11-1-2012

An Assessment of Procedural Justice: Do Perceptions of Job Relatedness Alter Criterion-Related Validities?

Dana M. Grambow
University of Missouri-St. Louis

An Assessment of Procedural Justice: Do Perceptions of Job Relatedness Alter Criterion-Related Validities?

Dana M. Grambow
M.A., Industrial & Organizational Psychology, University of North Carolina at Charlotte, 2005
B.S., Psychology, Campbell University, 2000

A dissertation submitted to the Graduate School at the University of Missouri-St. Louis in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychology with an emphasis in Industrial and Organizational

October 2012

Advisory Committee

Therese Macan, Ph.D., Chairperson

James Breaugh, Ph.D.

Alice Hall, Ph.D.

John Meriac, Ph.D.

An Assessment of Procedural Justice:

Do Perceptions of Job Relatedness Alter Criterion-Related Validities?

Dana M. Grambow

University of Missouri – St. Louis

Abstract

Two studies examined an application of procedural justice to the field of human resources by investigating participant reactions to different selection assessments administered in the same session. Following Gilliland's model of organizational justice applied to human resources, the first study investigated how reactions to a cognitive ability assessment and a measure of Conscientiousness may alter the participants' test-taking motivation for these assessments. Extending beyond Gilliland's proposals, the first study assessed if these changes in test-taking motivation ultimately altered the validity of the assessments and tested a model depicting changes in the effect on reactions when two assessments are taken together. Using 347 undergraduate students enrolled in psychology and business classes from two mid-sized universities, the validities were not significantly altered based on the different assessment orders. The second study attempted to determine if the validity of assessments could be maximized by administering the assessments with a business frame-of-reference and by administering them in a specific order. Using 176 participants similar to those in the first study, the results indicated that using a contextualized measure of Conscientiousness followed by a business related cognitive ability assessment did not maximize the assessments' validities. Despite the general lack of support for the relationships assessed by these studies, exploratory analyses indicate that future research is warranted before ruling out the impact of perceptions of procedural justice on the validities of assessments administered in the same session.

An Assessment of Procedural Justice:

Do Perceptions of Job Relatedness Alter Criterion-Related Validities?

Organizational justice literature originally focused on how violations of the justice rules could impact employees within an organization (Greenberg, 1990). This literature has been expanded to include applicants' perceptions of justice based on the organizations' selection system (Gilliland, 1993). While Gilliland's seminal article has been the impetus for research addressing the impact of justice in the selection process, there are still questions that remain unanswered. There has been a call for research that addresses immediate individual outcomes of applicants' perceptions of justice (Ryan & Ployhart, 2000; Chan & Schmitt, 2004; Truxillo, Steiner, & Gilliland, 2004) as well as how applicant reactions can impact the selection system itself (Schmitt & Chan, 1999). Through the use of two studies, the present paper extends the research of others who have begun to address these issues by investigating the effects of procedural justice perceptions on the test-taking motivation when two assessments are given in the same session. Test-taking motivation was chosen as the individual outcome variable of interest due to its potential impact on the validity of the assessment being used (Schmit & Ryan, 1992). The present studies have three purposes: (1) to extend the knowledge base beyond the existing literature by assessing any unintended consequences on criterion-related validity of typical organizational selection practices, (2) by assessing how those consequences impact the selection system as a whole and, (3) if present, to determine if those consequences can be avoided.

Establishment of Organizational Justice Constructs

Early justice literature primarily addressed distributive justice, with roots in equity theory, and was largely focused on outcomes such as pay. The focus of research then shifted to include procedural justice and the implications of fairness perceptions for processes such as performance appraisals (Greenberg, 1990). In his review of the justice research, Greenberg made distinctions in how to address procedural justice. His review revealed three characteristics of procedural justice: formal characteristics of the process, explanations of the process, and interpersonal treatment. Perceptions of fairness hinged on the actual process itself, whether the process was explained in a sincere manner with a logical reason, and whether or not the individual was treated honestly and with courtesy by others involved in the process.

Since Greenberg's (1990) review, researchers have reconsidered the types of justice that may play a role in organizations. Initially, researchers pushed to include a third form of justice. Interactional justice, previously included in procedural justice, is said to reflect the fairness in communication by demonstrating elements such as honesty and respect (Cohen-Charash & Spector, 2001). Other researchers have argued that organizational justice is comprised of four factors formed by splitting interactional justice into interpersonal justice and informational justice (Colquitt, Conlon, Wesson, Porter, & Ng, 2001). It should be noted that the aforementioned changes in the perception of organizational justice did not introduce any new rules of justice. As there have been no new rules of organizational justice introduced with the emergence of interactional, interpersonal, or informational justice and because scholars have not been able to reach

an agreement on the number of factors of organizational justice, the present studies rely on Greenberg's (1990) two factor delineation of distributive and procedural justice.

Organizational Justice Applied to Selection Systems

Using Greenberg's (1990) review as a foundation, Gilliland (1993) presented a model whereby the distributive and procedural justice concepts were extended beyond the boundaries of an organization's existing structure (e.g., performance appraisals) by applying these constructs to the selection system of an organization. In his proposed model, outcomes such as hiring decisions and discrimination were seen as either conforming to the distributive justice rules or as violating those rules. The amount of compliance with the distributive justice rules was then said to impact fairness perceptions of the selection process outcomes. Ultimately, these fairness perceptions were proposed to affect applicant reaction outcomes during and after the selection process along with outcomes related to the applicants' self perceptions. Gilliland's model followed a similar path for procedural justice as he proposed that the determinants of procedural justice rules were: 1) the type of test(s) used in the selection process, 2) the policies of Human Resources and the personnel within Human Resources which impact formal characteristics, 3) explanation of the process, and 4) interpersonal treatment, respectively. While Gilliland specifically addressed how Human Resource personnel could impact the interpersonal treatment component of procedural justice, it stands to reason that any representative of the organization involved in administering the selection assessment(s) could have a similar impact. A review of the components of interpersonal treatment outlined by Gilliland (i.e., interpersonal effectiveness of the administrator, two-way communication, and propriety of questions) reveals how this aspect of procedural justice

need not be limited to Human Resource personnel. Additionally, Gilliland noted that the distributive justice rules may moderate the relationship between the procedural justice rules and the applicant's perceptions of overall fairness. He also suggested that the procedural justice rules could impact distributive justice in the same way. The moderation was thought to stem from violations of some aspects of justice altering the salience of other justice aspects. For example, if a person is hired for a job (distributive justice), the procedural justice rules will likely have less impact on whether or not they see the overall process as fair than if they had not been hired for the job.

While Gilliland (1993) did not test his proposed model in this seminal piece, he did lay the foundation for future research by listing specific examples of each of the components of the procedural justice rules (see Figure 1). The overall model also included potential moderators such as the applicants' previous experiences with selection processes and the stage of the selection process they were in when a rule violation occurred. Gilliland's (1993) theory was heavily focused on the impact of procedural justice rules and that was evidenced in his specific propositions. He proposed that different types of selection tests would result in differing perceptions of job relatedness. While Gilliland proposed differences based on test type, he also proposed that differences in tests would have the greatest impact on outcomes when a procedural rule was violated (e.g., when a selection test is seen as being unrelated to the job). This sentiment was echoed by Truxillo, Steiner, and Gilliland (2004) who proposed that there are boundary conditions that determine when perceptions of fairness matter with regards to outcomes. One of the proposed boundary conditions stated that fairness would only impact an outcome if the process was perceived to be unfair, as opposed to being seen as less fair

relative to another procedure. Gilliland (1993) suggested that when procedures were seen as unfair, test motivation would be lower, job acceptance and application recommendations would decrease, and legal battles would increase.

Tests of Gilliland's Model

Gilliland's (1993) model is so complex that it would be difficult, if not impossible, to test empirically the entire proposed model. However, it is possible to examine one "branch" of the model from the characteristics of the selection system through to the applicant reactions. One branch of Gilliland's model that has been tested, and is the focus of the present studies is the relationship between reactions to an assessment based on the procedural justice rule of job relatedness, fairness perceptions, and participant reactions in the form of increased/decreased test-taking motivation. This paper replicates the current research of this branch and extends it to assess the effect on criterion-related validity and the impact of using multiple selection assessments in a single testing session (see Figure 2). The present paper focuses on two selection assessments commonly used by organizations, cognitive ability and personality assessments. Job relatedness is the procedural justice rule that is the focus of the present studies. This particular rule was chosen not only to be able to extend the existing research of Gilliland's model, but also because Gilliland suggested that the rule of job relatedness may have a stronger impact on perceptions of fairness than any of the other procedural justice rules. Test-taking motivation is an outcome that is also consistent with previous research on Gilliland's model; however, the reason for including test-taking motivation in the present studies is that it may be seen as one of the most important outcomes given its potential to alter the validity of a selection assessment (Schmit & Ryan, 1992).

