aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Improving Graduate-School Admissions by Expanding Rather Than Eliminating Predictors

Perspectives on Psychological Science
1–7

DOI: 10.1177/17456916221081873
www.psychologicalscience.org/PPS
SAGE

Christopher D. Nye and Ann Marie Ryan
Department of Psychology, Michigan State University

Abstract
The article by Woo et al. (this issue) reviews the existing research on graduate-school admissions measures. The goal of this commentary is to expand on their review and suggest several ways of supplementing the Graduate Record Examination (GRE) to both increase the predictive validity of admissions decisions and improve the diversity of a graduate program. We rely on several decades of research to suggest assessing both conscientiousness and vocational interests and combining the scores from these predictors with the GRE to inform admissions decisions. In addition, we also propose several ways of expanding recruitment efforts to attract qualified underrepresented minority applicants to improve the diversity of the applicant pool.

Keywords
graduate student admissions, GRE, personality, vocational interests, recruiting

Woo et al. (2022; this issue) provide a comprehensive summary of the research on graduate-school admissions measures. We agree with their conclusion that there is "very limited evidence for psychometric bias" (proof p. 14) but substantial evidence for the predictive validity of the GRE. However, readers may wonder how graduate programs can successfully use the GRE, which shows strong validity for predicting academic success, without disproportionately screening out underrepresented minority (URM) applicants because of the large subgroup differences in GRE scores. Similar issues regarding the use of cognitive ability testing for employee selection have been examined extensively over the past several decades in the organizational psychology literature. Therefore, this research can provide guidance on how to supplement the GRE in a way that will increase the predictive validity of admissions decisions and also improve the diversity of a graduate program.

To be clear, we agree with Woo et al. that the most promising way to address subgroup differences on the GRE is to address the underlying sources of these differences (e.g., disparities in educational opportunities, physical environments, or socioeconomic status). Nevertheless, addressing these systemic societal issues will take time, and things can be done right now to improve the admissions process and also mitigate the effects of subgroup differences on the GRE. Specifically, we propose expanding the range of predictors that are used to make admissions decisions, being intentional about how the predictor scores are combined, and using targeted recruiting of qualified URM applicants to improve the overall quality and diversity of the applicant pool.

Expanding the Predictors of Academic Success

One concrete way to increase the validity of the admissions process and address diversity goals is to expand the range of predictors that are assessed and used to make admissions decisions. The GRE is an assessment of cognitive ability, which has been defined as the "ability to learn" (Schmidt, 2002), and therefore has obvious relevance for predicting academic performance. Nevertheless, successful graduate-school performance requires more than just cognitive ability. It also requires interest in doing the work, a willingness to work hard,

Corresponding Author:
Christopher D. Nye, Department of Psychology, Michigan State University
Email: nyechris@msu.edu

and the ability to meet deadlines and program milestones (e.g., a Master's thesis), among other things. These characteristics are not assessed by the GRE. Therefore, although the GRE has validity for predicting academic success, it also provides an incomplete assessment of an applicant. Consequently, instead of removing the GRE from the admissions process, a better approach would be to add predictors that can provide a more comprehensive evaluation of an applicant's potential for academic success. In addition to increasing the predictive validity of the admissions process, combining a broader range of predictors can also help to *reduce* subgroup differences in admissions decisions and *increase* graduate student diversity in a program. We next summarize research on two key predictors that are not currently assessed well during the admissions process but that have been shown to have both validity for predicting academic success and small subgroup differences across race and/or gender groups.

Conscientiousness

Academic success requires not just an ability to learn but also hard work, persistence, organization, and dependability, among other things. These characteristics are all aspects of the personality trait known as conscientiousness, and a substantial amount of research has supported the importance of this trait for predicting performance both at work and in school. In a recent report from the National Academies of Sciences, Engineering, and Medicine (2017), a panel of experts identified conscientiousness as one of the key characteristics related to academic success in college. Likewise, a meta-analysis conducted by Poropat (2009) found that the corrected correlation between conscientiousness and academic performance was .22. Incidentally, this was only slightly smaller than the meta-analytic correlation between cognitive ability and academic performance (corrected correlation = .25) in that same study. Likewise, in a recent longitudinal study of undergraduate performance, my colleagues and I (Nye et al., 2021) showed that conscientiousness was related to several academic outcomes, including course grades, citizenship behavior, and intentions to leave the university, even after controlling for cognitive ability and interests. Moreover, we demonstrated that motivation partially mediated many of these effects—students who scored higher on conscientiousness performed better in school because they were more likely to attend their classes and exert effort on their coursework. Although these studies focused on undergraduate academic performance, conscientiousness is likely even more relevant for graduate-school performance, where there is less structure and more independence.

