

Goldsmiths Research Online

*Goldsmiths Research Online (GRO)
is the institutional research repository for
Goldsmiths, University of London*

Citation

Harris-Watson, Alexandra M.; Kung, Mei-Chua; Tocci, Michael C.; Boyce, Anthony S.; Weekly, Jeff A.; Guenole, Nigel and Carter, Nathan T.. 2022. The Interaction Between Conscientiousness and General Mental Ability: Support for a Compensatory Interaction in Task Performance. *Journal of Business and Psychology*, 37(4), pp. 855-871. ISSN 0889-3268 [Article]

Persistent URL

<https://research.gold.ac.uk/id/eprint/30638/>

Versions

The version presented here may differ from the published, performed or presented work. Please go to the persistent GRO record above for more information.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Goldsmiths, University of London via the following email address: gro@gold.ac.uk.

The item will be removed from the repository while any claim is being investigated. For more information, please contact the GRO team: gro@gold.ac.uk

**The Interaction between Conscientiousness and General Mental Ability:
Support for a Compensatory Mechanism in Explaining Task Performance**

Alexandra M. Harris-Watson¹, Mei-Chuan Kung², Michael C. Tocci³, Anthony S. Boyce⁴, Jeff
A. Weekley⁵, Nigel Guenole^{6,7}, Nathan T. Carter⁸

¹ Department of Communication Studies, Northwestern University

² Talent Management, PSI Services

³ Human Resources, Procter & Gamble

⁴ Talent Assessment, Amazon

⁵ Organizations, Strategy and International Management, University of Texas at Dallas

⁶ Talent and Transformation Business, IBM Global Business Services

⁷ Institute of Management, Goldsmiths, University of London

⁸ Department of Psychology, University of Georgia

Author Note

Alexandra M. Harris-Watson  <https://orcid.org/0000-0001-6136-8069>.

The work of Alexandra M. Harris-Watson was supported in part by the National Science Foundation Graduate Research Fellowship Program (DGE-1443117), and the work of Nathan T. Carter was supported in part by the National Science Foundation (SES1561070). Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the view of the National Science Foundation.

Correspondence concerning this article should be addressed to Alexandra M. Harris-Watson, email: alexandra.harris@northwestern.edu.

Abstract

We propose a *compensatory* interactive influence of conscientiousness and GMA in task performance such that conscientiousness is most beneficial to performance for low-GMA individuals. Drawing on trait by trait interaction theory and empirical evidence for a compensatory mechanism of conscientiousness for low GMA, we contrast our hypothesis with prior research on a conscientiousness-GMA interaction and argue that prior research considered a different interaction type. We argue that observing a compensatory interaction likely requires: (a) considering the appropriate interaction form, including a possible curvilinear conscientiousness-performance relationship; (b) measuring the full conscientiousness domain (as opposed to motivation proxies); (c) narrowing the criterion domain to reflect task performance; and (d) appropriate psychometric scoring of variables to increase power and avoid type 1 error. In four employee samples ($N_1 = 300$; $N_2 = 261$; $N_3 = 1,413$; $N_4 = 948$), we test a conscientiousness-GMA interaction in two employee samples. In three of four samples, results support a nuanced compensatory mechanism such that conscientiousness compensates for low to moderate GMA, and high conscientiousness may be detrimental to or unimportant for task performance in high-GMA individuals.

Keywords: Conscientiousness, general mental ability, interaction, performance

**The Interaction Between Conscientiousness and General Mental Ability:
Support for a Compensatory Interaction in Task Performance**

Many reviews and meta-analyses have affirmed the importance of personality for job performance (e.g., Judge et al., 2013; Ones, Dilchert, Viswesvaran, & Judge, 2007). Within the Five Factor Model (FFM) of personality, conscientiousness—the tendency toward organization, planning, persistence, sense of duty, and achievement motive—has been shown to have the strongest, positive relationship with job performance across occupations (Barrick & Mount, 1991; Judge et al., 2013).

Today, investigations of personality-performance relationships have moved beyond studying simple correlations to qualified relationships in the hopes of better understanding what causes relationships to strengthen or weaken (Barrick, 2005). In particular, *trait by trait interactions* have recently received attention for their potential to clarify personality-performance relationships (Shoss & Witt, 2013). For example, Witt et al. (2002) suggest that the influence of conscientiousness on job performance should be considered in conjunction with an individual's standing on agreeableness. They find that individuals are most effective when high in both conscientiousness and agreeableness and that low agreeableness can negate the benefits of high conscientiousness.

We argue that the same justification for studying trait by trait interactions easily extends to an interaction between conscientiousness and general mental ability (GMA) in explaining task performance. Given the dominance of these two antecedents in performance research, understanding their joint influence on performance should be of fundamental interest to researchers and practitioners. We build upon evidence of a compensatory relation between

conscientiousness and GMA (e.g., Moutafi et al., 2004) to propose conscientiousness is most beneficial to performance for low-GMA individuals.

Notably, the current study's primary contribution is its consideration of a *compensatory* interaction specifically. We are not the first to suggest an interactive influence of conscientiousness and GMA on performance generally (e.g., Sackett et al., 1998; Mount et al., 1999). Prior studies have not found strong support for a conscientiousness-GMA interaction. However, we argue past findings are attributable to important theoretical and methodological differences between prior research and the current study. After establishing the theoretical foundation for a compensatory interaction, we attempt to reconcile a compensatory interaction with past research by reviewing these differences.

Reviewing the theoretical context for prior studies reveals that most extended a *motivation-ability* interaction hypothesis to personality rather than investigating a *conscientiousness-GMA* interaction specifically. This distinction is critical because the motivation-ability interaction framework guided hypotheses inconsistent with the compensatory interaction proposed here. Thus, we hypothesize a different interaction form than was investigated by prior studies. Moreover, we argue that investigating a conscientiousness-GMA interaction requires: (a) considering that GMA moderates a curvilinear conscientiousness-performance relation; (b) defining conscientiousness as more than motivation; (c) distinguishing between performance dimensions; and (d) using appropriate psychometric scoring.

Modern Support for a Conscientiousness-GMA Interaction

A variety of moderators have been considered to clarify the personality-performance relationship. One category of qualifiers that has been posited for decades is personality traits as moderators of other personality-performance relationships, or trait by trait interactions. The

theoretical rationale for trait by trait interactions is rooted in the intuitive idea that “one word or one trait provides a limited representation of a person” and that a “configuration” of traits is more “relevant for understanding and predicting workplace variables” than is treating traits as independent, competing predictors (Shoss & Witt, 2013, p. 392). Moreover, researchers note two empirical justifications for investigating trait by trait interactions: (1) secondary loadings of personality items onto other traits and (2) growing evidence that interactive influences of personality traits show incremental validity in predicting a variety of employee outcomes (Penney et al., 2011). Despite limited empirical research, several studies have found support for such trait by trait interactions in performance (Judge & Erez, 2007; Witt, 2002; Witt et al., 2002).

Much of the theoretical rationale for considering the simultaneous influence of personality traits on job performance extends to the simultaneous influence of personality and GMA. Some of the most popular frameworks for understanding the personality-GMA interface today have advocated for a complex, conditional relationship (e.g., Ackerman, 1996; Chamorro-Premuzic & Furnham, 2004) much like the theoretical basis for trait by trait interactions. For example, Ackerman builds on earlier frameworks of intelligence in his intelligence-as-process, personality, and interests, and intelligence-as-knowledge (PPIK) theory to propose that personality may affect the intellect development. That is, Ackerman suggests personality is generally unrelated to innate GMA (intelligence-as-process) but is related to acquired knowledge (intelligence-as-knowledge) because individuals with particular traits will be influenced to invest in knowledge acquisition.

Chamorro-Premuzic and Furnham take an even broader perspective on the intelligence-personality interface by highlighting distinctions between intellectual ability (i.e., fluid intelligence or what we refer to as GMA), intellectual knowledge (i.e., crystallized intelligence),

IQ test performance, and self-assessed intelligence. They argue personality is likely related to each intelligence type such that high or low GMA may differentially encourage development of certain personality traits. Finally, in his chapter reviewing the relationship between personality and intelligence, DeYoung (2011) presents and refutes three common arguments for separating the two: (a) the dichotomy between “cognitive” and “non-cognitive” traits; (b) differential measurement; and (c) their relationship to maximal and typical performance. DeYoung asserts that personality traits include cognitive components, that measurement should not be confounded with constructs, and that intelligence can impact typical performance just as personality might impact maximal performance. Ultimately, he concludes that “viewing intelligence as a personality trait is a viable, if relatively uncommon, conceptual strategy” (p. 720). Regardless of the framework used to understand the intelligence-personality interface, it seems clear these fundamental individual differences are interrelated in complicated ways.

Conscientiousness, in particular, has been posited to be related to GMA. Chamorro-Premuzic and Furnham (2004) suggest that conscientiousness develops in part to compensate for low GMA. This hypothesis stems in part from empirical findings regarding the conscientiousness-GMA relationship. Although early research suggested no relationship (Ackerman & Heggestad, 1997), more recent studies have shown evidence for a small, negative relationship between conscientiousness and GMA (e.g., Carretta & Ree, 2017; Chamorro-Premuzic & Furnham, 2008; Furnham et al., 2007; Rammstedt et al., 2016). This relationship seems inconsistent with our understanding of how conscientiousness and GMA relate to performance in that both correlate positively with performance yet negatively with each other (Moutafi et al., 2004). To account for this inconsistency, researchers have proposed a compensation mechanism by which individuals with low GMA develop conscientiousness.

To understand how conscientiousness may act as a compensation mechanism, it is useful to review what we mean by “GMA” and “conscientiousness” and how each relates to performance. GMA refers to one’s mental capacities and is frequently conceptualized as a basic ability for abstract reasoning, problem solving, or adaptability (Ones, Dilchert, & Viswesvaran, 2012). Although researchers distinguish between specific mental abilities as well as fluid versus crystallized intelligence, these constructs are highly correlated. As explained above, specific abilities and crystallized intelligence are thought to develop in part due to the importance of GMA in various disciplines. Just as GMA fosters more specific ability types, GMA is thought to be strongly related to job performance because it fosters the development of job knowledge (Hunter, 1986). Thus, individuals with high GMA have more declarative and procedural knowledge, which mediates the GMA-performance relation (Ones et al., 2012).