Gilliland (1994) offers the first evidence to support some of the propositions presented in his model related to the present studies (see Figure 2) by assessing the relationship between the procedural justice rule of job relatedness and perceived fairness and organizational outcomes. Using undergraduate students who were applying for a clerical position, Gilliland (1994) assessed the impact of procedural justice measured by job relatedness and the explanation offered on various outcome variables such as: perceived distributive fairness, perceived procedural fairness, recommendation intentions, self-efficacy, and job performance. Having the students actually apply for a clerical position presumably provided motivation for participants to perform as an applicant to a job would perform since those who were hired were paid for hours worked. The students were given a cognitive ability test, a work sample test, and an overt integrity test. While there was no relationship between perceptions of procedural justice and job performance as proposed in his model, the proposed relationship between applicant perceptions and self-efficacy were supported. Specifically, Gilliland found that assessments that were job related were positively related to self-efficacy for the participants who were accepted for the position and negatively related to self-efficacy for participants who were rejected.

As progress was made in the attempt to test empirically aspects of Gilliland's (1993) model, some researchers began a line of research designed to provide further support for the model's propositions related to the focus of the present paper. One research study addressed the relationship between job relatedness and test performance. Chan (1997) used a sample of undergraduate students to test the effect of participant reactions to a cognitive ability test along with a personality test. The students were given the Wonderlic cognitive ability assessment and the NEO-FFI personality inventory. After

completing the assessments the students were instructed to rate the tests on perceptions of predictive validity. Unlike the results reported by Gilliland (1994), participant's perceptions of predictive validity affected performance on the cognitive ability test supporting Gilliland's (1993) original proposition in his model. Because the main focus of Chan's (1997) study was to assess racial differences in perceptions, and because there were not expected to be any differences in perceptions across race for the personality tests, Chan did not evaluate the relationship between participants' perceptions of predictive validity and their subsequent performance on the personality assessment.

Chan, Schmitt, Sacco, and DeShon (1998) continued the evaluation of students' reactions to cognitive ability and personality tests. The NEO-FFI was used in this study as well, but the Wonderlic was replaced with a six scale measure of cognitive ability where the items were worded to appear business related. The study attempted to motivate students to perform well by offering a monetary reward to top performers. To assess initial perceptions, students were first given a set of cognitive ability items to read so they would have familiarity with the test in question, but they were specifically asked not to answer the items. Using entry level managers in a large U.S. based telecommunications company as the job reference, the participants then completed a reaction measure that assessed their initial perceptions of predictive validity, fairness, and face validity measured by the job relatedness of the test content. The full cognitive ability test was then administered followed by the same reaction measure. After a ten minute break the students completed a parallel form of the cognitive ability test. This procedure was then repeated using a personality test. The results of Chan's (1997) study were replicated in that participant reactions affected performance on the cognitive ability test even when a

different assessment of cognitive ability was used. However, the study did not find the same relationship for the personality test. Rather, the study found that there was no relationship between participants' reactions to the personality test and their performance on the personality test.

In an additional study using only a cognitive ability test, Chan, Schmitt, DeShon, Clause, and Delbridge (1997) provided a full assessment of the branch from Gilliland's (1993) model that is the focus of the present paper. As in the other studies by Chan and his colleagues (1997, Chan et al., 1998), students were used as participants and monetary rewards were offered to induce motivation. The authors found that face validity of an initial assessment predicted performance on a parallel form of the cognitive ability test given in the same session. However, the results of a structural equation model revealed that this relationship was fully mediated by the students' test-taking motivation. The mediation was included in the hypothesized model due to the researchers' assumption that test-taking motivation would directly affect the participants' performance on the assessment. While the aforementioned studies have provided some answers to the questions presented in Gilliland's (1993) model, they do not address changes that may occur in the model when different types of selection assessments are administered in the same testing session. The present paper addresses this need by using two different types of selection assessments.

Test Type

Use of Cognitive Ability and Personality Assessments

While Chan et al. (1998) provide an assessment of Gilliland's model demonstrating that perceptions of job relatedness impact fairness perceptions and

ultimately test-taking motivation, there is a lack of empirical evidence demonstrating what happens to these relationships when different types of assessments are administered together. If the reactions to different assessments administered in the same session result in no differences than had they been administered alone, then organizations need not concern themselves with this factor; however, if differences do exist then organizations may want to consider these impacts of applicant reactions when designing their selection process in addition to other factors such as validity and the potential for adverse impact. Consistent with existing research the present studies focus on cognitive ability and personality assessments in order to address this deficiency in the current literature.

The decision to include a cognitive ability assessment is based on their well-established criterion-related validity. Hunter and Hunter (1984) established that cognitive ability assessments are valid predictors for most job families. In a review of nine different job families, the mean validity coefficients for cognitive ability assessments ranged from a high of .61 for salespersons to a low of .27 for sales clerks. The criterion-related validity of cognitive ability assessments was further established in Schmidt and Hunter's 1998 review that found that cognitive ability had a correlation of .51 with overall job performance and .56 with training performance. Though the aforementioned studies were limited to results found in the United States, similar results were reported for countries in the European Community. Salgado, Anderson, Moscoso, Bertua, and Fruyt (2003) conducted a meta-analysis on studies using participants from the European Community and found that general mental ability had validity coefficients with job performance that ranged from .34 for sales positions to .12 for police officers. The validity coefficients for training performance ranged from .46 for information clerk to .13 for police in the

European Community. While the results reported by Salgado et al. (2003) were markedly lower than those reported for job families in the United States, the validity coefficients still suggest that cognitive ability assessments are useful predictors for most job families though it is evident that there may be some professions (e.g., police officer and information clerk) for which the usefulness is attenuated. More recently, research has shown that these results are generalizable to populations in the United Kingdom and Germany as well. Bertua, Anderson, and Salgado (2005) conducted a meta-analysis of the relationship between cognitive ability and performance using samples from the United Kingdom in which they found the correlation between cognitive ability and job performance to be .22 while the correlation between cognitive ability and training performance was .29. A meta-analysis using German samples to assess these constructs was conducted as well and the correlation of general mental ability with job performance and training success were .33 and .31 respectively (Hülsheger, Maier, & Stumpp, 2007). The results of the aforementioned meta-analyses demonstrate the robust nature of the relationship between cognitive ability and performance, making it a desirable component of a selection system.

Personality assessments do not share the same criterion-related validity estimates as cognitive ability assessments, yet their value to practitioners is still evident. In response to a renewed interest in the use of personality assessments in the selection process, Hurtz and Donovan (2000) sought to address limitations in previous studies (e.g., Barrick & Mount, 1991; Salgado, 1997; Tett, Jackson, & Rothstein, 1991). Hurtz and Donovan concurred with previous findings that Conscientiousness had the highest validity coefficient from the five factor model of personality. However, with a validity

coefficient of .15 for job performance and .02 for training performance the usefulness of Conscientiousness as part of a selection battery was questioned. Similar results were reported in a meta-analysis including studies from the United Sates and the European Community where the validity coefficient of Conscientiousness was reported to be .17 for job performance (Salgado, 2003). The lower validities found may be due to respondents answering with different frame-of-references (Schmit, Ryan, Stierwalt, & Powell, 1995). Despite these dismal validity coefficients, personality assessments may be useful in other ways.

Schmidt and Hunter (1998) reported that by combining a measure of Conscientiousness with a measure of cognitive ability the criterion-related validity for job performance increased 18% over that of the cognitive ability measure alone. Though some researchers have questioned the usefulness of personality assessments in selection, even the skeptics concur that by using a personality test along with a measure of cognitive ability organizations will likely achieve validities higher than could be attained by using either measure alone (Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). The amount of incremental validity of Conscientiousness over that of cognitive ability has been shown to be even greater when the criterion is contextual performance as opposed to task performance. Contextual performance has been described as extra role behaviors that support an environment in which the tasks can be carried out and transcend across job titles (Borman & Motowildo, 1993). Avis, Kudisch, and Fortunato (2002) used cognitive ability and Conscientiousness scores from incumbent cashiers in a retail company in their assessment of incremental validity. The cashiers' supervisors then completed a rating of composite performance for these employees. A

factor analysis demonstrated that contextual performance was included in the measure of composite performance. The 15 items assessing contextual performance had a correlation of .32 with Conscientiousness whereas task performance and Conscientiousness were correlated at .20. When assessing the incremental validity of Conscientiousness over that of cognitive ability, the researchers found that the Multiple *R* for task performance and contextual performance were .24 and .32 respectively. This resulted in a 41% increase in Multiple *R* for task performance and a 167% increase in Multiple *R* for contextual performance when Conscientiousness was included with cognitive ability scores.