In addition to the validity of conscientiousness, this trait also tends to show small subgroup differences across race and gender. Hough et al. (2001) reviewed this research and found standardized mean differences in effect sizes (d) of 0.06[1] between Black and White samples and −0.08 between men and women. Using Cohen's (1988) guidelines for interpreting effect sizes, these differences would be considered negligible. In addition to the trivial differences across subgroups, Keiser et al. (2016) found that controlling for conscientiousness could reduce prediction biases across men and women during the admissions process.

The research cited above suggests that considering both conscientiousness and cognitive ability during the admissions process can help to improve the validity of the process and also mitigate differences across demographic groups (i.e., sex and race). However, Woo et al. suggest that faking is a major concern with using personality in high-stakes settings (Morgeson et al., 2007). The concern here is that applicants for admissions or jobs can identify the best (or most socially desirable) responses and are motivated to distort their responses to increase their chances of being selected. Despite this concern, recent psychometric research has identified ways of mitigating the effects of faking on high-stakes personality assessments. For example, instead of presenting applicants with a single personality statement (e.g., "*I tend to work hard*"), forced-choice measures require individuals to choose the statement out of a pair (e.g., "*I tend to work hard*" vs. "*I get along well with other students*") that best describes them. This forced-choice format makes faking more difficult because it is harder for applicants to identify the "best" response out of the two equally desirable alternatives. The resistance to faking of forced-choice measures has been demonstrated in both high-stakes employment settings (e.g., Stark et al., 2014; Trent et al., 2020) and in meta-analytic research (Cao & Drasgow, 2019). However, not all forced-choice measures are created equal—it is important to incorporate multidimensional statement pairs that are matched on both the extremity and social desirability of the statements to maximize resistance to faking (Cao & Drasgow, 2019).

As noted by Woo et al., coaching is another important concern for the use of personality assessments. Although coaching is often treated as a form of faking, it is also qualitatively different in that applicants are coached to choose certain responses or types of items rather than try to determine the most socially desirable responses for themselves (i.e., faking). For this reason, Zickar and Robie (1999) found that score inflation was higher for individuals who were coached on how to fake compared with a control group that was asked to fake but not coached. Despite these differences, the

forced-choice format would likely also mitigate the effects of coaching by asking applicants to pick between two statements that are matched on their social desirability. In addition, using the forced-choice format in combination with computer adaptive testing (CAT) could make coaching more difficult (Drasgow et al., 2009). With CAT, each applicant sees a different set of items that are chosen from a large pool of items to match their estimated level of the trait. When CAT is combined with the forced-choice format, both statements in an item pair can be selected adaptively, creating a large number of potential statement combinations that can be administered to applicants. The large number of potential pairings and the fact that each applicant would potentially see a unique set of items can both help to mitigate the effects of coaching and other forms of cheating on applicants' scores (Drasgow et al., 2009). However, more research is needed to evaluate the potential effects of coaching on computer-adaptive, forced-choice assessments. In addition, even with these sophisticated psychometric tools, it will be important to maintain the security of both the content and scoring algorithms used for these assessments to mitigate the potential influence of coaching efforts.