Because GMA is thought to primarily affect performance via acquired job knowledge, compensating for low-GMA requires (a) alternative strategies for acquiring job knowledge or (b) strategies for performing without high levels of job knowledge. Conscientiousness may provide both strategy types. DeYoung (2015) asserts personality traits describe “stable patterns of emotion, motivation, cognition, and behavior” that facilitate goal fulfillment. Specifically, conscientiousness concerns “governing behavior across long time spans or according to the relatively arbitrary explicit rules that are a function of the complexity of human cultures” (p. 13). DeYoung suggests the lower-order conscientiousness aspect *industriousness* concerns the prioritization of goals, and *orderliness* concerns following rules. Industriousness may allow individuals to see the benefit of job knowledge and push themselves to acquire it despite low ability. Indeed, numerous studies have linked conscientiousness to training performance (Barrick & Mount, 2001), suggesting conscientiousness may play a role in developing job knowledge. On

the other hand, orderliness may allow individuals to strictly adhere to policies and procedures despite incomplete understanding of why those actions are appropriate. Low-GMA individuals may develop conscientious behavior to achieve outcomes similar to their higher-GMA peers (Moutafi et al., 2004, Chamorro-Premuzic & Furnham, 2008).

Some researchers have suggested the negative conscientiousness-GMA relationship merely reflects “compensatory selection” into common research samples (e.g., mid- to upper-level managers, employees selected for developmental opportunities, or other occupationally elite groups). Murray et al. (2014) assert that obtaining membership in such populations “can be done through combinations of ability and hard work... but hard work can compensate for relatively low ability and high ability can compensate for relatively less hard work.” (p. 18). They argue this phenomenon causes a negative conscientiousness-GMA relation to appear in samples of high-achieving participants that does not reflect the overall population.

Regardless of their relative accuracy, both explanations for the conscientiousness-GMA relationship suggest an interactive, compensatory¹ influence of conscientiousness and GMA on occupational achievement. In both explanations, GMA influences conscientiousness expression such that low-GMA individuals are more reliant on conscientiousness for goal attainment relative to high-GMA individuals. Thus, task performance, a desirable occupational outcome that corresponds to gainful employment and reward, depends on not only conscientiousness and GMA independently but also their compensatory combination. Just as researchers have called for

¹ The term *compensatory* is sometimes used to refer to an additive model in which effects are still independent (e.g., GMA and motivation are independently beneficial to performance; see Van Iddekinge et al., 2017). We use *compensatory* here to refer to an interaction in which conscientiousness is most or generally beneficial to performance among those with low-GMA but less so among those with high-GMA. That is, the conscientiousness-performance relationship is *dependent* on GMA level.

investigating the joint influence of personality traits on job performance, it seems imperative to consider the joint influence of conscientiousness and GMA on performance.

Reconciling Prior Findings and a Compensatory Hypothesis

Although very few studies have considered a compensatory conscientiousness-GMA interaction like that discussed above, the notion of a conscientiousness-GMA interaction broadly is not novel. Readers may recall two seminal studies that investigated such an interaction: Sackett et al. (1998) in the *Journal of Applied Psychology* and Mount et al. (1999) in the *Journal of Management*. Both found no support for an interaction of conscientiousness-related traits and GMA for performance.

Still, some work tentatively supports a conscientiousness-GMA interaction. In addition to earlier studies that found support for a conscientiousness-GMA interaction for career success (O'Reilly & Chatman, 1994) and supervisor ratings of performance (Wright et al., 1995), recent studies revisiting the topic have found support. Perry et al. (2010) found support for an interactive influence of the achievement facet of conscientiousness and GMA on task performance, and de Haro et al. (2013) found evidence of a GMA-conscientiousness interaction for salary and career satisfaction. Mixed support also exists for a conscientiousness-GMA interaction in academic performance (e.g., Beaujean et al., 2011; Bergold & Steinmayr, 2018; Di Domenieco & Fournier, 2015; Zhang and Ziegler, 2015; Ziegler et al., 2009). Tentative support for such interactions in academics is particularly relevant given its similarity to job performance in relation to FFM traits and GMA (Kuncel et al., 2004).

These varied findings suggest research on a conscientiousness-GMA interaction is far from conclusive. Next, we review four explanations for conflicting findings. First and foremost, we assert that prior research has been informed by a different theoretical tradition than that of the

current study: that performance is a multiplicative function of motivation and ability. This difference in theory suggests different forms of the expected interaction. Moreover, we argue for the importance of considering a *curvilinear* conscientiousness-performance relationship such that there may be a quadratic compensatory interaction. Finally, we summarize three methodological concerns that may account for inconsistent evidence of an interaction: differences in the measured conscientiousness domains; the breadth of the criterion domain; and measurement concerns related to assessment appropriateness, particularly in personality measures.

Interaction Form

Most past research on the interactive influence of GMA and conscientiousness on performance originates from the expectancy theory hypothesis that performance is a function of ability and *motivation* (Maier, 1955; Vroom, 1995). This multiplicative function suggests that ability affects performance only in so far as employees are motivated to apply that ability. Perhaps related to the establishment of the FFM and the resurgence of personality research, in the late 1980s and 1990s researchers began utilizing motivation-related personality traits—namely conscientiousness and related traits—in investigations of the motivation-by-ability function (Hollenbeck et al., 1988; O'Reilly & Chatman, 1994; Mount et al., 1999; Sackett et al., 1998; Wright et al., 1995). More recent work on a conscientiousness-GMA interaction has also invoked motivation-based justifications (e.g., Beaujean et al., 2011; de Haro et al., 2013; Di Domenieco & Fournier, 2015).

Importantly, the motivation-ability interaction theory and the compensatory mechanism described above explicate different—even opposing—forms of a conscientiousness-GMA interaction. As explained by Van Iddekinge et al. (2017) in their meta-analysis of the topic, the motivation-ability interaction hypothesizes a non-compensatory multiplicative model such that

“performance is predicted to be low whenever ability *or* motivation is low” (p. 2). That is, the motivation-performance relationship will be strongest when an individual possesses moderate to high GMA. In contrast, a compensatory interaction suggests conscientiousness will most benefit the performance of low-GMA persons.

Because the forms of the interactions grounded in these two perspectives are opposed, evidence against the motivation-ability interaction does not necessarily run counter to a compensatory mechanism. As Sackett et al. (1998) noted, some interactions identified in investigations of a motivation-ability interaction have shown inconsistent forms. Further, more recent support for the conscientiousness-GMA interaction in performance shows some evidence consistent with a compensatory role of conscientiousness (e.g., Beaujean et al., 2011; de Haro et al., 2013; Postlewhaite et al., 2009). Thus, it seems even results from studies supporting a motivation-ability interaction are perhaps more appropriately understood in terms of a compensatory mechanism. As our hypotheses will suggest, we endorse the compensatory mechanism based on theoretical grounds and empirical evidence for this functional form for other task-related criteria (e.g., academic and safety performance).

A second potential limitation of earlier work on the form of conscientiousness-GMA interactions is that, to our knowledge, no prior work has considered whether GMA moderates a curvilinear conscientiousness-performance relationship. Recent findings point to such curvilinear relationships such that high levels of conscientiousness are either detrimental or no longer beneficial for performance (Yuan, 2018; Carter et al., 2014; Le et al., 2011) and other outcomes (Carter et al., 2016). It is possible a curvilinear conscientiousness-performance relationship is most pronounced among high-GMA individuals. Lower-GMA individuals may show linear conscientiousness-performance relationships (i.e., benefit from even very high levels of

conscientiousness), whereas the performance of high-GMA individuals does not benefit from or is hindered by high conscientiousness. Given emerging evidence that conscientiousness is curvilinearly related to performance, a compensatory conscientiousness-GMA interaction may manifest such that there are differential quadratic conscientiousness-performance relationships dependent on GMA. Evidence supporting such a quadratic interaction would suggest a role of quadratic relationships in obscuring findings so far.

FFM-Based Conscientiousness Measures

Another factor that may have contributed to inconsistent findings regarding a conscientiousness-GMA interaction is different operationalizations of conscientiousness. As alluded to in our discussion of the motivation-ability interaction tradition, in many cases researchers were more interested in conscientiousness as a proxy for motivation than they were in conscientiousness itself. For instance, Sackett et al. (1998) focused on dependability and need-for-achievement, which they distinguished from the broader conscientiousness trait. Indeed, Mount et al. (1999) noted, “none of the [prior] studies used construct valid measures of conscientiousness. Instead, components of conscientiousness were examined such as need for achievement” (p. 711). Intriguingly, in their recent meta-analysis of the motivation-ability interaction theory of performance, Van Iddekinge et al. (2017) state that most research on the topic used trait-based measures of motivation, including conscientiousness. However, like Mount and colleagues, Van Iddekinge et al. (2017) note that conscientiousness is a “broad, multifaceted construct” (p. 11) and chose not to include studies that operationalized motivation as conscientiousness in their meta-analysis.

Whereas earlier work on this topic involved mostly motivation-related components of conscientiousness, here we are interested in the full domain. As discussed above, both the lower-

order conscientiousness aspects might aid individuals in compensating for low GMA. Notably, recent studies that conceptualized conscientiousness as the broader trait found support for the interaction in both career success (de Haro et al., 2013) and academic performance (Beaujean et al., 2011; Bergold & Steinmayr, 2018; Di Domenieco & Fournier, 2015). Moreover, the forms of some of these interactions have been consistent with a compensatory interaction (de Haro et al., 2013; Beaujean et al., 2011). Thus, prior findings may have differed had they used broader measures of conscientiousness.

Breadth of the Criterion Domain

Of studies that have considered the interactive influence of conscientiousness and GMA on job performance, the breadth of the performance construct domain has varied widely. Most studies have used supervisor ratings that combine task and extra-role performance. Whereas task performance refers to effectiveness completing core job requirements, contextual performance refers to additional behaviors that impact organizational performance (Borman & Motowidlo, 1993). For instance, the performance composites used in Mount et al. (1999) included both task performance items (e.g., “know-how” and “job knowledge”) and extra-role performance (e.g., “organizational commitment” and “motivates others”).