In addition to the incremental validity demonstrated by using a personality assessment in conjunction with a cognitive ability measure, this combination also has the potential to decrease adverse impact resulting from the cognitive ability measure. Potosky, Bobko, and Roth (2005) estimated that by administering a personality assessment in conjunction with a cognitive ability assessment, the adverse impact potential could be lowered by five percent. While this estimate is not as desirable as organizations may prefer, combined with results that have shown a potential increase in validity, there is evidence that organizations could benefit by using both cognitive ability and personality assessments. It is also important to note that the use of personality assessments as part of the selection process is wide spread in organizations and appears to be steadily increasing (Rothstein & Goffin, 2006). Due to the consistent findings for Conscientiousness over that of the other Big Five personality variables, the present studies focus on Conscientiousness for the personality assessment.

In addition to the arguments set forth above, the focus of the present studies on cognitive ability and Conscientiousness serves to advance research on applicant reactions

by allowing for comparison to the extant literature. Many of the studies stemming from Gilliland's (1993) model of procedural justice in human resources have focused on the use of cognitive ability assessments alone (Gilliland, 1994; Chan et al., 1997) or cognitive ability assessments along with personality assessments that include a measure of Conscientiousness (Chan, 1997; Chan et al., 1998).

Applicant Reactions to Cognitive Ability and Personality Assessments

With regard to cognitive ability and Conscientiousness, another important element to examine is the literature addressing applicants' reactions to these assessments. Without varying reactions, one would not anticipate different results from administering the assessments together, versus separately, so in order to extend the research in this area one must use measures that result in different reactions. Keeping in mind Gilliland's (1993) assertion that the greatest effects of procedural justice are felt when a rule of justice is violated, care needs to be taken in selecting assessments that are sufficiently different from one another to elicit different responses from the participants. While Gilliland's claim leaves an impression that an assessment would have to be seen as completely unrelated to the job to result in differences great enough to impact the outcome measures of his model, there is some evidence from other areas of procedural justice that the impact can be felt even when there are only moderate differences. Ployhart and Ryan (1998) manipulated a testing environment so that participants felt either a positive inconsistency (receiving more time to complete a test), negative inconsistency (receiving less time to complete a test), or a consistent situation. Of interest to Gilliland's claim regarding rule violations is that the researchers found those in the positive inconsistency condition (fairness $M = 3.97$) and those in the negative inconsistency condition (fairness

M = 3.05) reported statistically significant differences in their intent to participate in similar selection practices in the future while those in the consistent condition who experienced no rule violation (fairness *M* = 3.88) did not have significantly different results from either inconsistency condition on intent to participate in similar selection practices in the future. It should be noted that the violation of the rule did not result in significantly different outcomes, and while there were significant differences between fairness ratings in the positive inconsistency and negative inconsistency conditions both of the ratings were above 3.00 which on a five-point scale would have been "neither agree nor disagree" or a "neutral" response. Yet these conditions still induced different reactions to the outcome variable specified above. Another study, looking at subgroup differences based on type of test, found lower subgroup differences on a constructed response exam compared to a multiple choice exam and these differences were attributed to more favorable ratings on the constructed response exam for job relatedness (*M* = 2.51 vs. 2.40), fairness (*M* = 3.46 vs. 3.09), and test-taking motivation (*M* = 4.17 vs. 3.96) for the African American participants (Edwards & Arthur, 2007). The results presented in both of the aforementioned studies provide a gauge by which one could determine if the differences in applicant reactions between a cognitive ability assessment and a Conscientiousness assessment are large enough to result in different outcomes even if neither of the two assessments is seen as completely unrelated to the job.

Despite the importance of understanding applicant reactions, the extant literature is not consistent in how the types of applicant reactions are defined. It is therefore necessary to first clarify the procedural justice rule of job relatedness as it pertains to the present studies.

Gilliland viewed job relatedness as perceptions of content validity (i.e., does the content of the test appear to be related to the content of the job) and perceptions of criterion-related validity (i.e., does the test appear to measure characteristics necessary to perform the job well), while face validity was determined by whether or not it appeared that the test measured what it is actually measuring (i.e., does a cognitive ability test look like it is measuring cognitive ability). While Gilliland made a distinction between job relatedness and face validity, there is less clarity between the two constructs within the literature. Table 1 provides the items of face validity used by Smither, Reilly, Millsap, Pearlman, and Stoffey (1993). The items of the job-relatedness content scale developed by Bauer, Truxillo, Sanchez, Ferrara, and Campion (2001) that was based on Gilliland's (1993) model are also presented in Table 1. As demonstrated by the overlap of the content in the items presented in this table (the first and second items of each scale), Smither et al. defined face validity as the extent to which the content of the test is seen as job related. Despite Gilliland's assertion that these terms are not equivalent, the operational definitions of the terms, job related and face valid, frequently overlap. The overlap of these definitions may be that in a selection context, the assessment used should be measuring characteristics relevant to the job. Thus, if the measure meets Gilliland's definition of face validity it should also meet his requirements for job relatedness as the content of the assessment should appear related to the content of the job and the assessment should appear to measure characteristics that would lead to successful job performance. Therefore, the literature reviewed below will include research assessing Gilliland's (1993) definition of job relatedness regardless of the terminology used.

Job Relatedness of Cognitive Ability and Personality Assessments

Existing research examining participants' perceptions of job relatedness for cognitive ability assessments and personality assessments have shown that not only are these assessments viewed differently, but also that cognitive ability tests are consistently rated higher on scales of job relatedness than personality assessments. In a meta-analysis Hausknecht, Day, and Thomas (2004) reviewed primary studies assessing the favorability (e.g., job relatedness) of cognitive ability and personality assessments. As the scales used in the primary studies varied, the researchers recalibrated the results so that all primary studies were on a five-point scale. They found that the sample weighted favorability mean for cognitive ability tests (M = 3.14, SD = 1.00) was higher than for personality assessments (M = 2.88, SD = .99). Additionally, the 95% confidence intervals for these ratings do not overlap indicating that there is a significant statistical difference. Because the scales of the primary studies were adjusted to reflect a five-point scale, there are no anchors tied to these means. However, typically on a five-point scale one would expect that a mean above 3.0, the mid-point, would indicate that the assessment was seen as job related while a mean below 3.0 would indicate that the assessment was not seen as job related. Thus, the results of the meta-analysis could suggest that cognitive ability assessments and measures of Conscientiousness are useful assessments for the present studies given Gilliland's (1993) statement that violations of the procedural justice rule (i.e., job relatedness) will have a greater impact on outcomes (i.e., fairness perceptions) than if one assessment was simply rated as less job related than the other, especially considering the aforementioned results of Edwards and Arthur (2007) regarding participant reactions and subgroup differences.

While the meta-analysis conducted by Hausknecht et al. (2004) provides a good picture of the differences in job relatedness on cognitive ability assessments and Conscientiousness measures, one should still consider the primary studies that were included in the meta-analysis. Using managers from a large corporation, Smither et al. (1993) found that out of eight assessments of cognitive ability, only two of the assessments were not rated significantly higher on job relatedness than the personality assessment. Having provided the managers with descriptions of the assessments, which most participants had prior exposure to, and asking the participants to rate the assessments using a five-point scale, only quantitative comparisons and lettersets did not have significantly higher means than the personality assessment (M = 3.02, 2.99, and 2.91 respectively). The results of Smither et al.'s (1993) study hold in samples outside of the United States as well. Steiner and Gilliland (1996) asked undergraduate students in both the United States and in France to think about a job they would like to have when they completed their education. The participants were then presented with descriptions of 10 different selection assessments and asked to complete a questionnaire assessing the favorability of each of the assessments using a seven-point scale. Included in the favorability scale was an item assessing face validity. The results from participants in both the United States and France indicated that written ability tests (described as "Paper-and-pencil tests that evaluate your intelligence on your reasoning, verbal, or mathematical skill," p. 136) were perceived as being more job related than personality tests. The United States sample provided a mean rating of 4.80 for the written ability tests and 3.80 for the personality tests, while the differences in the respective mean ratings for the French sample were not as drastic at 4.40 and 4.20. The procedures used by Steiner

and Gilliland (1996) were later replicated with undergraduate students from Spain and Portugal. Moscoso and Salgado (2004) found that participants in Spain and Portugal also provided higher mean ratings of face validity on the written ability test than on the personality assessment. Written ability tests received a mean rating of 5.13 from the Spaniards and 4.62 from the Portuguese, while the respective ratings for the personality tests were 4.76 and 4.31.