Vocational interests

Another widely researched predictor of academic success is vocational interests. Vocational interests reflect preferences for certain activities, tasks, occupations, and environments. In other words, vocational interests are associated with the types of work that people like to do and the types of subjects they like to study. The National Academies of Sciences, Engineering, and Medicine (2017) report on college student success identified a strong interest in the subject matter as one of the eight most promising characteristics that are linked to college success. In a meta-analysis on the relationship between interests and performance, my colleagues and I (Nye et al., 2012) found a baseline correlation of .32 between vocational interest fit (i.e., the match between individuals' interests and their coursework or major) and academic performance. Moreover, these effects hold even after controlling for cognitive ability and conscientiousness (Van Iddekinge et al., 2011). In fact, we also found (Nye et al., 2021) that interest fit was one of the strongest predictors of grades, citizenship behavior, and intent to stay at the university even after controlling for ACT scores and conscientiousness in a sample of undergraduates. Given the incremental validity of vocational interests over both cognitive ability and conscientiousness, considering interests during the admissions process can help to identify applicants who are truly interested in the types of tasks that are performed in graduate school (e.g., writing papers/reports, conducting research, developing theories, conducting statistical analyses) and improve the prediction of success as a graduate student. This is important because many individuals may have inaccurate views of the work that is done in graduate school or their interest in those activities (see Nye, Perlus, & Rounds, 2018), and their performance, satisfaction, and motivation may suffer as a result (Nye et al., 2021).

As with conscientiousness, vocational interests tend to show a different pattern of subgroup differences compared with cognitive ability. In their meta-analytic review, Jones et al. (2021) found that average score differences (d) between Black and White samples ranged from −0.23 to 0.18 across interest types. For the interest profile of psychology occupations, which consists of social, artistic, and investigative interests, these differences were just −0.15, 0.00, and 0.11. However, Roth et al. (2017) found slightly larger differences when comparing Hispanic and Asian samples with White samples (e.g., Cohen's d ranged from −0.04 to 0.31). For the interest profile of psychology occupations, these differences were 0.21, 0.14, and 0.17, respectively, for comparisons with Hispanic samples and 0.11, 0.15, and .016 for comparisons with Asian samples. Nevertheless, these differences still represent small effects according to Cohen's (1988) guidelines. In contrast, much larger differences have been observed when comparing scores between men and women, ranging from −0.68 to 0.84 across the six interest types (average d = 0.45) but with differences as high as 1.11 for more specific interests (e.g., interest in engineering; Su et al., 2009). For social, artistic, and investigative interests, these differences were −0.68, −0.35, and 0.26, respectively. In other words, although the differences were larger, the interest profile for psychology occupations mostly favors women. We found that, despite these differences—and partially because of them—excluding vocational interests from admissions decisions resulted in an omitted-variable problem that actually *increased* gender biases in admissions decisions (Nye, Butt, et al., 2018).

This commentary focuses on conscientiousness and vocational interests because these constructs are supported by vast bodies of research. However, other predictors may also be useful for increasing validity and reducing subgroup differences in graduate-school admissions. For example, both biodata and situational-judgment measures have shown validity for predicting academic performance (Schmitt et al., 2009; Zhang & Kuncel, 2020). In addition, some evidence suggests that these measures can reduce subgroup differences in the admissions process (Schmitt et al., 2009; Sinha et al., 2011). However, other evidence suggests that these assessments are highly susceptible to coaching (Cullen et al., 2006). Structured interviews have also been

proposed as methods with high validity and somewhat lower subgroup differences (Sackett et al., 2021), but these assessments have the same challenges related to impression management and coaching and are resource intensive to administer and score, unless delivered in an asynchronous video format. Nevertheless, these and other assessments can still be useful for the graduate-school admissions process. Although space does not permit a comprehensive discussion of all possible predictors, the key point is that the alternative predictors that are used for graduate-school admissions must both add validity to the GRE and help to reduce subgroup differences in the admissions process.

Combining Predictors

In view of the research cited above, combining the GRE with measures of conscientiousness and vocational interests could provide several advantages for graduate-school admissions. Some readers may be thinking that these characteristics are "already covered" in graduate-school admissions through the use of personal statements, letters of recommendation, or unstructured interviews. However, as noted by Woo et al., research has found that these alternative predictors have low construct validity, have weak correlations with graduate-school performance, and are influenced by rater biases. Although these predictors are not quality measures of conscientiousness or vocational interests, psychometrically sound measures of these characteristics exist or can be readily developed (i.e., for specific programs or needs) and can be administered both economically and efficiently.