Distinguishing between performance dimensions is now considered critical to identifying personality-performance relationships (Motowidlo & Van Scotter, 1994). More recent studies utilizing narrower performance criteria, such as task performance (e.g., Perry et al., 2010) or academic performance, have shown some support for a conscientiousness-GMA interaction. Still, few studies have used task performance as the exclusive criterion. Insufficient distinction between performance dimensions may have impeded the identification of an interaction. Here, we concentrate on task performance given that theoretical and empirical evidence suggest the

compensatory interactive influence of conscientiousness and GMA is most tenable for task-related performance behavior.

Assessment Appropriateness

Finally, two key statistical issues limit the conclusiveness of prior findings. First, all prior studies on the interactive influence of conscientiousness and GMA on job performance have used raw composite scores as their predictor and criterion variables. Morse et al. (2012) showed that use of raw scores can lead to inflated type 1 error rates and reduced power for tests of moderation. Type 1 error increases greatly when assessments are “inappropriate” for those assessed (i.e., when items are more extreme than the majority of persons being measured). Recent research suggests that most personality measures fit this definition of assessment inappropriateness (Carter et al., 2017) and utilizing the IRT-derived scores can help to alleviate type 1 error and low power. Because all prior investigations of personality-GMA interactions utilized raw scores, the inconsistency of past findings may reflect type 1 error and/or reduced power. Thus, we utilize IRT-scoring or ensure assessment appropriateness for all measures in the current study to increase accuracy in detection of interactions (Embretson, 1996; Morse et al., 2012; Carter et al. 2017).

The Current Study

Despite less supportive earlier work, some recent evidence supports the interactive influence of conscientiousness and GMA on job performance. Evidence from other fields also suggests personality and GMA impact performance interactively. Differences in theoretical and methodological approaches provide some insight into the discrepancy between these more recent results and earlier findings. Thus, we believe the interactive influence of conscientiousness and GMA on task performance requires further investigation. Drawing on theoretical and empirical

support for a compensatory conscientiousness-GMA interaction, we expect conscientiousness will be most beneficial for performance among low-GMA individuals and relatively less beneficial for performance among high-GMA individuals, or stated formally:

Hypothesis: GMA moderates the conscientiousness-task performance relationship such that conscientiousness is more beneficial to lower-GMA than higher-GMA individuals.

In addition to updating the hypothesized form of the conscientiousness-GMA interaction to be consistent with a compensatory mechanism rather than motivation-ability interaction theory, the current study allows for the possibility that a compensatory interaction might be quadratic.

Because our hypothesis for a conscientiousness-GMA interaction draws on the full conscientiousness domain (as opposed to its motivational components), we strove to measure the trait's full breadth. In the current study, we draw on the 6-2-1 structure (Judge et al., 2013) as the basis for the conscientiousness measures. The 6-2-1 structure assigns two aspects (DeYoung, et al., 2007) and six facets (McCrae & Costa, 1992) to each FFM trait, providing a useful guide for evaluating content. This structure suggests conscientiousness is composed of the aspects industriousness (containing self-discipline, competence, and achievement facets) and orderliness (containing order, dutifulness, and deliberation facets). For each conscientiousness measure used here, we aimed to balance content between these aspects and—where possible—across facets.

Finally, consistent with the methodological limitations of prior studies outlined above, in the current study, we narrow the criterion exclusively to task performance and score assessments using appropriate IRT analyses. Next, we test our hypothesis regarding a compensatory, interactive influence of GMA and conscientiousness on task performance using four samples ranging from approximately 300 to 1,500 employees from across different organizations and occupational contexts.

Method

In the current study, we test our hypothesis in four samples from applied datasets at large national or international organizations. Because each organization used a proprietary measure of personality that varied slightly from the FFM traditionally used in research, the first and last author reviewed the content of all items and selected only those that fit within the 6-2-1 framework of conscientiousness as described above. We then conducted IRT analyses to evaluate dimensionality and item properties (e.g., discrimination parameters reflecting the degree to which each item loads onto the underlying trait and difficulty parameters reflecting the level of the latent trait needed to endorse an item). We eliminated items with the goal of retaining those that reflected a range of conscientiousness facets, item extremity, and sufficient discrimination parameters (e.g., no less than .20 and no more than 4.0) and reviewed model-data fit of the final measure. Similarly, we reviewed GMA measures for consistency with fluid intelligence and performance measures for relevance to task performance. For all measures, IRT models were selected according to the measure's response format and the nature of the construct consistent with best practices in IRT analysis (Embretson & Reise, 2000; LaPalme et al., 2018; Stark et al., 2006).

Samples

Sample 1

Personality data and supervisor performance ratings for Sample 1 were collected from 3,963 employees at a large, multinational consumer goods organization (47.3% female; 41.5% United States) as part of a large-scale validation study. Represented positions were entry to mid-level management in a range of department types (e.g., sales, marketing, product supply, administrative, etc.). GMA data were collected as part of a previous validation effort for an in-

house GMA assessment. Personality and performance scores were derived using the full data. However, only 300 employees with available personality data also took part in the GMA validation effort. Thus, GMA data was not available for all participants and those with GMA data were subset for use in regression analyses ($N = 300$; 49% female; 38% United States).

Sample 2

Sample 2 data were collected from 5,570 applicants to entry-level, manufacturing positions at a large glass and fiberglass company in the United States (27.6% female). Personality and GMA scores were derived using the full sample. Performance data were collected from hired applicants, whom were subset for regression analyses ($N = 261$; 21.8% female).

Sample 3

Sample 3 data were collected from 3,104 employees in assistant or store manager positions at both large and specialty retail stores. Data included responses to a personality assessment, a GMA assessment and supervisor ratings of performance. Predictor and criterion IRT scores were calculated using all data available for each construct. However, only 1,413 of these participants had scores available for all three constructs.

Sample 4

Sample 4 data were collected from 948 employees in entry-level positions at food service and retail companies (39.7% female). Personality and GMA assessments as well as supervisor ratings were available for the full sample.

Measures and Scoring

Sample 1

Personality. In Sample 1, participants completed an in-house computer adaptive test (CAT) designed to assess a range of trait composites (e.g., leadership, integrity, customer service orientation, etc.). All items were self-report (e.g., “Pay attention to details”). Participants indicated agreement to statements on a scale of 1 (*strongly disagree*) to 6 (*strongly agree*). In CATs, participants complete a subset of items selected by a computer algorithm to target the individual’s trait levels. Consequently, participants complete different sets of items, and substantial missingness is inherent to CAT data. To accommodate missingness, we dichotomized item responses such that options 1 to 3 were coded ‘0,’ and 4 to 6 were coded ‘1.’ After reviewing item content and item parameters and removing items accordingly, the final conscientiousness measure included 17 items and was scored using the ideal point model for dichotomous data (Maydeu-Olivares et al., 2006).² Prior research has shown that ideal point responding more accurately describes responses to personality constructs using agreement scales (LaPalme et al., 2018). Global fit statistics suggested good model-data fit for the ideal point model the conscientiousness measure responses (RMSEA = .006, TLI = .998, CFI = .998; Hu & Bentler, 1999), and trait estimates³ were calculated using the fitted model.

General Mental Ability. In Sample 1, GMA was measured using composite scores from a 40-item internally IRT-validated and in-person GMA assessment including numeric, logic, and figural subscales.

Task Performance. In Sample 1, a 5-item task performance measure was selected from available supervisor ratings using the GRM. Supervisors rated employees from 1 = *weak* to 5 =

² Recent research has shown that ideal point models generally fit personality responding using Likert-type (agreement) scales better than do dominance models (Stark et al., 2006; LaPalme et al., 2018) and are important for detecting curvilinear relationships (Carter et al., 2014; Carter et al., 2016). Consequently, we use ideal point IRT models and note fit relative to dominance-based IRT models where appropriate. We recommend Roberts et al. (1999) for a thorough review of the theoretical distinction between ideal point and dominance models.

³ Theta values can be interpreted much like z-scores (e.g., a standardized theta of 1 is close to 1 standard deviation).

exceptional for each item. Global fit statistics could not be calculated due to the low number of items. Thus, we estimated a confirmatory factor analysis model for ordinal data as an approximation to the GRM, which showed acceptable model-data fit (TLI = .974; CFI = .987; RMSEA is not appropriate due to low *df*; Kenny et al., 2015). Performance scores were estimated using the fitted exploratory GRM.

Sample 2

Personality. Participants completed a 128-item personality measure with nine subscales, of which two approximately reflected the two aspects of conscientiousness. The first and last authors reviewed item content for these subscales and removed items that were inconsistent with these aspects or showed inappropriate item parameters. The final measure included 17 items (e.g., “If something needs to be done, I do it immediately” and “I worry about small details”). Items utilized a sliding scale of agreement from 0 to 100. To conduct IRT analyses, sliding scale responses were divided evenly into six categories. To ensure an ideal point model was still appropriate for the recoded data, we compared the fit of the generalized graded unfolding model (GGUM; Roberts, Donoghue, & Laughlin, 2000) to the dominance-based graded response model (GRM; Samejima, 1997). RMSEA confidence intervals for GGUM and GRM overlapped (GGUM RMSEA = .042, 95% CI [.039, .047]; GRM RMSEA = .037, 95% CI [.034, .041]), but relative fit indices suggested better fit of the GGUM (TLI = .997, CFI = .998) relative to the GRM (TLI = .936, CFI = .952). Thus, the GGUM was used.

General Mental Ability. In Sample 2, participants completed a 25-item GMA assessment, including numerical and figural reasoning subscales. Numerical items involved identifying which number in a series does not belong, and figural reasoning items involved reviewing three figures to identify the next figure in the sequence. After using IRT analyses to

evaluate the GMA measure's items, 19 items were retained reflecting both numerical and figural reasoning items. The three-parameter logistic (3PL) model is often appropriate for dichotomously scored multiple choice assessments (Waller, 1989) because it includes a c -parameter that accounts for guessing (Embretson & Reise, 2000). Alternatively, the two-parameter logistic model (2PL) may be appropriate if c is near zero. The 3PL showed evidence of better fit relative to the 2PL (AIC = 110,061; BIC = 110,312 for the 2PL; AIC = 109,883, BIC = 110,260 for the 3PL), and a RMSEA of .048 suggested good absolute fit for the 3PL. Relative fit statistics (CFI, TLI) were not appropriate because null RMSEA was less than .158 (Kenny et al., 2015). Trait estimates were calculated using the 3PL.