Each of the primary studies reviewed above relied on ratings of job relatedness obtained from participants who simply read descriptions of the various assessments. While some of the participants may have encountered similar selection assessments before (e.g., Smither et al., 1993) other participants may have had no knowledge of the items from the assessments. It is conceivable that different reactions may occur when one simply reads a description of an assessment than when one encounters actual items from the assessment. If participants perceive cognitive ability assessments as more job related than personality assessments after simply reading sample items, it stands to reason that these reactions may be intensified when the applicants have repeated exposure to these types of items as they complete the actual assessment.

Outcomes of Applicant Reactions

Fairness Perceptions

Despite its depiction in Gilliland's (1993) model, the literature has provided little indication that perceptions of an assessment's job relatedness relate to perceptions of fairness in the selection process, the second link in the present study (Figure 2). Perhaps this lack of empirical testing stems from Gilliland not including a direct proposition related to this relationship in his seminal piece. Another explanation is that the terms,

organizational justice and fairness, have been used synonymously by some researchers. Truxillo, Bauer, and Sanchez (2001) demonstrated this tendency by using a measure of procedural justice to assess perceived fairness of selection assessments. Truxillo et al. also included a measure of overall fairness in their study, however, this measure included questions designed to assess the overall testing process and therefore does not distinguish between the two tests that were given during the testing process. This lack of distinction in their overall fairness measure prohibits drawing conclusions as to the relationship between perceptions of job relatedness (included in the measure of procedural justice) for an assessment and subsequent fairness perceptions of that same assessment.

Chan et al. (1998) provided some evidence of a relationship between perceptions of job relatedness and perceptions of fairness. Separate scales of face validity perceptions, predictive validity perceptions, and fairness perceptions were given to participants as part of pre-test and post-test measures of applicant reactions to both a cognitive ability test and a Conscientiousness assessment. For both the cognitive ability and the Conscientiousness assessments, the perceptions of face validity and fairness displayed correlations of .37 and higher when the reactions were measured at the same time (e.g., pre-test or post-test). The relationships between perceived predictive validity and fairness were of a larger magnitude (r = .40 - .64). Table 2 provides all of the correlations between perceptions of job relatedness and perceptions of fairness. Additionally, Edwards and Arthur (2007) found the correlation between job relatedness and fairness to be .27 for their multiple choice test and .32 for their constructed response test. It should also be noted that these correlations are not as high as to suggest that face validity perceptions and fairness perceptions are the same construct, therefore there is

still a need to assess both of these constructs when testing this portion of the model as proposed by Gilliland (1993).

Test-taking Motivation

The third, and final, relationship in the portion of Gilliland's (1993) model addressed in the present studies (see Figure 2) is that applicants' fairness perceptions will then impact their test-taking motivation on the selection assessment(s). Only a single study has been found to assess the relationship between fairness and test-taking motivation. Edwards and Arthur (2007) reported correlations between these two constructs to be .26 for their multiple choice assessment and .25 for their constructed response assessment. While most researchers have not addressed this relationship, perhaps for the same reasons many have not assessed the aforementioned relationship between job relatedness and fairness perceptions, the literature has provided additional indirect evidence of this relationship. There is evidence that perceptions of procedural justice can affect test performance (Macan, Avedon, Paese, & Smith, 1994; Chan et al., 1998) and it is suggested that the observed differences in performance may be due to test-taking motivation that has gone unmeasured (Bell, Ryan, & Weichmann, 2004). While these relationships have only been demonstrated for cognitive ability tests, it seems to reason that Gilliland's (1993) proposition would extend to Conscientiousness measures evidenced in the 2004 meta-analysis conducted by Hausknecht et al. where they found that perceptions of face validity correlated with test motivation (r = .31).

Extension of Gilliland's (1993) Model – The Present Study

Applicant Reactions Influencing Subsequent Assessments

The present study replicates the aspects of Gilliland's (1993) model outlined above. Ryan and Huth (2008) critiqued the literature on applicant reactions to date and called attention to the aspects of this literature base that need further examination. They indicated that researchers studying perceptions of job relatedness need to move beyond having participants simply look at a description of a test and provide a rating. One specific area noted in this review is a need to understand how perceptions of job relatedness can impact, or alter, other perceptions. By replicating the compilation of extant research on Gilliland's model, the present study addresses this need in the research by investigating how perceptions of job relatedness can alter perceptions of fairness and indirectly alter test-taking motivation.

In addition to testing Gilliland's (1993) model, which suggests that different selection assessments lead to different perceptions of job relatedness resulting in differing perceptions of fairness and ultimately differing levels of test-taking motivation, scientist and practitioners would be well served to understand how applicant reactions can extend beyond what is proposed in Gilliland's model. Of particular importance is to understand how applicants' reactions may influence test-taking motivation when applicants are given multiple assessments in a single testing session. The option of testing applicants on multiple assessments in a single session may be of interest to practitioners due to the cost and time associated with administering selection assessments. Additionally, the literature suggests benefits to administering assessments in the same session. Rosse, Miller, and Stecher (1994) found that applicants who received a cognitive ability test in conjunction

with a personality test and unstructured interview rated the overall selection process as more favorable on aspects of appropriateness than did applicants who only had the personality test and unstructured interview. Rosse et al. suggested that practitioners could benefit by combining these assessments in their selection process. Cropanzano and Wright (2003) echo this sentiment in their review of existing research on procedural justice and human resources. Ryan and Huth (2008) also noted the need to investigate applicant reactions to tools used in combination with one another, as opposed to being administered alone. Specifically, they indicated a need to look at applicants' reactions based on the sequence in which the assessments are given and to consider deciding where in the selection process to use the tool based on validity evidence. The present study meets this need by not only looking at how participants react to two assessments given in the same session, but by also investigating how those reactions can impact the validity of the assessments based on their sequence in the administration of the assessments.

Chan et al. (1997) provided some evidence that participants' reactions to one assessment can carry over to a second assessment given in the same testing session. The authors used an undergraduate sample to test the relationship between performance on an initial test, perceptions of face validity for that test, test-taking motivation, and subsequent performance on a second test. The initial motivation level of the participants was manipulated by offering a cash reward to the top performers. The participants completed a cognitive ability test, reported their perceptions of face validity and their test-taking motivation, and then completed a parallel form of the cognitive ability test. In order to provide a reference for assessments of face validity, the participants were read a job description for a managerial position. Chan et al. (1997) found that test-taking

motivation mediated the relationship between perceptions of face validity for the first test and performance on the second test. The mediation was tested using structural equation modeling based on the hypothesis that test-taking motivation resulting from perceptions of face validity on one assessment would persist long enough to impact motivation on the parallel test. This relationship remained even when the authors controlled for performance on the first test (Chan et al., 1997).

While these results indicated that applicant reactions to one procedure can affect test-taking motivation as well as performance on subsequent procedures, some caution must be used when interpreting the results of the study as the authors did not assess the participants' initial levels of motivation or their test-taking motivation for the second assessment. Thus the authors were unable to control for initial levels of motivation and were not able to determine if the test-taking motivation reported lasted throughout the administration of the second assessment. It is necessary to control for initial levels of test-taking motivation to ensure that the test-taking motivation measured was a result of the perceptions of face validity as opposed to an extraneous variable. Being unable to identify the source of motivation would limit the ability of an organization to enhance test-taking motivation if they desired to do so. Also, since the results of Chan et al.'s (1997) study do not provide conclusive evidence that the test-taking motivation did indeed persist throughout the administration of the parallel test it cannot be concluded that the researchers' assumption of a direct relationship between test-taking motivation and test performance exists. Additionally, the subsequent test used in this study was a parallel form of the first assessment which is not typical for a multiple assessment approach to selection. Nonetheless, the results do suggest that initial perceptions of face

validity may impact performance on subsequent tests through their influence on test-taking motivation, and further research is needed to determine if these relationships hold when different types of selection assessments are used as opposed to parallel forms of an assessment. The present paper provides this needed research.

Additional evidence of applicant reactions affecting subsequent assessments was inadvertently provided by LaHauis, Perreault, and Ferguson (2003). They assessed the impact of explanations on perceptions of procedural justice. The researchers administered a cognitive ability test and a personality test to students. The students were placed into one of three conditions whereby they received no explanation, a specific explanation for how the tests were related to the job, or a general explanation for the tests. Participants' reactions to perceptions of fairness, measured by the Selection Procedural Justice Scale (Bauer et al., 2001), were assessed twice after they completed each assessment. The results showed that providing a general explanation for the cognitive ability test did impact perceptions of fairness, but this relationship was fully mediated by perceptions of content and predictive validity. Of particular interest given Chan et al.'s (1997) study that only used parallel forms of a cognitive ability test, is that the researchers found order effects for perceived predictive validity on the personality test. Because the researchers did not intend to examine the order effects, and in fact they counterbalanced their research design to offset any potential order effects, the direction of the relationship was not reported. While this doesn't address how the reactions relate to performance on the second assessment, it does demonstrate that reactions can carry over to affect perceptions of subsequent assessments across test types.