In addition to increasing the validity of admissions decisions, combining conscientiousness and vocational interests with the GRE could also help to reduce subgroup differences and adverse impact in admissions decisions because of their smaller score differences across racial groups. For example, Sackett and Ellingson (1997) showed that weighting conscientiousness and cognitive ability equally, rather than putting a higher weight on assessments such as the GRE, can dramatically reduce subgroup differences. Moreover, combining these predictors will not only reduce the race differences that would be observed with use of the GRE alone but also mitigate the observed differences between men and women on vocational interest measures (Nye, 2022). Likewise, recent work on pareto-optimization provides another approach to combining predictors (De Corte et al., 2007; Song et al., 2017), but a detailed discussion of this approach is beyond the scope of this commentary.

Another important consideration is *how* these predictors are combined to make decisions. In their article, Woo et al. compare mechanical and clinical approaches to combining admissions information. However, in the example scenario that they provide on proof page 3, they implicitly assume a multiple-hurdle approach in which applicants are screened sequentially on different predictors. This sequential screening can involve either selecting individuals with the highest scores on each of the predictors or ensuring that applicants pass a minimum cut score on one or more of the predictors. An alternative technique is to use a compensatory approach in which all predictors are considered simultaneously (i.e., the GRE, conscientiousness, and interests are considered at the same time). An advantage of the compensatory approach is that high scores on one predictor (e.g., conscientiousness) can make up for low scores on another predictor (e.g., the GRE). This compensatory effect may be particularly important for graduate-school admissions, where scores on some traditional admissions criteria, including the GRE, personal statements, resumes, and letters of recommendation, can be adversely affected for some applicants who do not have the same access to opportunities (e.g., research experience) or resources (e.g., financial resources) that can strengthen their applications. In this case, the compensatory approach can help to identify individuals who are likely to be successful graduate students despite low scores on one or more of the predictors that are used. Note that this approach is not necessarily the same as the "holistic" approach described by Woo et al. in that the compensatory predictors can be combined either mechanically or clinically—the key distinction is that all predictors are considered simultaneously rather than sequentially.

In some cases, a compensatory approach may not be feasible because of the large number of applicants or because using several assessments up front would be costly and time-consuming for applicants, which are burdens that would disproportionately affect URMs. In contrast, a multiple-hurdle approach can be more cost effective for both universities and applicants if the added assessments are administered online during the application process, as is done by many organizations using personality or interest measures to screen job applicants. The multiple-hurdle approach can also be more effective than the compensatory approach at increasing the selection rates of URM applicants. However, this will depend on several factors, including the predictors that are used and the number of applicants who are screened out at each stage (Finch et al., 2009; Sackett & Roth, 1996).

Recruiting Diverse and Qualified Applicants

Combining the GRE, conscientiousness, and vocational interests can reduce the overall score differences between Black and White applicants to a relatively

small effect but does not completely eliminate them (Newman & Lyon, 2009). Therefore, additional steps are necessary to further increase the diversity of admitted students. As mentioned by Woo et al., one way to do this is to recruit a "healthy pipeline of URM students" (proof p. 16) who have the knowledge, skills, and abilities that are required in a specific graduate program. Past research has demonstrated that recruitment based on demographic characteristics alone is insufficient to improve the diversity of those selected for a job. Instead, recruitment efforts should focus on recruiting for *both* diversity and qualifications to further reduce adverse impact (Newman & Lyon, 2009).

Organizational and social psychological research on attracting qualified URMs is seldom applied in graduate-student recruitment. In particular, graduate programs need to be more intentional about how they design websites, information sessions, and other recruiting efforts to consider what the program is signaling to potential applicants. Identity-contingency cues (Emerson & Murphy, 2014) that are communicated in recruiting efforts can provide applicants with information about whether their identity will be accepted and valued in the program. At a basic level, organizations are more attractive to URM applicants when they include diversity statements in their recruitment materials (Kim & Gelfand, 2003) or show diverse groups of individuals in websites, videos, and social-media pages (Avery, 2003; Avery et al., 2004; Walker et al., 2009). However, to truly have an impact on the recruitment of URMs, programs need to attend to much more.

First, graduate programs need to consider *all* of the cues that are presented to potential applicants. Emerson and Murphy (2014) describe four categories of cues: (a) those that signal representation, (b) those that make identity and stereotypes salient, (c) those that signal organizational values and beliefs, and (d) those related to structures and policies. These various types of cues can provide conflicting information to applicants. For example, a photo of diverse students and a prominent diversity statement on a program's website may not signal identity safety if there is insufficient information about the support provided for URM students or if applicant emails go unanswered (e.g., Chaney et al., 2016; Gutiérrez & Saint Clair, 2018). Purdie-Vaughns et al. (2008) noted that the effects of providing inconsistent cues are unclear, but some research suggests that a single threatening cue can undermine other identity-safety cues (Wilton et al., 2019; Wout et al., 2014).