Task Performance. In Sample 2, a 7-item task performance measure was selected from supervisor ratings using the GRM. Supervisors rated the frequency of employees' task-related behavior on a scale of 1 = *never* to 7 = *always*. Due to the low number of items in the measure, global fit statistics could not be calculated. Thus, we conducted confirmatory factor analysis for ordinal data, which showed acceptable model-data fit (TLI = .939; CFI = .960; RMSEA is not an appropriate due to low df ; Kenny et al., 2015). Scores were estimated from the GRM.

Sample 3

Personality. In Sample 3, participants completed a 225-item assessment representing a range of personality traits. We selected 13 items that reflected the conscientiousness aspects orderliness (e.g., "I keep my officer or work space clean, neat, and orderly") and industriousness (e.g., "I do my best work when my tasks are challenging and difficult"). Items included variable response option types (e.g., agreement; yes/no; frequency; and situational judgement statements). Items that invoke frequency and rankings as response options conform to dominance response processes because they require test-takers to indicate whether they have more or less of a trait

directly (e.g., “daily” always means more frequently than does “weekly”). Thus, we expected a dominance model to fit better than an ideal point model. We compared the fit of the GRM and the GGUM for the final conscientiousness scale in Sample 1. Relative fit statistics (TLI and CFI) were not appropriate indicators of model-data fit for the GGUM because null RMSEA was less than .158 (Kenny et al., 2015). However, absolute fit statistics suggested better fit of the GRM (RMSEA = .048, 95% CI [.043, .053]) relative to the GGUM (RMSEA = .063, 95% CI [.057, .070]). Thus, conscientiousness theta estimates were calculated from the unidimensional GRM.

General Mental Ability. In Sample 3, participants completed a 90-item GMA test including a variety of item types. From this pool, 68 items were selected consistent with the content of the Weschler Adult Intelligence Verbal IQ subtest (Enns & Reddon, 1998; Wechsler, 1955), including verbal comprehension and working memory content. GMA was scored using the 2PL IRT model. Although the 3PL is also appropriate for dichotomously scored data, the data did not show evidence of meaningful guessing, and model-data fit statistics suggested better fit of the 2PL relative to the 3PL (BIC = 174,060 for the 2PL; BIC = 174,579 for the 3PL). We found excellent model-data fit for the GMA measure in Sample 1 (RMSEA = .044, TLI = .960, CFI = .962). Thus, GMA estimates were derived from the 2PL.

Task Performance. In Sample 3, an 11-item measure of task performance was derived from a larger bank of supervisor ratings using the GRM. Global fit statistics suggested excellent model-data fit of the GRM for the task performance measure (RMSEA = .032; TLI = .973; CFI = .991).

Sample 4

Personality. In Sample 4, participants completed a 322-item assessment including a variety of item types (e.g., agreement, situational judgement) and range of personality traits.

Items were selected using the same process as Sample 3. However, only self-report agreement items (scale of 1 = *strongly disagree* to 5 = *strongly agree*) were included in the final measure. The final measure included 11 items (e.g., “I am often late for scheduled appointments” and “I get discouraged easily”). Because Sample 2 utilized traditional agreement statements, comparison of ideal point and dominance-based scoring was not necessary (LaPalme et al., 2018), and the final measure was scored using the GGUM. Global fit statistics suggested good model-data fit (RMSEA = .042., TLI = .980, CFI = .985).

General Mental Ability. In Sample 4, participants completed 55 logic and numerical reasoning items; 48 items were selected based on item parameters. Like Sample 3, fit statistics again suggested better fit of the 2PL relative to the 3PL (BIC = 36843.03 for the 2PL; BIC = 36956.62 for the 3PL). Global fit statistics suggested excellent model-data fit for the GMA measure in Sample 2 (RMSEA = .025, TLI = .982, CFI = .983), and GMA estimates were calculated using the 2PL.

Task Performance. In Sample 4, participants were rated by supervisors on 36 items reflecting various performance dimensions. Consistent with a prior study utilizing the same performance items (CITATION REDACTED FOR BLIND REVIEW), we selected five items that best reflected task performance. Due to the low number of items and resulting low *df*, global fit statistics were not available. Consequently, we estimated a confirmatory factor analysis model for ordinal data as a proxy for the GRM, which showed acceptable model-data fit (TLI = .992; CFI = .996; RMSEA is not appropriate due to low *df*; Kenny et al., 2015). High internal consistency was also evident ($\alpha = .90$).

Data Analysis

Moderated regression models were tested in all samples. Due to type 1 and type 2 error that can result from correlated variables in interaction terms (Cortina, 1993), a quadratic GMA term was included. All variables were standardized prior to calculating quadratic and interaction terms. One-tailed p-values are reported for regression analyses consistent with directional hypotheses (see Cho & Abe, 2013 for a discussion on the misuse of two-tailed tests).

Notably, prior research has found that moderation tends to show inherently small effect sizes using traditional indices (e.g., median ΔR^2 and f^2 of .002) even when those effects are significant and have meaningful impact (Aguinis et al., 2005). Thus, in the current study, we report traditional effect size indices as well as several alternatives to aide interpretation of meaningful impact. We utilize the recently proposed R^2_{mo} to report the variability in the predictor-criterion relationship that is accounted for by a moderating term (Liu & Yuan, 2020). Liu and Yuan assert that R^2_{mo} is a more conceptually appropriate representation of moderator effect size relative to the more traditional index ΔR^2 , which represents the proportion of *all* variance in the criterion accounted for only by the multiplicative interaction term and contributes to the misleading perception of small moderation effective sizes. Additionally, we highlight inflection points in quadratic relationships and conclude with a simulation that compares the predictive utility of alternative models with the aim of illustrating the real-world impact of the hypothesized effects.

Results

Table 1 displays correlations and descriptive statistics for all samples. Table 2 shows results for regressions analyses in all samples. In Sample 1, results showed a significant quadratic interaction, $b = -0.11$, $p = .042$. In Sample 2, results did not show evidence of a significant conscientiousness-GMA interaction. Results for Sample 3 showed a significant

quadratic interaction, $b = -0.04$, $p = .014$. Results for Sample 4 also showed a quadratic interaction, $b = -0.05$, $p = 0.026$ as well as a significant linear interaction, $b = 0.06$, $p = .037$.

As noted above, prior research has found that traditional indices of interaction effect sizes yield very small effect sizes even for interactions that have meaningful impact (Aguinis et al., 2005) and, further, that these indices may be conceptually inappropriate for representing moderation effect size (Liu & Yuan, 2020). As calculated using ΔR^2 , the current results show effect sizes for the conscientiousness-GMA interaction terms (including both linear and quadratic conscientiousness) that are slightly larger than median interaction effect sizes found in the literature for interaction effects ($\Delta R^2 = .016$ in Sample 1; $\Delta R^2 = .005$ in Sample 3; $\Delta R^2 = .006$ in Sample 4). That is, across the three samples that demonstrated evidence of a significant interaction between quadratic conscientiousness and GMA, the interaction terms accounted for between 0.05% and 1.6% of the variance in task performance. However, Liu and Yuan argue that ΔR^2 is conceptually inappropriate and thus misleading regarding true moderation effect size. Instead, they propose R^2_{mo} which represents the proportion of variability in just the relationship between the predictor and criterion that is accounted for by the moderator. In the current study, $R^2_{mo} = 30.1\%$ in Sample 1; 66.3% in Sample 3; and 61.9% in Sample 4. Thus, across the three samples that demonstrated a significant conscientiousness-GMA interaction, GMA accounted for between 30.1% and 66.3% of the varying relationship between conscientiousness and performance.

Figure 1 illustrates the moderating effect of GMA on the conscientiousness-performance relationship and indicates inflection points for Samples 1, 3, and 4. In all three samples, among high-GMA employees, there was a quadratic conscientiousness-performance relationship such that conscientiousness was positively related to task performance until an inflection point. After

the inflection point, conscientiousness became detrimental to performance in Sample 1 (i.e., detrimental to top 27.73% of population in conscientiousness) and was no longer beneficial to performance in Samples 3 and 4 (i.e., not beneficial for top 12.76% and 10.24% of population in conscientiousness for Samples 3 and 4, respectively). Among average-GMA employees, Sample 1 showed a quadratic conscientiousness-performance relationship such that, after an inflection point, conscientiousness was no longer beneficial to performance (i.e., not beneficial for top 21.01% of population in conscientiousness). However, in Samples 3 and 4, conscientiousness showed an essentially linear relationship with performance among average-GMA employees.

Finally, among low-GMA employees, conscientiousness showed a generally positive relationship with performance in Sample 1. In Samples 3 and 4, conscientiousness also showed a positive relationship with performance among low-GMA employees but only *above* an inflection point (only top 57.34% and 60.67% of population in conscientiousness benefitted). Thus, in all three samples, the form of the conscientiousness-GMA interaction was consistent with a compensatory quadratic interaction such that conscientiousness was more beneficial to low- to average-GMA employees relative to high-GMA employees.

Discussion

Conscientiousness and GMA represent two of the most fundamental individual differences and researched predictors of performance. Understanding how they interact is critical to advancing basic psychological research as well as a thorough explanation of job performance and its predictors. The current results support a compensatory quadratic conscientiousness-GMA interaction in task performance such that for high-GMA employees there exists a point at which increased levels of conscientiousness are no longer beneficial to performance. Three of four samples (75%) showed support for a significant quadratic interaction between GMA and

conscientiousness, roughly the level of power expected from these samples. Indeed, an admittedly rough “mini meta-analysis” of the quadratic interaction effects suggested an average weighted coefficient of -0.04 , $p=.008$, 95% CI $[-0.09; -0.01]$, with non-significant heterogeneity.⁴ Across the three samples, GMA accounted for between 30.1% and 66.3% of the varying relationship between conscientiousness and performance according to the new effect size R^2_{mo} proposed by Liu & Yuan (2020). Thus, results emphasize both that GMA moderates the relationship between conscientiousness and performance and the importance of considering curvilinear effects to identify a compensatory interaction between conscientiousness and GMA. Not considering curvilinearity may have obscured this finding in prior research.