Drawing on the inferred relationship between applicant reactions to job relatedness for the first assessment and test-taking motivation on the second assessment (as evidenced by actual performance), it is proposed that participant reactions to the first assessment will influence their reported reactions on the subsequent assessment. Taking into consideration the existing evidence that cognitive ability measures are seen as more job related than measures of Conscientiousness (Hausknecht et al., 2004; Smither et al., 1993; Steiner & Gilliland, 1996; Moscoso & Salgado, 2004) and the demonstrated relationships between job relatedness and fairness (Chan et al., 1998; Edwards & Arthur, 2007) as well as fairness and test-taking motivation (Edwards & Arthur, 2007), it is expected that higher levels of test-taking motivation after taking a cognitive ability assessment will persist through the administration of a Conscientiousness measure and, conversely, that lower levels of test-taking motivation after taking a Conscientiousness measure will persist through the administration of a cognitive ability assessment.

Hypothesis 1a: Participants who take the cognitive ability assessment first will report higher levels of test-taking motivation on the Conscientiousness measure than those who are administered the Conscientiousness measure first.

Hypothesis 1b: Participants who take the Conscientiousness measure first will report lower levels of test-taking motivation on the cognitive ability assessment than those who are administered the cognitive ability assessment first.

Applicant Reactions Affecting Validity Coefficients

In addition to understanding how applicant reactions can impact subsequent selection assessments, existing research has indicated that test-taking motivation can influence the validity of selection assessments (the final link in Figure 2), and that this

result has different outcomes for cognitive ability assessments than for Conscientiousness measures. Schmit and Ryan (1992) found that test-taking motivation affected the validity of cognitive ability and personality tests differentially in that cognitive ability tests had higher validity coefficients if the applicant was motivated to perform well on the test while the validity of personality tests was lowered when the participants reported increased levels of motivation. To date no other studies have attempted to replicate this finding in its entirety, though a more recent study did find evidence that Conscientiousness (measured by the Six Factor Personality Questionnaire) was a valid predictor of job performance only when levels of reported test-taking motivation were low (O'Neill, Goffin, & Gellatly, 2010). Given evidence that applicants react to cognitive ability and personality assessments differently, and that this subsequently affects their test-taking motivation it is necessary to replicate the findings presented by Schmit and Ryan.

Hypothesis 2a: The higher test-taking motivation on the Conscientiousness measure reported by participants who were previously administered the cognitive ability assessment will result in lower validity coefficients for the Conscientiousness measure than the validity coefficients obtained when the Conscientiousness measure is administered first.

Hypothesis 2b: The lower test-taking motivation on the cognitive ability assessment reported by participants who were previously administered the Conscientiousness measure will result in lower validity coefficients for the cognitive ability assessment than the validity coefficients obtained when the cognitive ability assessment is administered first.

Assessing the Impact on the Selection System

The hypotheses outlined above address concerns regarding the changes in criterion-related validity of assessments when they are administered together compared to what the validity coefficient would have been had the assessments been administered alone. In addition to the theoretical benefit of understanding how these relationships could work within Gilliland's model, this is also of great practical concern as organizations often make decisions about which selection assessments to use based on validity evidence provided in technical manuals (where an assessment is tested alone) and then combine them in practice as a result of convenience. However, it is also important to consider the impact of the proposed outcomes on the selection system as a whole. Theoretically, it is important to understand the extent to which initial reactions persist once other elements of procedural justice are introduced, especially when the new elements are related to the same rule of justice already being enacted. Figure 3 depicts a proposed model of participant reactions to multiple assessments given in the same session, based on Gilliland's (1993) model. The proposed model extends beyond that of Gilliland by including an assessment of the indirect impact of participant reaction on criterion-related validity and by suggesting that fairness and test-taking motivation expressed as a result of perceived job relatedness on the first assessment (Time 1) will impact performance not only on the first assessment, but also on the second assessment through the impact of fairness and test-taking motivation expressed at Time 2. Thus, the total effect on the performance of the second test not only reflects the fairness and test-taking motivation reactions resulting from perceptions of its job relatedness, but also from the initial reactions of fairness and test-taking motivation experienced at Time 1.

Due to the additional effects on the performance of the second test, it is expected that the path coefficient from test performance at Time 2 to criterion performance will be lower than when the test was administered at Time 1.

Hypothesis 3a: Within the proposed model of participant reactions to multiple assessments in the same session, the path coefficient for the cognitive ability assessment will be largest when the cognitive ability assessment is administered first.

Hypothesis 3b: Within the proposed model of participant reactions to multiple assessments in the same session, the path coefficient for the Conscientiousness measure will be largest when the Conscientiousness measure is administered first.

In addition to the theoretical importance of understanding how initial reactions may persist throughout a testing session, the proposed model in Figure 3 could also offer some insight for practitioners. Of particular concern for practitioners is the hypothesized diminished validity of the Conscientiousness measure when it is administered after the cognitive ability assessment. Due to the fact that Conscientiousness does not share the high validity coefficients with performance that cognitive ability does, it is feasible that any decrease in the validity of Conscientiousness could eliminate the incremental validity it has over cognitive ability.

Research Question 1: Does the path coefficient from a measure of Conscientiousness to the criterion performance within the model of participant reactions to multiple assessments in the same session approach zero when the Conscientiousness measure is administered after a measure of cognitive ability?

Methods

Participants

There were 354 participants in this study. Two participants did not complete the cognitive ability measure and five participants completed the assessments out of order. These seven participants were removed from all analyses. Of the 347 remaining participants, 172 completed the Conscientiousness measure before the measure of cognitive ability (Conscientiousness First) while 175 completed the measure of cognitive ability before the Conscientiousness measure (Cognitive Ability First). Participants were students enrolled in psychology or business classes at the University of Missouri – St. Louis ($N = 323$) or psychology classes at the University of North Carolina – Charlotte ($N = 24$). The majority of the participants were female ($N = 224$) and Caucasian ($N = 202$). Participants ranged in age from 18 to 58 ($M = 25.56$, $SD = 7.62$). Table 3 contains additional demographic information for Conscientiousness First and Cognitive Ability First. The two groups did not significantly differ in demographic distributions: gender $\chi^2(1, N = 347) = .44$, $p = .51$; race $\chi^2(5, N = 346) = 4.08$, $p = .54$; age $t(345) = -.389$, $p = .70$.

Measures

Job Relatedness. Perceptions of job relatedness were measured by combining Smither et al.'s (1993) scales for face validity and perceived predictive validity, with Bauer et al.'s (2001) scales for job relatedness predictive and job relatedness content (see Appendix A). These four scales use a five-point response format ranging from strongly disagree to strongly agree. Although Bauer et al. (2001) created a measure specifically designed to assess procedural justice reactions, their measure only included four items

addressing job relatedness (two each for job relatedness content and job relatedness predictive) compared to the ten items in Smither et al.'s two scales. There is also less evidence of reliability for the scales developed by Bauer et al. than by Smither et al. Only two studies reported reliability information for the job related scales developed by Bauer et al. (Truxillo et al., 2001; Bauer, Truxillo, Paronto, Weekley, & Campion, 2004). The majority of the reliability coefficients reported for the two scales are above .85, though Bauer et al. (2004) did report a coefficient alpha of .75 for the job relatedness predictive scale. Reported reliabilities for Smither et al.'s (1993) scales for face validity and perceived predictive validity used in studies since their original publication do not fall below .82 and these measures have still been used to assess job relatedness in the literature even after the publication of Bauer et al.'s scale (Chan, 1997; Holtz et al., 2005, Edwards & Arthur, 2007). As the content of the measure created by Smither et al. and Bauer et al. do not vary much (see Table 1), the Smither et al. scales were combined with the items from Bauer et al. (2001). Coefficient alphas ranged from .88 to .92 in the present study. Given that many of the variables in this study were used to measure reactions to more than one assessment, all alphas from the present study are presented in the results section.

Fairness. Perceptions of fairness were measured by Chan et al.'s (1998) three item fairness perceptions scale (see Appendix B). This scale also uses a five-point response format ranging from strongly disagree to strongly agree. Chan et al.'s scale was chosen over Smither et al.'s (1993) two item procedural justice scale due to the higher reliability coefficients reported and because Chan et al.'s scale was adapted in part from Smither et al.'s scale. Coefficient alphas ranged from .67 to .83 in the present study.