Second, it is critical that the signals communicated by graduate programs are authentic and accurately portray the organization because "counterfeit diversity" can hurt recruitment efforts (Kroeper et al., 2020). Moreover, graduate programs need to signal that their stated support for diversity is actually enacted. Wilton et al. (2019) noted that expressed prodiversity views (e.g., a diversity statement) are likely to be viewed with skepticism if they are not also supported by evidence-based diversity cues (e.g., data on demographics, objective support for URM). Note that this does not mean that an organization with low representation cannot signal a welcoming environment—Purdie-Vaughns et al. (2008) demonstrated that individuals do search for other cues to disambiguate the meaning of low URM representation in an organization.

Conclusion

In sum, we agree with many of the points made by Woo et al. about the graduate-school admissions process. Past research has provided support for the predictive validity of the GRE and "very limited evidence for psychometric bias." Nevertheless, we also recognize the limitations of the GRE and emphasize that fundamental changes need to occur at the broader societal level to address the subgroup differences that have been observed on cognitive assessments. With this in mind, the goal of this commentary was to describe several alternative predictors and approaches to selection that have substantial empirical support for their use. These alternatives, including measuring conscientiousness and vocational interests and recruiting qualified URM applicants, are not proposed as replacements for the GRE. In fact, given the low correlations between conscientiousness, vocational interests, and GRE scores, these alternative predictors cannot make up for the reduced validity that would occur if the GRE were to be eliminated from the admissions process. Instead, our goal was to describe approaches that can supplement the GRE, increase the predictive validity of the admissions process, and promote diversity in graduate programs. Although these approaches can help to address the diversity of a particular program or institution, they will not address the broader disparities that still influence scores on the GRE, access to relevant research experiences, and other inequities that affect the diversity of graduate-school applicant pools. Nevertheless, these changes can be implemented immediately as we continue to focus on the broader societal changes that need to occur, particularly regarding systemic racism.

Transparency

Action Editor: Matthew Rhodes
Editor: Laura A. King
Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Note

1. For all values of Cohen's *d* presented in this article, the direction of the effect is scaled such that positive values indicate that the majority group members (i.e., men or White samples) tend to score higher than members of a minority group (e.g., women or Black samples). Negative values indicate that minority group members tend to score higher.

References

Avery, D. R. (2003). Reactions to diversity in recruitment advertising-Are differences Black and White? *Journal of Applied Psychology*, *88*(4), 672–679. https://doi.org/10.1037/0021-9010.88.4.672

Avery, D. R., Hernandez, M., & Hebl, M. R. (2004). Who's watching the race? Racial salience in recruitment advertising. *Journal of Applied Social Psychology*, *34*(1), 146–161. https://doi.org/10.1111/j.1559-1816.2004.tb02541.x

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

Chaney, K. E., Sanchez, D. T., & Remedios, J. D. (2016). Organizational identity safety cue transfers. *Personality and Social Psychology Bulletin*, *42*(11), 1564–1576. https://doi.org/10.1177/0146167216665096

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cullen, M. J., Sackett, P. R., & Lievens, F. (2006). Threats to the operational use of situational judgment tests in the college admission process. *International Journal of Selection and Assessment*, *14*, 142–155. https://doi.org/10.1111/j.1468-2389.2006.00340.x

De Corte, W., Lievens, F., & Sackett, P. R. (2007). Combining predictors to achieve optimal trade-offs between selection quality and adverse impact. *Journal of Applied Psychology*, *92*(5), 1380–1393. https://doi.org/10.1037/0021-9010.92.5.1380

Drasgow, F., Nye, C. D., Guo, J., & Tay, L. (2009). Cheating on proctored tests: The other side of the unproctored debate. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *2*, 46–48. https://doi.org/10.1111/j.1754-9434.2008.01106.x