Notably, Sample 2 did not show support for the hypothesized interaction. Two possible explanations for the differences in results between Sample 2 and other samples are range restriction and lack of power. Personality and GMA batteries in Sample 2 were used to select employees and included data on all applicants, but only performance data for hired employees were available. Thus, Sample 2 data included direct range restriction. Admittedly, although we had evidence of direct range restriction in Sample 2, we cannot know for sure whether range restriction might have affected findings in the other samples. Additionally, Sample 2 was the smallest of all four samples ($N = 261$). To evaluate the potential impact of range restriction and power on our findings, we conducted post-hoc analyses using multivariate imputation and simulation analyses (see Appendix A for details). Overall, results suggest that range restriction may have affected results but was not enough to fully account for the lack of significant effects

⁴ Additionally, using the Sample 2 estimate corrected for range restriction (see Appendix A, Table A1) for the meta-analysis shows similar results with an average weighted coefficient of -0.04 , 95% CI $[-0.06; -0.02]$, $p=.001$, with non-significant heterogeneity.

here. In contrast, results suggest that power may be of particular concern in Sample 2 ($N = 261$) relative to other samples.

In all three samples that exhibited evidence of a quadratic interaction, high-GMA individuals showed a curvilinear conscientiousness-performance relationship such that the benefit of conscientiousness either tapered off or became detrimental at very high levels. Similarly, in these samples the point at which conscientiousness became detrimental for performance (i.e., inflection point) increased at lower GMA levels. In Samples 3 and 4, inflection points for low GMA were slightly below mean conscientiousness levels. However, this inflection was an indicator that one must have at least that level of conscientiousness to see its benefits for performance (as opposed to no more than that level) and therefore was not inconsistent with our hypothesis. Thus, the form of the interactions in these samples was consistent with a nuanced compensatory conscientiousness-GMA interaction such that conscientiousness is more beneficial to average- relative to high-GMA individuals. Results also suggest conscientiousness can help low-GMA individuals compensate, though they are still unlikely to outperform peers with a combination of average to high GMA and moderately high conscientiousness.

Prior research has noted the difficulty in translating the meaningful impact of interaction effects due to inherently small effect sizes as calculated and interpreted using traditional metrics (Aguinis et al., 2005). We have presented one recently proposed alternative to traditional metrics here: R^2_{mo} as proposed by Liu and Yuan (2020) which suggests that between 30.1% and 66.7% of the variability in the conscientiousness-performance relationship is accounted for by GMA. Additionally, to demonstrate the practical consequences of assuming a linear conscientiousness-performance relation among high-GMA employees, we simulated top-down selection procedures

to compare model selection efficacy (see Appendix B). Simulation results suggest that in most cases the quadratic interaction model is better at predicting top performance than is a linear conscientiousness-performance model. That is, using a rank-order strategy to select applicants with the highest scores on conscientiousness is likely to be inappropriate if they are also high in GMA. Instead, selection practitioners should consider using a multiplicative model or, alternatively, using a cut-off criterion in which individuals with conscientiousness scores below some prespecified values are eliminated from consideration. Thus, results suggest that moderation of the conscientiousness-performance relationship by GMA does have a meaningful impact on prediction of performance scores.

Overall, these results have broad and important implications for research on the relationship between conscientiousness, GMA, and performance. Results suggest that the conventional understanding of the relationship between conscientiousness and task performance does not hold at all GMA levels, particularly among individuals high in both GMA and conscientiousness. Assuming a linear relationship that is constant for all GMA levels is likely to yield misleading predictions of task performance and, consequently, selection decisions. Moreover, our results provide support for the broader theory of a conditional, compensatory relationship between conscientiousness and intelligence (e.g., Chamorro-Premuzic & Furnham, 2004). Individuals who are low in intelligence may be somewhat more likely to develop conscientiousness as strategy for goal attainment, such as through high discipline, rule-following, and greater work ethic. At the broadest level, results reinforce the importance of considering complex, interactive trait relationships to further our understanding of how traits influence behavior (Shoss & Witt, 2013).

A potentially fruitful area for extending the current findings into practice may be to develop interventions that target behaviors associated with conscientiousness and to tailor those interventions according to GMA level. For example, low- to average-GMA individuals may benefit from interventions that teach long-term goal prioritization, whereas individuals with high GMA may benefit from training that helps them to recognize and suppress extremely conscientious behaviors. Although personality traits are often treated as immutable (Chapman, et al., 2014), research has shown that personality can change over the lifespan (Roberts & Mroczek, 2008) and that targeted “bottom-up” interventions may be useful for intentional changes (Magidson et al., 2014). These “bottom-up” interventions target specific behaviors rather than the underlying trait directly. Thus, interventions that target specific problematic behaviors may be useful and, ultimately, even lead to changes in trait levels overall. Additionally, if these interventions showed success in producing long-term behavioral and performance changes, they would further support the possibility of a compensatory developmental relationship between conscientiousness and intelligence. Still, greater insight into the specific lower-order components of conscientiousness that drive a compensatory interaction with GMA is needed before targeted interventions can be developed.

The current results also point to the importance of several methodological updates relative to earlier studies that may account for differences in findings. First, results suggest that including quadratic terms was critical to identifying significant conscientiousness-GMA interactions for task performance. In fact, evidence for the hypothesized compensatory interaction was found only for quadratic conscientiousness. In the context of the ongoing dialogue about when and how conscientiousness may exhibit curvilinear relationships with performance and wellbeing outcomes (Carter et al., 2014; Carter et al., 2016; Le et al., 2011;

Nickel et al., 2018), these results suggest that curvilinearity may also depend on other individual differences. Additionally, we utilized a broader operationalization of conscientiousness than have most prior studies on a conscientiousness-GMA interaction, and we constrained the criterion domain to task performance only. Finally, the current study used appropriate IRT-scoring approaches wherever possible to reduce type 1 error and increase power. We advise other researchers interested in this topic to likewise include quadratic terms in analyses, carefully consider the breadth of the performance domain and conscientiousness measures, and utilize appropriate IRT-scoring.

Limitations and Future Directions

Despite the methodological advances of the current study over earlier investigations of this topic, the current study involved several limitations. First, due to the nature of the applied datasets used in the current study, all samples involved different measures of conscientiousness and GMA. Although the breadth of measures enhances the generalizability of results, it also prevents direct comparison between the samples. Moreover, although we strove to balance items within each measure across aspects (i.e., industriousness and orderliness) and facets, this was not always possible, particularly at the facet level. Within the orderliness aspect, measures generally over-represented the order facet and were inconsistent in the representation of other facets. In contrast, within industriousness, items tended to represent a greater range of facets. Thus, differences in item content might also account for slight differences in the current findings. Future research should consider using standardized measures across samples.

Due to uneven facet-level representation in the current study, we do not conduct analyses at the aspect or facet level. However, considering lower-order interactions may provide crucial insights to the compensatory role of conscientiousness. Research suggests that narrower aspects

of performance are best predicted by personality facets rather than traits (Barrick & Mount, 2005). Further, in their study on curvilinear relationships between conscientiousness facets and well-being, Carter et al. (2016) found evidence of differential facet relationships. For example, some researchers have suggested that the facet competence is associated with perfectionism at high levels. Perhaps high levels of competence are not beneficial to high-GMA individuals because it leads to an obsession with “perfection” when their standards for “good enough” are still superior to their lower-GMA peers. Future research should consider whether lower-order components of conscientiousness may differentially drive a conscientiousness-GMA interaction.

Additionally, each sample in the current study reflected different job and occupation types. Sample 1 included entry- to mid-level managers in a corporate environment. In contrast, Samples 3 and 4 represented food service and retail jobs. Sample 2 participants were employees in entry-level manufacturing positions. As with the breadth of conscientiousness measures included in the samples, the variety of job types enhances generalizability but limits direct comparison between samples. Thus, future studies should explicitly consider the effect of job and occupational context on this interaction. It is possible that differences in job type may explain the non-supportive findings in Sample 2 relative to other samples. For example, perhaps the manufacturing jobs in Sample 2 require a great degree of repetition or precision such that conscientiousness is always beneficial to performance regardless of GMA level. Likewise, other jobs requiring a high level of accuracy (e.g., surgeon, air traffic controller) may not exhibit these compensatory interaction patterns because even among high-GMA individuals, high conscientiousness levels improve rule- and detail-orientation.

Different job types may also account for lower inflection points in Sample 1 relative to inflection points identified in Samples 3 and 4. Relative to manufacturing and service jobs, the

knowledge work jobs represented in Sample 1 may be more likely to show a strong quadratic relationship between conscientiousness and performance among high-GMA employees. That is, complex jobs requiring a high level of knowledge work and creativity may show more pronounced compensatory effects because strictly ascribing to rules directly impedes performance or abstract decision-making. Indeed, researchers have suggested that conscientiousness may have detrimental effects on both creativity (Feist, 1998) and adaptability in decision-making (LePine et al., 2000).⁵ Further, some research has found that conscientiousness may benefit creativity for those with low but not high creative abilities (King et al., 1996). Thus, we encourage future researchers to investigate how conscientiousness-GMA interaction effects may vary by job type and, accordingly, different performance operationalizations.

Conclusion

Current results suggest that there is a compensatory interactive influence of conscientiousness and GMA on task performance. Specifically, there are differential, curvilinear conscientiousness-performance relationships at varying levels of GMA. Among high-GMA individuals, conscientiousness shows diminishing returns for performance (i.e., a downward curvilinear relationship). However, among average-GMA individuals, conscientiousness shows a generally linear, positive relationship with performance, and among low-GMA individuals, conscientiousness may be increasingly beneficial to performance at high levels (i.e., shows an upward curvilinear relationship). We believe these findings have the potential to spark research that shifts our understanding of the relation between conscientiousness, GMA, and performance.