Test-Taking Motivation. Participants' test-taking motivation was measured by the Motivation scale of the Test Attitude Survey (TAS; Arvey, Strickland, Drauden, & Martin, 1990; see Appendix C). This 10 item scale assesses the participants' desire to perform well on an assessment using a seven-point disagree - agree response format. There is evidence that it can discriminate between motivational levels as demonstrated by comparing scores across applicants and incumbents (Arvey et al., 1990). Additionally, the TAS Motivation scale has been used repeatedly in existing literature to assess test-taking motivation (Schmit & Ryan, 1992; Chan et al., 1997; Edwards & Arthur, 2007). Coefficient alphas ranged from .86 to .91 in the present study.

Perceived Performance. Participants' perceived performance on each of the assessments was measured with Sanchez, Truxillo, and Bauer's (2000) four item scale (see Appendix D). This scale uses a five-point response format ranging from strongly disagree to strongly agree. Coefficient alphas ranged from .85 to .93 in the present study.

Conscientiousness Context Scale. Participants completed a five item scale asking them to indicate the context they were thinking of when responding to the personality items. Using a five-point response format (strongly disagree to strongly agree), participants were asked to indicate the extent to which they were thinking about the following contexts while responding: at work, at school, at home, with friends, every day (see Appendix E).

Procedural Justice Scales. The following subscales from Bauer et al.'s (2001) Selection Procedural Justice Scale were added for exploratory purposes (see Appendix F); Information Known (extent to which the participant knew what to expect on the assessment), Chance to Perform (extent to which the participant could demonstrate his or

her abilities on the assessment), Consistency (extent to which the participant believed the assessment was administered consistently across participants), Treatment (extent to which the participant felt he or she was treated politely and with respect during the assessment), and Propriety of Questions (extent to which the participant believed the content of the assessment was appropriate).

Conscientiousness Measure. The Conscientiousness scale from the NEO - Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992) was used to assess participants' Conscientiousness. This scale was chosen due to its demonstrated reliability over time. Use of the NEO-FFI was also consistent with other studies in the area of applicant reactions (Chan, 1997; Chan et al., 1998). Additionally, the Conscientiousness measure of the NEO-FFI has been shown to predict task performance (Chamorro-Premuzic & Furnham, 2008) and contextual performance (LePine & Van Dyne, 2001), which is also assessed in the present study. Coefficient alphas ranged from .84 to .85 in the present study.

Cognitive Ability Measure. Cognitive ability was measured by the Watson-Glaser Critical Thinking Appraisal – Form S (Watson-Glaser Form S). This 40 item measure can be given untimed and includes five subtests: inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments. In choosing a measure of cognitive ability, measures were only considered if they had a strong history of reliability and validity. Two measures of cognitive ability met these criteria and are also used extensively in published research studies, Wonderlic Personnel Test and the Watson-Glaser Critical Thinking Appraisal. Schmidt and Hunter (2004) claim that the Wonderlic Personnel Test is the best general mental ability assessment available and that it is also

the most prevalently used measure of general mental ability. However, the Wonderlic Personnel Test is a speeded test allowing only 12 minutes in which to complete 50 items, presenting some concerns for the present study. Given that the Wonderlic is a speed test, as opposed to a power test, in which the score is calculated by the total number of items correct, it is likely to adversely impact participants for whom English is not their native language. A negative impact such as that would likely result in participants who are not native English speakers becoming outliers in the study, thereby limiting the generalizability of the results found. This is relevant as the University of Missouri – St. Louis website indicates that in the Fall semester of 2007 4.10% of the students on the main campus were international students.

There are also theoretical concerns with using a speed test for the present study. Most studies of applicant reactions today have used only sample items or descriptions in their studies, and these descriptions often lack any indication of being a timed test (e.g., Steiner & Gilliland, 1996; Moscoso & Salgado, 2004). A timed test could therefore alter participants' perceptions of job relatedness especially if they do not believe that a job would require one to work under such demanding time restrictions. A second theoretical concern is that a speed test may be seen as violating another procedural justice rule, opportunity to perform. If participants feel that the opportunity to perform is not given because they did not have enough time to attempt all of the items, it could confound the relationship between job relatedness and fairness being assessed in the present study. Due to the need of a valid and reliable test that is a power test as opposed to a speed test, the Watson-Glaser Form S was used. Coefficient alphas ranged from .71 to .75 in the present study.

Task Performance Measure. The participants' self-reported college grade point average (GPA) was used as the task performance criterion measure. The use of GPA as a criterion measure with student samples is consistent with the literature in this area. Specifically, Schmit and Ryan (1992) used college GPA as the criterion measure when they reported their finding of the differential effects of test-taking motivation on cognitive ability and personality assessments with regard to their criterion-related validities. While other studies have used actual tasks designed specifically for the criterion measure (Gilliland, 1994), use of such a task presents a theoretical concern for the present study. The present study proposes that initial reactions to the procedural justice rule of job relatedness will persist throughout the testing session and impact performance on subsequent tests through test-taking motivation. A task designed to assess performance may therefore also be impacted by the participants' initial reactions, thereby not allowing for an assessment of typical task performance. This potential carry-over effect is not likely to occur in organizations for two reasons. First, applicants to an organization are likely to encounter other aspects of organizational justice after the selection tests are administered that could overshadow their initial reactions to the procedural justice rule of job relatedness. Most notably would be that in order for an applicant's reaction to the testing process to impact their job performance they would have to be hired for the job. This distributive justice component would likely alter perceptions of fairness. For example, Ployhart and Ryan (1998) found that those hired for a job reported a mean fairness rating of 3.80 while the rating for those not hired was 2.89. The second reason that the potential carry-over of reactions to the selection assessments is not likely to impact task performance in an organization is that unlike a research design

in which the task performance would be measured within a relatively short time frame of administering the assessment, organizations may not evaluate performance for anywhere from 60 days to one year after the employee is hired. Therefore, in order to ensure that a measure of typical performance is obtained, college GPA was used to assess task performance to more closely simulate the performance measures obtained in organizations.

Contextual Performance Measure. A 10 item measure adapted from Avis et al. (2002) was used to assess contextual performance (see Appendix G). Their 15 item measure uses a nine-point response format ranging from needs improvement to outstanding, has demonstrated high levels of reliability ($\alpha = .97$), and was found to be related to both cognitive ability and Conscientiousness. This measure was completed by the participant's supervisor at work. If a participant was unable to find a supervisor who was willing to complete the questionnaire, the participant was instructed to ask an instructor, a coworker, or classmate who had observed them at work or in class to complete the measure. Five items were removed from the original measure so that the items could be adapted to apply to either a work or educational setting. The response format was modified to a five-point scale ranging from (1) does not meet expectations to (5) exceeds expectations. This modification was necessary to improve interpretation of the survey when comparing the results to other surveys in this study. Having a supervisor complete the questionnaire was necessary to avoid socially desirable responding. Additionally, the questionnaire asked the respondent to indicate their relationship to the participant so that any differences based on the type of respondent could be determined and controlled for, if necessary. Coefficient alpha was .90 in the present study.

Measure of Expected Performance. A 10 item measure developed based on the skills, abilities, and activities of an entry level management position (O*NET, 2008) was used to assess expected performance. This measure was completed by both the participant (see Appendix H), and by the person asked to complete the contextual measure of performance (see Appendix I). Respondents were asked to rate the participant according to how they believed the participant would perform in a work environment on a five-point scale ranging from (1) would not meet expectations to (5) would exceed expectations. Coefficient alphas ranged from .86 to .92 in the present study.

Procedure

Participants were recruited from undergraduate business and psychology classes at the University of Missouri – St. Louis, and from undergraduate psychology courses at the University of North Carolina – Charlotte.

Upon entering the research session, all participants received informed consent forms. Only after the participant read and signed both copies of the consent form were they allowed to participate in the study. Participants then received a job description for an entry level management position, Office Manager (Appendix J). All participants received the same job description and no additional information about the job was provided even if requested by a participant in order to control for potential contextual confounds, such as amount of information or two-way communication provided (Ryan & Ployhart, 2000). At this point, participants were reminded that the 25 participants with the highest scores would receive $20 and they received basic instructions for the session (Appendix K). Participants were then given a test booklet and instructed to answer all items in the test booklet as though they were applying for the job they just read. The test booklet

contained the following scales in this order: TAS Motivation scale, assessment one, TAS Motivation scale, perceived performance, job relatedness and fairness scales, procedural justice scales, assessment two, TAS Motivation scale, perceived performance, job relatedness and fairness scales, and procedural justice scales. The Conscientiousness context scale followed the TAS Motivation scale rating perceptions for the Conscientiousness measure for all participants. The final two pages of the test booklet asked the participants to report their expected performance, college GPA, and to provide the other demographic information listed in Appendix L in order to assess any individual differences in the relationships being tested. The cognitive ability and Conscientiousness measures were counterbalanced across participants so that 50% of participants received the NEO-FFI Conscientiousness scale for assessment one (Conscientiousness First) while the remaining participants received the Watson-Glaser Form S for assessment one (Cognitive Ability First). Motivation was assessed prior to the first assessment so that differences in initial levels of motivation could be controlled, if warranted.