Emerson, K. T. U., & Murphy, M. C. (2014). Identity threat at work: How social identity threat and situational cues contribute to racial and ethnic disparities in the workplace. *Cultural Diversity and Ethnic Minority Psychology*, *20*(4), 508–520. https://doi.org/10.1037/a0035403

Finch, D. M., Edwards, B. D., & Wallace, J. C. (2009). Multistage selection strategies: Simulating the effects on adverse impact and expected performance for various predictor combinations. *Journal of Applied Psychology*, *94*(2), 318–340. https://doi.org/10.1037/a0013775

Gutiérrez, A. S., & Saint Clair, J. K. (2018). Do organizations' diversity signals threaten members of the majority group? The case of employee professional networks. *Journal of Business Research*, *89*, 110–120. https://doi.org/10.1016/j.jbusres.2018.04.003

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, *9*(1-2), 152–194. https://doi.org/10.1111/1468-2389.00171

Jones, K. S., Newman, D. A., Su, R., & Rounds, J. (2021). Black-White differences in vocational interests: Meta-analysis and boundary conditions. *Journal of Business and Psychology*, *36*, 589–607. https://doi.org/10.1007/s10869-020-09693-5

Keiser, H. N., Sackett, P. R., Kuncel, N. R., & Brothen, T. (2016). Why women perform better in college than admission scores would predict: Exploring the roles of conscientiousness and course-taking patterns. *Journal of Applied Psychology*, *101*, 569–581. https://doi.org/10.1037/apl0000069

Kim, S. S., & Gelfand, M. J. (2003). The influence of ethnic identity on perceptions of organizational recruitment. *Journal of Vocational Behavior*, *63*(3), 396–416. https://doi.org/10.1016/S0001-8791(02)00043-X

Kroeper, K. M., Williams, H. E., & Murphy, M. C. (2022). Counterfeit diversity: How strategically misrepresenting gender diversity dampens organizations' perceived sincerity and elevates women's identity threat concerns. *Journal of Personality and Social Psychology*, *122*(3), 399–426. https://doi.org/10.1037/pspi0000348

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, *60*(3), 683–729. https://doi.org/10.1111/j.1744-6570.2007.00089.x

National Academies of Sciences, Engineering, and Medicine. (2017). *Supporting students' college success: The role of assessment of intrapersonal and interpersonal competencies*. The National Academies Press.

Newman, D. A., & Lyon, J. S. (2009). Recruitment efforts to reduce adverse impact: Targeted recruiting for personality, cognitive ability, and diversity. *Journal of Applied Psychology*, *94*(2), 298–317. https://doi.org/10.1037/a0013472

Nye, C. D. (2022). Assessing interests in the twenty-first century workforce: Building on a century of interest measurement. *Annual Review of Organizational Psychology and Organizational Behavior*, *9*, 415–440. https://doi.org/10.1146/annurev-orgpsych-012420-083120

Nye, C. D., Butt, S. M., Bradburn, J., & Prasad, J. (2018). Interests as predictors of performance: An omitted and underappreciated variable. *Journal of Vocational Behavior*, *108*, 178–189. https://doi.org/10.1016/j.jvb.2018.08.003

Nye, C. D., Perlus, J. G., & Rounds, J. (2018). Do ornithologists flock together? Examining the homogeneity of interests in occupations. *Journal of Vocational Behavior*, *107*, 195–208. https://doi.org/10.1016/j.jvb.2018.04.004

Nye, C. D., Prasad, J., & Rounds, J. (2021). The effects of vocational interests on motivation, satisfaction, and academic performance: Test of a mediated model. *Journal of Vocational Behavior*, *127*, Article 103583. https://doi.org/10.1016/j.jvb.2021.103583

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives of Psychological Science*, *7*, 384–403. https://doi.org/10.1177/1745691612449021

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338. https://doi.org/10.1037/a0014996

Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Ditlmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology*, *94*(4), 615–630. https://doi.org/10.1037/0022-3514.94.4.615

Roth, P. L., Van Iddekinge, C. H., DeOrtentiis, P. S., Hackney, K. J., Zhang, L., & Buster, M. A. (2017). Hispanic and Asian performance on selection procedures: A narrative and meta-analytic review of 12 common predictors. *Journal of Applied Psychology*, *102*(8), 1178–1202. https://doi.org/10.1037/apl0000195