⁵ We thank an anonymous reviewer for prompting us to consider this literature.

References

- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*(2), 219-245.
<https://doi.org/10.1037/0033-2909.121.2.219>
- Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*(2), 227-257. [https://doi.org/10.1016/S0160-2896\(96\)90016-1](https://doi.org/10.1016/S0160-2896(96)90016-1)
- Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology*, *90*(1), 94-107. <https://doi.org/10.1037/0021-9010.90.1.94>
- 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. (2005). Yes, personality matters: Moving on to more important matters. *Human Performance*, *18*(4), 359-372. https://doi.org/10.1207/s15327043hup1804_3
- Beaujean, A., Firmin, M. W., Attai, S., Johnson, C. B., Firmin, R. L., & Mena, K. E. (2011). Using personality and cognitive ability to predict academic achievement in a young adult sample. *Personality and Individual Differences*, *51*(6), 709-714.
<https://doi.org/10.1016/j.paid.2011.06.023>
- Bergold, S., & Steinmayr, R. (2018). Personality and intelligence interact in the prediction of academic achievement. *Journal of Intelligence*, *6*(2), 27.
<https://doi.org/10.3390/jintelligence6020027>

- Borman, W. C., & Motowidlo, S. M. (1993). Expanding the criterion domain to include elements of contextual performance. *Personnel Selection in Organizations; Jossey-Bass*.
- Carretta, T. R., & Ree, M. J. (2018). The relations between cognitive ability and personality: Convergent results across measures. *International Journal of Selection and Assessment*, 26(2-4), 133-144. <https://doi.org/10.1111/ijsa.12224>
- Carter, N. T., Dalal, D. K., Boyce, A. S., O'Connell, M. S., Kung, M.-C., & Delgado, K. M. (2014). Uncovering curvilinear relationships between conscientiousness and job performance: How theoretically appropriate measurement makes an empirical difference. *Journal of Applied Psychology*, 99(4), 564-586. <https://doi.org/10.1037/a0034688>
- Carter, N. T., Dalal, D. K., Guan, L., LoPilato, A. C., & Withrow, S. A. (2017). Item response theory scoring and detection of curvilinear relationships. *Psychological Methods*, 22(1), 191-203. <https://doi.org/10.1037/met0000101>
- Carter, N. T., Guan, L., Maples, J. L., Williamson, R. L., & Miller, J. D. (2016). The downsides of extreme conscientiousness for psychological well-being: The role of obsessive compulsive tendencies. *Journal of Personality*, 84(4), 510-522. <https://doi.org/10.1111/jopy.12177>
- Chamorro-Premuzic, T., & Furnham, A. (2004). A possible model for understanding the personality-intelligence interface. *British Journal of Psychology*, 95(2), 249-264. <https://doi.org/10.1348/000712604773952458>
- Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences*, 44(7), 1596-1603. <https://doi.org/10.1016/j.paid.2008.01.003>

- Chapman, B. P., Hampson, S., & Clarkin, J. (2014). Personality-informed interventions for healthy aging: Conclusions from a National Institute on Aging work group. *Developmental Psychology, 50*(5), 1426-1441. <https://doi.org/10.1037/a0034135>
- Cho, H. C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate?. *Journal of Business Research, 66*(9), 1261-1266. <https://doi.org/10.1016/j.jbusres.2012.02.023>
- Cortina, J. M. (1993). Interaction, nonlinearity, and multicollinearity: Implications for multiple regression. *Journal of Management, 19*(4), 915-922. [https://doi.org/10.1016/0149-2063\(93\)90035-L](https://doi.org/10.1016/0149-2063(93)90035-L)
- de Haro, J.-M., Castejón, J.-L., & Gilar, R. (2013). General mental ability as moderator of personality traits as predictors of early career success. *Journal of Vocational Behavior, 83*(2), 171-18. <https://doi.org/10.1016/j.jvb.2013.04.001>
- DeYoung, C. G. (2011). Intelligence and personality. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence* (pp. 711–737). New York, NY: Cambridge University Press.
- DeYoung, C.G. (2015). Cybernetic big five theory. *Journal of Research in Personality, 56*, 33-58. <https://doi.org/10.1016/j.jrp.2014.07.004>
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*(5), 880-896. <https://doi.org/10.1037/0022-3514.93.5.880>
- Di Domenico, S. I., & Fournier, M. A. (2015). Able, ready, and willing: Examining the additive and interactive effects of intelligence, conscientiousness, and autonomous motivation on

- undergraduate academic performance. *Learning and Individual Differences*, 40, 156-162.
<https://doi.org/10.1016/j.lindif.2015.03.016>
- Embretson, S. E. (1996). Item response theory models and spurious interaction effects in factorial ANOVA designs. *Applied Psychological Measurement*, 20(3), 201-212.
<https://doi.org/10.1177/014662169602000302>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*: Mahwah, NJ: Lawrence Erlbaum.
- Enns, R. A., & Reddon, J. R. (1998). The factor structure of the Wechsler Adult Intelligence Scale - revised: One or two but not three factors. *Journal of Clinical Psychology*, 54(5), 447-459. [https://doi.org/10.1002/\(sici\)1097-4679\(199806\)54:4<447::aid-jclp7>3.0.co;2-k](https://doi.org/10.1002/(sici)1097-4679(199806)54:4<447::aid-jclp7>3.0.co;2-k)
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2(4), 290-309.
https://doi.org/10.1207/s15327957pspr0204_5
- Furnham, A., Dissou, G., Sloan, P., & Chamorro-Premuzic, T. (2007). Personality and intelligence in business people: A study of two personality and two intelligence measures. *Journal of Business and Psychology*, 22(1), 99-109.
<https://doi.org/10.1007/s10869-007-9051-z>
- Hollenbeck, J. R., Brief, A. P., Whitener, E. M., & Pauli, K. E. (1988). An empirical note on the interaction of personality and aptitude in personnel. *Journal of Management*, 14(3), 441-451. <https://doi.org/10.1177/014920638801400307>
- Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>

- Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*(3), 340-362. [https://doi.org/10.1016/0001-8791\(86\)90013-8](https://doi.org/10.1016/0001-8791(86)90013-8)
- Judge, T. A., & Erez, A. (2007). Interaction and intersection: The constellation of emotional stability and extraversion in predicting performance. *Personnel Psychology, 60*(3), 573-596. <https://doi.org/10.1111/j.1744-6570.2007.00084.x>
- Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: Integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology, 98*(6), 875-925. <https://doi.org/10.1037/a0033901>
- Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44*(3), 486-507. <https://doi.org/10.1177/0049124114543236>
- King, L. A., Walker, L. M., & Broyles, S. J. (1996). Creativity and the five-factor model. *Journal of Research in Personality, 30*(2), 189-203. <https://doi.org/10.1006/jrpe.1996.0013>
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology, 86*(1), 148-161. <https://doi.org/10.1037/0022-3514.86.1.148>
- LaPalme, M., Tay, L., & Wang, W. (2018). A Within-Person Examination of the Ideal-Point Response Process. *Psychological Assessment, 30*(5), 567-581. <https://doi.org/10.1037/pas0000499>

Le, H., Oh, I.-S., Robbins, S. B., Ilies, R., Holland, E., & Westrick, P. (2011). Too much of a good thing: Curvilinear relationships between personality traits and job performance.

Journal of Applied Psychology, 96(1), 113-133. <https://doi.org/10.1037/a0021016>

LePine, J. A., Colquitt, J. A., & Erez, A. (2000). Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel Psychology, 53*(3), 563-593. <https://doi.org/10.1111/j.1744-6570.2000.tb00214.x>

Liu, H., & Yuan, K. H. (2020). New measures of effect size in moderation analysis.

Psychological Methods. https://doi.org/10.1037/met0000371

Magidson, J. F., Roberts, B. W., Collado-Rodriguez, A., & Lejuez, C. W. (2014). Theory-driven intervention for changing personality: Expectancy value theory, behavioral activation, and conscientiousness. *Developmental Psychology, 50*(5), 1442-1450.

Developmental Psychology, 50(5), 1442-1450.

<https://doi.org/10.1037/a0030583>

Maier, N. R. F. (1955). *Psychology in Industry: A Psychological Approach to Industrial*

Problems: Houghton Mifflin.

Maydeu-Olivares, A., Hernández, A., & McDonald, R. P. (2006). A multidimensional ideal point item response theory model for binary data. *Multivariate Behavioral Research, 41*(4),

445-472. https://doi.org/10.1207/s15327906mbr4104_2

McCrae, R. R., & Costa Jr, P. T. (1992). Discriminant validity of NEO-PIR facet scales. *Educational and Psychological Measurement, 52*(1), 229-237.

<https://doi.org/10.1177/001316449205200128>

Morse, B. J., Johanson, G. A., & Griffeth, R. W. (2012). Using the Graded Response Model to Control Spurious Interactions in Moderated Multiple Regression. *Applied Psychological Measurement, 36*(2), 122-146. <https://doi.org/10.1177/0146621612438725>

Measurement, 36(2), 122-146. <https://doi.org/10.1177/0146621612438725>

- Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology, 79*(4), 475-480. <https://doi.org/10.1037/0021-9010.79.4.475>
- Mount, M. K., Barrick, M. R., & Strauss, J. P. (1999). The joint relationship of conscientiousness and ability with performance: Test of the the interaction hypothesis. *Journal of Management, 25*(5), 707-721. <https://doi.org/10.1177/014920639902500505>
- Moutafi, J., Furnham, A., & Paltiel, L. (2004). Why is conscientiousness negatively correlated with intelligence?. *Personality and Individual Differences, 37*(5), 1013-1022. <https://doi.org/10.1016/j.paid.2003.11.010>
- Murray, A. L., Johnson, W., McGue, M., & Iacono, W. G. (2014). How are conscientiousness and cognitive ability related to one another? A re-examination of the intelligence compensation hypothesis. *Personality and Individual Differences, 70*, 17-22. <https://doi.org/10.1016/j.paid.2014.06.014>
- Nickel, L. B., Roberts, B. W., & Chernyshenko, O. S. (2019). No evidence of a curvilinear relation between conscientiousness and relationship, work, and health outcomes. *Journal of Personality and Social Psychology, 116*(2), 296-312. <https://doi.org/10.1037/pspp0000176>
- O'Reilly III, C. A., & Chatman, J. A. (1994). Working smarter and harder: A longitudinal study of managerial success. *Administrative Science Quarterly, 60*, 603-627. <https://doi.org/10.2307/2393773>
- Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology, 60*, 995-1027.

