly*, among others. In addition, some of Don's research has been featured in media outlets such as the *Wall Street Journal, Time Magazine, and National Public Radio*. Don was formerly the Director of Human Resources of a 250 + employee nonprofit organization.

Zhonghao Wang is an assistant professor of management at the Marilyn Davies College of Business at the University of Houston-Downtown. He received his PhD in human resources and labor relations from Michigan State University in 2023. Zhonghao studies job crafting, transfer of training, and response effort and has several publications in peer-reviewed journals.