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, *50*(3), 707–721. https://doi.org/10.1111/j.1744-6570.1997.tb00711.x

Sackett, P. R., Laczo, R. M., & Lippe, Z. P. (2003). Differential prediction and the use of multiple predictors: The omitted variables problem. *Journal of Applied Psychology*, *88*(6), 1046–1056. https://doi.org/10.1037/0021-9010.88.6.1046

Sackett, P. R., & Roth, L. (1996). Multi-stage selection strategies: A Monte Carlo investigation of effects on performance and minority hiring. *Personnel Psychology*, *49*(3), 549–572. https://doi.org/10.1111/j.1744-6570.1996.tb01584.x

Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2021). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*. Advance online publication. https://doi.org/10.1037/apl0000994

Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, *15*(1-2), 187–210. https://doi.org/10.1207/S15327043HUP1501&02_12

Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., & Zorzie, M. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students. *Journal of Applied Psychology*, *94*, 1479–1097. https://doi.org/10.1037/a0016810

Sinha, R., Oswald, F., Imus, A., & Schmitt, N. (2011). Criterion-focused approach to reducing adverse impact in college admissions. *Applied Measurement in Education*, *24*, 137–161. https://doi.org/10.1080/08957347.2011.554605

Song, Q. C., Wee, S., & Newman, D. A. (2017). Diversity shrinkage: Cross-validating pareto-optimal weights to enhance diversity via hiring practices. *Journal of Applied Psychology*, *102*(12), 1636–1657. https://doi.org/10.1037/apl0000240

Stark, S., Chernyshenko, O. S., Drasgow, F., Nye, C. D., White, L. A., Heffner, T., & Farmer, W. L. (2014). From ABLE to TAPAS: A new generation of personality tests to support military selection and classification decisions. *Military Psychology*, *26*(3), 153–164. https://doi.org/10.1037/mil0000044

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, *135*, 859–884. https://doi.org/10.1037/a0017364

Trent, J. D., Barron, L. G., Rose, M. R., & Carretta, T. R. (2020). Tailored Adaptive Personality Assessment System (TAPAS) as an indicator for counterproductive work behavior: Comparing validity in applicant, honest, and directed faking conditions. *Military Psychology*, *32*(1), 51–59. https://doi.org/10.1080/08995605.2019.1652481

Van Iddekinge, C. H., Putka, D. J., & Campbell, J. P. (2011). Reconsidering vocational interests for personnel selection: The validity of an interest-based selection test in relation to job knowledge, job performance, and continuance intentions. *Journal of Applied Psychology*, *96*, 13–33. https://doi.org/10.1037/a0021193

Walker, H. J., Feild, H. S., Giles, W. F., Armenakis, A. A., & Bernerth, J. B. (2009). Displaying employee testimonials on recruitment web sites: Effects of communication media, employee race, and job seeker race on organizational attraction and information credibility. *Journal of Applied Psychology*, *94*, 1354–1364. https://doi.org/10.1037/a0014964

Wilton, L. S., Sanchez, D. T., Unzueta, M. M., Kaiser, C., & Caluori, N. (2019). In good company: When gender diversity boosts a company's reputation. *Psychology of Women Quarterly*, *43*(1), 59–72. https://doi.org/10.1177/0361684318800264

Woo, S. E., LeBreton, J., Keith, M., & Tay, L. (2022). Bias, fairness, and validity in graduate admissions: A psychometric perspective. *Perspectives on Psychological Science*, *17*(♦), ♦♦♦. https://doi.org/10.1177/17456916211055374

Wout, D. A., Murphy, M. C., & Barnett, S. (2014). When having Black friends isn't enough: How threat cues undermine safety cues in friendship formation. *Social Psychological and Personality Science*, *5*(7), 844–851. https://doi.org/10.1177/1948550614535820

Zhang, C., & Kuncel, N. R. (2020). Moving beyond the brag sheet: A meta-analysis of biodata measures predicting student outcomes. *Educational Measurement: Issues and Practice*, *39*, 106–121. https://doi.org/10.1111/emip.12313

Zickar, M. J., & Robie, C. (1999). Modeling faking good on personality items: An item-level analysis. *Journal of Applied Psychology*, *84*, 551–563. https://doi.org/10.1037/0021-9010.84.4.551