- Ones, D. S., Dilchert, S., & Viswesvaran, C. (2012). Cognitive Abilities. In N. Schmitt (Ed.), *The Oxford Handbook of Personnel Assessment and Selection* (pp. 179-224). Oxford University Press.
- Penney, L. M., David, E., & Witt, L. (2011). A review of personality and performance: Identifying boundaries, contingencies, and future research directions. *Human Resource Management Review, 21*(4), 297-31. <https://doi.org/10.1016/j.hrmr.2010.10.005>
- Perry, S. J., Hunter, E. M., Witt, L. A., & Harris, K. J. (2010). P = f(Conscientiousness × Ability): Examining the facets of conscientiousness. *Human Performance, 23*(4), 343-36. <https://doi.org/10.1080/08959285.2010.501045>
- Pfaffel, A., Kollmayer, M., Schober, B., & Spiel, C. (2016). A missing data approach to correct for direct and indirect range restrictions with a dichotomous criterion: A simulation study. *PLOS ONE, 11*(3), e0152330. <https://doi.org/10.1371/journal.pone.0152330>
- Postlethwaite, B., Robbins, S., Rickerson, J., & McKinniss, T. (2009). The moderation of conscientiousness by cognitive ability when predicting workplace safety behavior. *Personality and Individual Differences, 47*(7), 711-716. <https://doi.org/10.1016/j.paid.2009.06.008>
- Rammstedt, B., Danner, D., & Martin, S. (2016). The association between personality and cognitive ability: Going beyond simple effects. *Journal of Research in Personality, 62*, 39-44. <https://doi.org/10.1016/j.jrp.2016.03.005>
- Roberts, J. S., Donoghue, J. R.s, & Laughlin, J. E. (2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement, 24*(1), 3-32. <https://doi.org/10.1177/01466216000241001>

- Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, 17(1), 31-35. <https://doi.org/10.1111/j.1467-8721.2008.00543.x>
- Sackett, P. R., Gruys, M. L., & Ellingson, J. E. (1998). Ability–personality interactions when predicting job performance. *Journal of Applied Psychology*, 83(4), 545-556. <https://doi.org/10.1037/0021-9010.83.4.545>
- Samejima, F. (1997). Graded response model. *Handbook of Modern Item Response Theory*. (pp. 85-100): Springer.
- Shoss, M. K., & Witt, L. A. (2013). Trait interactions and other configural approaches to personality. In N. D. Christiansen & R. P. Tett (Eds.), *Handbook of Personality at Work* (pp. 392-418). New York, NY: Routledge.
- Stark, S., Chernyshenko, O. S., Drasgow, F., & Williams, B. A. (2006). Examining assumptions about item responding in personality assessment: Should ideal point methods be considered for scale development and scoring?. *Journal of Applied Psychology*, 91(1), 25-39. <https://doi.org/10.1037/0021-9010.91.1.25>
- Van Iddekinge, C. H., Aguinis, H., Mackey, J. D., & DeOrtentiis, P. S. (2017). A meta-analysis of the interactive, additive, and relative effects of cognitive ability and motivation on performance. *Journal of Management*, 44(1), 249-279. <https://doi.org/10.1177/0149206317702220>
- Vroom, V. H. (1995). *Work and motivation*. San Francisco, CA: Jossey-Bass Inc., Publishers.
- Waller, M. I. (1989). Modeling guessing behavior: A comparison of two IRT models. *Applied Psychological Measurement*, 13(3), 233-243. <https://doi.org/10.1177/014662168901300302>

Wechsler, D. (1955). Manual for the Wechsler adult intelligence scale.

Witt, L. A., Burke, L. A., Barrick, M. A., & Mount, M. K. (2002). The interactive influences of conscientiousness and agreeableness on job performance. *Journal of Applied Psychology*, 87(1), 164-169. <https://doi.org/10.1037/0021-9010.87.1.164>

Witt, L. A. (2002). The interactive effects of extraversion and conscientiousness on performance. *Journal of Management*, 28(6), 835-851. [https://doi.org/10.1016/S0149-2063\(02\)00188-5](https://doi.org/10.1016/S0149-2063(02)00188-5)

Wright, P. M., Kacmar, K. M., McMahan, G. C., & Deleeuw, K. (1995). P= f(MXA): Cognitive ability as a moderator of the relationship between personality and job performance. *Journal of Management*, 21(6), 1129-1139. [https://doi.org/10.1016/0149-2063\(95\)90025-X](https://doi.org/10.1016/0149-2063(95)90025-X)

Yuan, X., Li, Y., Xu, Y., & Huang, N. (2018). Curvilinear effects of personality on safety performance: The moderating role of supervisor support. *Personality and Individual Differences*, 122, 55-61. <https://doi.org/10.1016/j.paid.2017.10.005>

van Buuren, S., & Groothuis-Oudshoorn, K. (2010). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1-68. <https://doi.org/10.18637/jss.v045.i03>

Zhang, J., & Ziegler, M. (2015). Interaction effects between Openness and Fluid Intelligence predicting scholastic performance. *Journal of Intelligence*, 3(3), 91-111. <https://doi.org/10.3390/jintelligence3030091>

Ziegler, M., Knogler, M., & Bühner, M. (2009). Conscientiousness, achievement striving, and intelligence as performance predictors in a sample of German psychology students: Always a linear relationship? *Learning and Individual Differences*, 19(2), 288-292. <https://doi.org/10.1016/j.lindif.2009.02.001>

Table 1

Descriptive Statistics and Correlations Between All Predictors and Criteria

Variable	Sample 1			Sample 2			1.	2.	3.	4.	5.	6.	7.	8.
	M	SD	r _{xx} (N)	M	SD	r _{xx} (N)								
1. C	0.00	1.00	<u>.80</u> (3,963)	0.00	1.00	<u>.86</u> (5,570)	-	-.08	.13*	.00	.01	-.05	-.11	.66**
2. GMA	0.00	1.00	<u>.99</u> (4,352)	0.00	1.00	<u>.77</u> (5,570)	-.08	-	.12*	-.04	.36**	.01	.65**	-.14*
3. Task Performance	0.00	1.00	<u>.83</u> (3,963)	0.00	1.00	<u>.93</u> (261)	.13*	.10*	-	.01	.01	.02	.08	.07
4. GMA ²	1.00	1.25	-	1.00	1.17	-	-.28**	.11	-.13*	-	-.08	-.07	-.01	-.01
5. C ²	1.00	1.25	-	1.00	1.42	-	.05	-.14*	-.14**	-.04	-	-.08	.13*	-.08
6. C x GMA	-0.08	0.96	-	-0.08	0.92	-	.14*	.06	.14**	-.21**	-.04	-	-.01	.17**
7. C x GMA ²	0.14	1.49	-	-0.05	1.33	-	-.22**	.63**	-.04	.22**	.01	-.13*	-	-.17**
8. C ² x GMA	0.06	1.33	-	0.01	1.30	-	.71**	-.09	.01	-.15**	.18**	.12*	-.24**	-
			Sample 3			Sample 4								
1. C	0.00	1.00	<u>.71</u> (2,757)	0.00	1.00	<u>.78</u> (948)	-	.16**	.08*	-.05	-.04	-.09**	.27**	.58**
2. GMA	0.00	1.00	<u>.93</u> (2,854)	0.00	1.00	<u>.88</u> (948)	.15**	-	.09**	-.07*	.32**	-.06	.61**	.25**
3. Task Performance	0.00	1.00	<u>.92</u> (1,725)	0.00	1.00	<u>.92</u> (948)	.10**	.08**	-	-.04	.02	.02	.07*	.02
4. GMA ²	1.00	1.49	-	1.00	1.30	-	.06*	.00	-.01	-	.02	.23**	-.15**	-.09**
5. C ²	1.00	1.37	-	1.00	1.49	-	-.05	-.12**	.00	.04	-	.19**	.12**	.06*
6. C x GMA	0.15	1.03	-	0.16	1.01	-	-.01	-.07*	.03	.23**	.16**	-	-.14**	.12**
7. C x GMA ²	-0.01	1.96	-	-0.10	1.70	-	.26**	.55**	.08**	-.11**	-.09**	-.16**	-	.37**
8. C ² x GMA	-0.07	1.82	-	-0.06	1.78	-	.59**	.21**	.03	-.09**	-.07**	-.13**	.45**	-

Note. Descriptive statistics are reported for standardized theta-values, except for interaction terms which are non-standardized

products of standardized values. Descriptive statistics and correlations are reported for only those participants with valid data for all

variables. Underlined values indicate coefficient alpha. Correlations and descriptive statistics reported for subset data used in

regression analyses. Sample 1 and 3 correlations are in the sub-diagonal. Sample 2 and 4 correlations are above the diagonal. P-values

are two-tailed. $N_1 = 300$. $N_2 = 261$. $N_3 = 1,413$. $N_4 = 948$. $C =$ Conscientiousness $r_{xx} =$ IRT-based marginal reliability estimated in full data. $N =$ sample size used for IRT estimation.