Along with the test booklet, participants received an envelope with the contextual and expected performance measures, instructions for the supervisor (or other rater), and a postage stamped envelope addressed to the researcher for returning the questionnaires. Additionally, the participants were given a sheet with the researcher's contact information and instructions reminding them of whom should receive the envelope (see Appendix M). All three study materials (test booklet, envelope and questionnaires, and information sheet) had the same unique numerical identifiers on them. This allowed the test booklet to be matched to the measures of contextual and expected performance, and

it also provided a way to contact the participants receiving a monetary reward without having to collect identifying information (such as name) from the participants.

Results

Preliminary Analyses

Ten measures were used in testing the hypotheses and research question for the present study, and respondents completed five additional measures for exploratory purposes. Table 4 contains the means, standard deviations, and reliability coefficients for Conscientiousness First and Cognitive Ability First for the 10 measures. With the exception of Conscientiousness First's perception of fairness for the NEO-FFI Conscientiousness scale ($\alpha = .67$), each of the measures demonstrated acceptable levels of reliability with coefficient alphas ranging from .71 (Watson-Glaser Form S for Cognitive Ability First) to .93 (Cognitive Ability First perceived performance on the NEO-FFI Conscientiousness scale). It should be noted that while Conscientiousness First's perception of fairness for the NEO-FFI Conscientiousness scale did not meet the acceptable levels of reliability, this measure contained only three items, which limits reliability. Additionally, the same three item measure showed higher levels of reliability when used to assess Conscientiousness First's perceptions of fairness for the Watson-Glaser Form S as well as Cognitive Ability First's perceptions of fairness for the NEO-FFI Conscientiousness scale and for the Watson-Glaser Form S ($\alpha = .83, .81$, and .73 respectively).

Preliminary analyses were also conducted in order to ensure that there were no effects due to which researcher oversaw the study sessions, the university at which the participant was enrolled, or the relationship that the participant had to the individual who

completed the measures of contextual performance and expected performance. Independent sample *t*-tests revealed that there were no statistically significant differences based on the researcher (see Table 5) or university affiliation (see Table 6). While the one-way analysis of variance (ANOVA) revealed that Instructors provided statistically significantly lower ratings than did Supervisors for Expected Performance (M = 3.51 & 4.26, SD = .54 & .62 respectively; $F(4, 78) = 2.67$, $p = .04$) and Contextual Performance (M = 3.91 & 4.49, SD = .69 & .46 respectively; $F(4, 86) = 2.64$, $p = .04$) of participants in Cognitive Ability First, these results were not consistent across all groups (see Table 7 and Table 8). Levene's test for homogeneity of variance was significant for Contextual Performance of Cognitive Ability First ($p = .02$), so Welch's test was also conducted as it allows for unequal variances[1]. Welch's test indicated that there were no statistically significant differences on Contextual Performance based on the relationship with the participant for Cognitive Ability First, $F(4, 11.65) = 1.47$. Based on these results, all participants were included in the analyses.

Extension of Gilliland's Model: Hypotheses 1a – 2b.

Correlations were calculated between initial levels of motivation (TAS Initial) and levels of motivation reported after completing the NEO-FFI Conscientiousness scale (TAS Conscientiousness) and the Watson-Glaser Form S (TAS Cognitive Ability). The participants' levels of motivation throughout the procedure were highly correlated and statistically significant (see Table 9), correlation coefficients values ranged from .64 (Cognitive Ability First correlations of Initial TAS rating with both TAS Conscientiousness and TAS Cognitive Ability) to .87 (Conscientiousness First correlation

[1] For each ANOVA conducted, Levene's test for homogeneity of variance was evaluated. In instances where the variances were unequal, the results for Welch's test are also presented.

of Initial TAS rating with TAS Conscientiousness). Therefore, initial levels of motivation were not controlled for when testing Hypotheses 1a and 1b.

To test the hypotheses that Cognitive Ability First would have a higher rating for TAS Conscientiousness than Conscientiousness First (Hypothesis 1a) and that Conscientiousness First would have a lower rating for TAS Cognitive Ability than Cognitive Ability First (Hypothesis, 1b), one-way ANOVAs were conducted. The lack of significant main effects (see Tables 10 and 11) indicate no significant differences in motivation for the NEO-FFI Conscientiousness scale (Hypothesis 1a) or the Watson-Glaser Form S (Hypothesis 1b) based on the order in which the assessments were administered. Levene's Test for homogeneity of variance indicated that there was not equal variance for Conscientiousness First and Cognitive Ability First on motivation for the NEO-FFI Conscientiousness scale; however, Welch's test (which does not assume equal variances) produced similar results F (1, 336.87) = 3.26. Therefore, no support was found for Hypotheses 1a or 1b.

To assess the hypothesized changes in criterion-related validity, correlations were computed between scores on the NEO-FFI Conscientiousness scale and the self-reported GPA as well as between Conscientiousness scores and the average rating for Contextual Performance (Hypothesis 2a) and between scores on the Watson-Glaser Form S and these same two performance measures (Hypothesis 2b). Confidence intervals were calculated around the correlation coefficients to determine if they were statistically different from one another based on the order in which the assessments were given. The confidence intervals for each of the comparisons overlapped (see Table 12), indicating that the correlations were not statistically different. The sign and magnitude of the correlation

coefficients were in the hypothesized direction using GPA as the criterion (r = .15 to .34 for Cognitive Ability First correlations with the NEO-FFI Conscientiousness scale and the Watson Glaser Form S respectively), but were opposite of the hypothesized direction using Contextual Performance as the criterion (r = .07 to .30 for Cognitive Ability First correlations with the Watson-Glaser Form S and the NEO-FFI Conscientiousness scale respectively). Therefore, no support was found for Hypotheses 2a or 2b.

Exploratory Analyses – Comparison to Extant Literature

Given the lack of support for the hypotheses in the present study, it is beneficial to examine the data further in relation to the foundation on which these hypotheses were derived. The first premise on which the present study is founded is that cognitive ability measures are seen as being more job related than personality measures (Hausknecht et al., 2004; Smither et al., 1993; Stiener & Gillilan, 1996; Moscoso & Salgad, 2004). Prior studies looked at personality measures (or descriptions of personality measures) in general as opposed to solely a measure of Conscientiousness, as with the present study. When comparing perceptions of job relatedness for the Watson-Glaser Form S to those for the NEO-FFI Conscientiousness scale, the results did not follow the pattern observed in the previous literature. Conscientiousness First reported significantly higher perceptions of job relatedness for the NEO-FFI Conscientiousness scale (M = 3.34, SD = .67) than the Watson-Glaser Form S (M = 2.76, SD = .80) t (171) = 10.56, p = .001, d = .79. The results for Cognitive Ability First, while not statistically significant, also tended to suggest higher perceptions of job relatedness for the NEO-FFI Conscientiousness scale (M = 2.84, SD = .82) than the Watson-Glaser Form S (M = 2.78, SD = .78) t (173) = -1.17, p = .24.

It is also important to note that the literature evaluated perceptions of job relatedness for selection measures independent of each other and of other procedural justice factors. Unlike the prior literature, the present study measured perceptions of job relatedness in a simulated selection context where participants' perceptions of job relatedness may have been influenced by the other assessment given during the session, or by other factors of procedural justice such as opportunity to perform. If the participants' perceptions of job relatedness were not influenced by extraneous factors such as the aforementioned variables, then there should not be significant differences in perceptions of job relatedness based on group membership. To that end, independent sample *t*-tests were conducted to determine if there were differences in perceptions of job relatedness between the two groups. Conscientiousness First did report significantly higher perceptions of job relatedness for the NEO-FFI Conscientiousness scale (M = 3.34, SD = .67) than Cognitive Ability First (M = 2.84, SD = .82) t (344) = 6.22, p = .001, d = .67. However, there were no statistically significant differences for perceptions of job relatedness of the Watson-Glaser Form S between Conscientiousness First (M = 2.76, SD = .80) and Cognitive Ability First (M = 2.77, SD = .78) t (345) = -.16, p = .87. The fact that significant differences were found for perceptions of job relatedness for the NEO-FFI Conscientiousness scale but not the Watson-Glaser Form S could be an indication that the cognitive ability measure evoked stronger reactions than the Conscientiousness measure. This could be supported by the lower perceptions of job relatedness on the NEO-FFI Conscientiousness scale for participants in Cognitive Ability First as they completed the cognitive ability measure first.