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

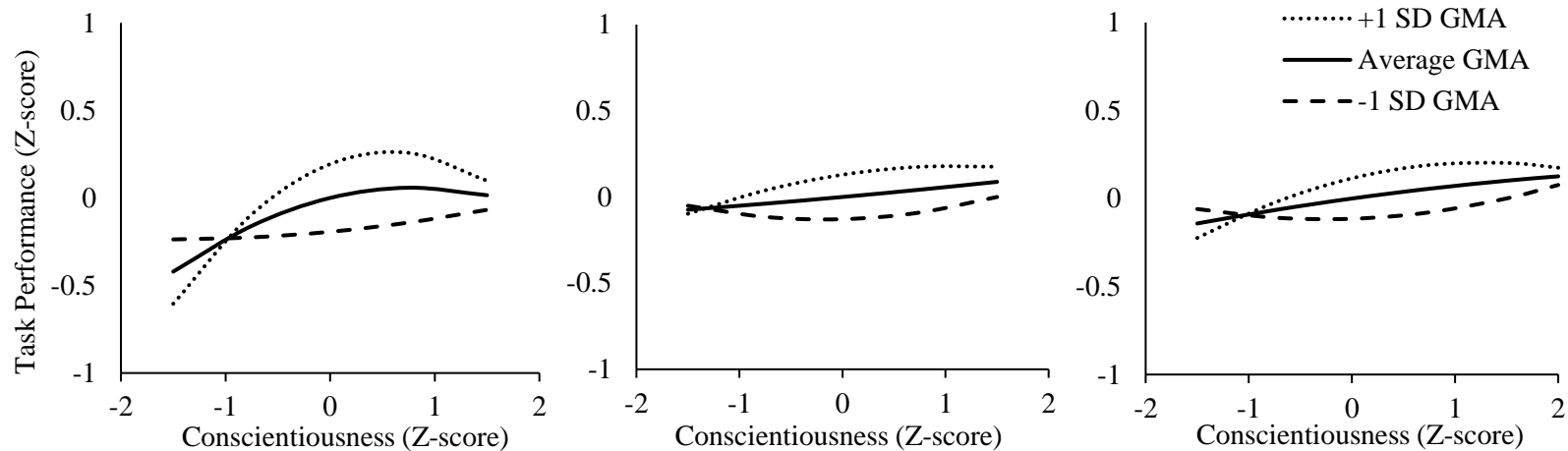
Table 2*Results of Task Performance Predicted by Conscientiousness and GMA*

Predictor	Sample 1			Sample 2			Sample 3			Sample 4		
	<i>b</i>	SE	<i>p</i>	<i>b</i>	SE	<i>p</i>	<i>b</i>	SE	<i>p</i>	<i>b</i>	SE	<i>p</i>
Intercept	0.17*	.09	.025	0.02	.094	.397	0.01	.04	.401	0.04	.045	.208
GMA	0.19**	.08	.009	0.16*	.075	.016	0.13**	.03	< .001	0.12**	.039	.002
C	0.15*	.08	.025	0.14*	.082	.039	0.05*	.03	.044	0.08*	.042	.029
C ²	-0.09*	.05	.037	-0.03	.035	.181	0.00	.02	.438	-0.01	.022	.335
GMA ²	-0.06 [†]	.05	.094	0.01	.049	.421	-0.02	.02	.133	-0.04 [†]	.024	.056
C x GMA ²	-0.06	.05	.113	0.01	.059	.459	0.03 [†]	.02	.057	0.02	.022	.255
C x GMA	0.09 [†]	.06	.070	0.03	.067	.335	0.04 [†]	.03	.086	0.06*	.034	.037
C ² x GMA	-0.11*	.06	.039	-0.02	.062	.394	-0.04*	.02	.014	-0.05*	.022	.016
ΔR^2		.016			.001			.005			.006	
R^2_{mo}		.301			.062			.663			.619	

Note. $N_1 = 300$. $N_2 = 261$. $N_3 = 1,413$. $N_4 = 948$. P-values are one-tailed. (See Cho & Abe, 2013 for a discussion on the use of one- vs.

two-tailed tests.) C = Conscientiousness.

[†] $p < .10$. * $p < .05$. ** $p < .01$.

Figure 1*Interactive Influence of Conscientiousness and GMA on Task Performance*

Note. Panels from left to right: Sample 1, Sample 3, and Sample 4. Range of graphed conscientiousness levels includes only values reflected by sample data. Inflection points (% population above) for Sample 1 (left panel): 0.59 (27.73%) for high-GMA; .81 (21.01%) for average-GMA; and -1.56 (94.02%) for low-GMA. Inflection points (% population above) for Sample 3 (middle panel): 1.14 (12.76%) for high-GMA; -9.00 (100.00%) for average-GMA; and -0.19 (57.34%) for low-GMA. Inflection points (% population above) for Sample 4 (right panel): 1.27 (10.24%) for high-GMA; 4.50 (0.00%) for average-GMA; and -0.26 (60.67%) for low-GMA.

Appendix A

Additional Analyses for Sample 2

Because personality and GMA batteries in Sample 2 were used to select employees and included data on all applicants, but only performance data for hired employees were available, there was evidence of direct range restriction in Sample 2. Minimum values for conscientiousness and GMA in the full applicant data (Conscientiousness: Min. = -2.58, Max = 4.00, $SD = .92$; GMA: Min = -2.46, Max = 1.99, $SD = .87$) were substantially lower than were minimum values among hired employees (Conscientiousness: Min = -1.22, Max = 3.77, $SD = .84$; GMA: Min = -1.74, Max. = 1.99, $SD = .82$). To determine whether range restriction might have affected findings, we conducted post-hoc analyses that included multivariate imputation by chained equations using the ‘mice’ package (van Buuren, 2018; Pfaffel et al., 2016) to impute performance data for the applicants not hired.

Results of analyses using imputed data to account for range restriction still did not show evidence of a significant conscientiousness-GMA interaction. However, results did show a marginally significant quadratic interaction, $b = -0.03$, $p = .082$, and GMA accounted for a proportion of the variability in the conscientiousness-performance relationship (45%) comparable to that of other samples. Results are shown in Table A1 below.

Additionally, to evaluate the potential impact of limited power on our findings in Sample 2, we conducted three Monte Carlo simulations using parameter estimates from Sample 1, 3, and 4 respectively and sample size from Sample 2 ($N = 261$). Across 1,000 replications, simulations suggested that when sample size is reduced to that of Sample 2, the true positive rate is .68 for Sample 1, .22 for Sample 3, and .24 for Sample 4. In other words, if the findings for Sample 1, 3, or 4 were true in the population, the sample size in Sample 2 is under-powered.

Table A1

Results of Task Performance Predicted by Conscientiousness and GMA in Sample 2 Corrected for Range Restriction

Predictor	<i>b</i>	SE	<i>p</i>
Intercept	-0.10 [†]	.11	.082
GMA	0.14	.10	.124
C	0.12	.10	.485
C ²	0.00	.06	.240
GMA ²	0.04	.06	.368
C x GMA ²	0.02	.06	.121
C x GMA	0.10	.08	.343
C ² x GMA	-0.03 [†]	.07	.082
ΔR^2		.009	
R^2_{mo}		.450	

Note. $N = 5,570$. P-values are two-tailed. C = Conscientiousness.

[†] $p < .10$. * $p < .05$. ** $p < .01$.

Appendix B

Top-Down Selection Simulation Analyses

For each dataset, 1,000 samples of 100 people each were randomly chosen as pools of “applicants.” We then used the coefficients for the “standard” linear model (i.e., using GMA and conscientiousness as linear predictors of performance) to calculate predicted performance and, separately, the coefficients from the quadratic interaction model to calculate predicted performance for each sample. Using these predicted values, we selected (i.e., “hired”) both the top 10% and top 20% of applicants. We then used the actual, observed performance for each selected applicant to calculate (a) mean performance, (b) maximum performance (i.e., the performance score for the top performer), and (c) the minimum performance (i.e., the performance score for the lowest performer). Then, we determined the “winner” of the two models regarding their selection decisions. For the minimum and maximum, the winner was the model that picked the “better best performer” and the “better worst performer,” respectively. For the mean, we considered the mean scores to be different only if they showed greater than a .10 *SD* difference (i.e., one standard error for the population of 100 applicants). Table B1 shows the results for simulated selection scenarios. The number outside the parentheses is the percentage of times (of the 1,000 simulated applicant pools) that the quadratic interaction model was the winner for the selection ratio. The number inside the parentheses is the percentage of times the standard linear model was the winner.⁶

The results suggest that, in most cases, the quadratic interaction model is better at predicting top performance, especially when the selection ratio is 20%. The quadratic interaction model generally resulted in higher average performance, the selection of the better top

⁶ The percentage of “ties” between the two models can be calculated by summing the two numbers and subtracting from 100%.

performer, and the selection of the better worst performer. In most instances, the quadratic interaction model won more than twice as many times as the linear model, indicating its general superiority in making top-down selection decisions. Thus, the simulation results suggest that, despite small effect sizes using traditional metrics (e.g., ΔR^2), the quadratic interaction has substantial practical value in its ability to more effectively predict performance. Notably, this is true even for Sample 2, despite the non-significance of the interaction. These results further point to the limited power of Sample 2 as a potential explanation for the non-significance of findings.

Table B1

Comparison of the Quadratic Interaction Model and the Standard Linear Model Efficacy in Simulated Top-Down Selection

Sample	% Times Model Resulted in Selecting Better Average Performance [†]		% Times Model Resulted in Selecting the Better Top Performer		% Times Model Resulted in Selecting the Better Lowest Performer	
	SR=10%	SR=20%	SR=10%	SR=20%	SR=10%	SR=20%
Sample 1	30%(29%)	52%(7%)	17%(29%)	76%(17%)	37%(12%)	68%(3%)
Sample 2	48%(26%)	55%(27%)	90%(3%)	68%(6%)	65%(15%)	75%(13%)
Sample 3	27%(26%)	57%(18%)	42%(22%)	68%(15%)	30%(23%)	68%(8%)
Sample 4	27%(26%)	57%(18%)	42%(22%)	68%(15%)	30%(23%)	68%(8%)

Note. Values outside parentheses correspond to the quadratic interaction model; values inside parentheses correspond to the standard linear model. Adding these values and subtracting from 100% results in the % of resulting selections that were a “tie;” values in bold indicate 50% or greater “wins” for the quadratic interaction model compared to the linear model. Boldface indicates the quadratic interaction model outperformed the standard linear model. SR indicates selection ratio out of 100 applicants.

[†]For this metric “better” is defined as a mean .10 *SD* higher (i.e., 1 standard error) than the competing mean; for all other metrics “better” is defined as any non-zero difference favoring the model.