The second premise on which the present study is founded is Schmit and Ryan's (1992) study indicating that motivation differentially influences the validity of cognitive ability measures and personality measures. This aspect has not been fully replicated; therefore, it is prudent to compare their results to the results of the present study. Before a comparison can be made, it is important to note that the cognitive ability and personality measures used by Schmit and Ryan are different from the measures used in the present study. Schmit and Ryan used the School and College Ability Test as a measure of cognitive ability, whereas the present study used the Watson-Glaser Form S. Additionally, Schmit and Ryan used the California Psychological Inventory (CPI) as their personality measure. While the CPI does not purport to measure Conscientiousness, there are facets of the NEO-FFI Conscientiousness scale that show correlations with factors of the CPI allowing for comparison (Costa & McCrae, 1992). Achievement via Conformance, as measured by the CPI, demonstrates correlations with the following facets of Conscientiousness: Dutifulness (r = .43), Achievement Striving (r =.49), Competence (r =.56), and Self-Discipline (r =.57). The CPI factor of Independence is correlated with the Conscientiousness factor Competence (r = .49) and Good Impression is correlated with the Conscientiousness factor Dutifulness (r = .39).

While the use of different measures between Schmit and Ryan's (1992) study and the present study prevent a direct comparison of results, there are similarities to warrant an examination of the relationships between motivation and performance on the measures of cognitive ability and personality. Schmit and Ryan found a non-significant correlation of .14 between motivation and scores on the School and College Ability Test. The present study revealed a similar correlation between motivation and scores on the Watson-Glaser

Form S for Conscientiousness First ($r = .18$, $p < .05$), but a smaller correlation between the same measures for Cognitive Ability First ($r = .09$, ns). The correlations between motivation and factors of the CPI reported by Schmit and Ryan were considerably lower than the correlations found in the present study between motivation and scores on the NEO-FFI Conscientiousness scale. Schmit and Ryan reported the following correlations with motivation: Achievement via Conformance ($r = .43$, $p < .01$), Independence ($r = .21$, $p < .05$), and Good Impression ($r = .28$, $p < .01$). However, correlations between motivation and the NEO-FFI Conscientiousness scale were .51 and .55 for Conscientiousness First and Cognitive Ability First of the present study, respectively. Overall, in the present study the relationships found between motivation and performance on the assessments appear to be inconsistent with those reported by Schmit and Ryan.

Exploratory Analyses – Investigation of Individual Differences

To understand the findings further, exploratory analyses were conducted. Perhaps individual differences may have influenced the results. One-way ANOVAs were conducted to determine if there were significant differences on the assessment scores, reaction measures, or criterion variables based on the participant's reported race. The race categories were reclassified into White, Black, Asian, and Other. Respondents who indicated that they were either Hispanic or Native American were recoded as Other due to the low number of participants in these categories (see Table 3). Most of the variables did not show significant differences between groups. The variables that did demonstrate significant differences are shown in Tables 13 through 16 for Conscientiousness First and Cognitive Ability First. While it is possible that some of the results would be significant by chance due to the number of analyses conducted, the results were examined to

determine if any of the significant results may also have theoretical relevance. Of particular interest are the results related to the participants' self-reported GPA and performance on the Watson-Glaser Form S. For both of these variables, there were significant differences based on the respondent's race for participants in Conscientiousness First and Cognitive Ability First. Black participants in both groups had significantly lower self-reported GPAs than did participants who identified themselves as either White or Asian. Additionally, participants from both groups who identified themselves as Black had lower scores on the Watson-Glaser Form S than did White participants. While not consistent across both groups, participants in Conscientiousness First who reported their race as Asian scored significantly lower on the NEO-FFI Conscientiousness scale than participants who reported their race as White or Black. These results provide some evidence of adverse impact with these assessments.

To investigate further any potential individual differences, *t*-tests were conducted to determine if there were any significant differences on the assessment scores, reaction measures, or criterion variables based on the participants' gender. Statistically significant differences based on gender were found only for two variables in each of the respective groups of participants. None of these variables overlapped across groups. For participants in Conscientiousness First, males scored significantly higher on the Watson-Glaser Form S (M = 26.14, SD = 6.03) than females (M = 24.02, SD = 5.38) t (170) = 2.35, p = .02, d = .37. Within Conscientiousness First, females reported higher levels of motivation on the TAS for Conscientiousness (M = 6.57, SD = .56) than did males (M = 6.34, SD = .80) t (169) = -2.19, p = .03, d = .33. For Cognitive Ability First, the statistically significant differences occurred on the TAS Initial measure and the TAS measure for cognitive

ability. On the TAS Initial measure, females reported significantly higher levels of motivation (M = 6.56, SD = .55) than males (M = 6.23, SD = .83) t (171) = -3.10, p = .002, d = .47. A similar pattern was found on the TAS measure for cognitive ability between females (M = 6.44, SD = .69) and males (M = 6.19, SD = .79) t (172) = -2.15, p = .03, d = .34.

Noting that White participants consistently reported higher GPAs than Black participants, consistently scored higher on the Watson-Glaser Form S than Black participants, and scored higher on the NEO-FFI Conscientiousness scale than Asian participants additional exploratory analyses were conducted to examine Hypotheses 1a, 1b, 2a, and 2b for White participants only. Ideally, the same exploratory analyses would be conducted independently for Black participants and for Asian participants; however, the sample sizes for these races were too small (ranging from 22 to 42 participants; see Table 3) for independent analyses. Table 17 displays the one-way ANOVA results for TAS motivation as a function of order assessment for White participants. No significant differences were found between the means of the two groups for TAS when assessing the participants' motivation on the NEO-FFI Conscientiousness scale or the Watson-Glaser Form S. In fact, Conscientiousness First had higher means on the TAS for conscientiousness than Cognitive Ability First (M = 6.47 and 6.33 respectively) which is counter to Hypothesis 1a. Conscientiousness First also had higher means on the TAS for cognitive ability than Cognitive Ability First (M = 6.48 and 6.35 respectively), thereby not lending support to Hypothesis 1b.

Despite the overall lack of support for Hypotheses 1a and 1b when using only participants who identified themselves as White, validity coefficients were still compared

across groups to understand the relationships more fully. Table 18 displays the validity coefficients and the 95% Confidence Intervals about the coefficients. While the validity coefficients are not significantly different from one another, the magnitude of the validity coefficient for Conscientiousness First is higher than for Cognitive Ability First when examining the relationship between scores on the NEO-FFI Conscientiousness scale and GPA. Additionally, the magnitude of the validity coefficient for Cognitive Ability First is higher than Conscientiousness First when examining the relationship between scores on the Watson-Glaser Form S and GPA. It should be noted that the validity coefficients calculated between Contextual Performance and the assessments did not differ, therefore there was no support for Hypotheses 2a or 2b.

Impact on the Selection System: Hypotheses3a-3b, Research Question 1.

To assess the hypotheses and research question regarding the impact of participant reactions on the selection system as presented in Figure 3, LISREL 8.72 (Jöreskog & Sörbom, 2004) was used to fit this model to the covariance matrix. The model was estimated twice, once using data from participants for which the NEO-FFI Conscientiousness scale was administered first (Conscientiousness First) and again using the data from participants for which the Watson-Glaser Form S was administered first (Cognitive Ability First). Before testing the hypothesized paths, it was necessary to examine the overall fit of the models. For each of the models, all missing data were removed. This resulted in the removal of 95 participants from Conscientiousness First and 96 participants from Cognitive Ability First, resulting in 77 participants in Conscientiousness First and 79 participants in Cognitive Ability First. The majority of participants were removed for missing measures of Contextual Performance and

Expected Performance – Other (85 from Conscientiousness First and 87 from Cognitive Ability First). The majority of the missing data could not be imputed as the measures of Contextual Performance and Expected Performance – Other were not completed or returned by someone on behalf of the participants. Due to the lack of any data on the two measures for these participants, the missing values could not be estimated through imputation methods.

The path estimates for the model using data from Conscientiousness First are presented in Figure 4 ($N = 77$). This model demonstrated poor fit ($\chi^2 = 458.00$, $df = 72$, $p < .01$; root mean square error of approximation (RMSEA) = .49, comparative fit index (CFI) = .06, non-normed fit index (NNFI) = -.18).[2] Given that the estimates of model fit are affected by sample size, a truncated version of the model was run in which the measures of Contextual Performance and Expected Performance Other were eliminated from the model. That is, these two measures were provided by someone other than the participant, and the limited number of returned surveys greatly reduced the sample size for the model presented in Figure 4. By eliminating these two measures, the sample size increased to 162; however, this model (see Figure 5 for path estimates) also demonstrated poor fit ($\chi^2 = 231.98$, $df = 47$, $p < .01$; RMSEA = .14, CFI = .78, NNFI = -.68). Figure 6 presents the path estimates for the model using data from Cognitive Ability First ($N = 79$), also exhibiting poor fit ($\chi^2 = 359.38$, $df = 72$, $p < .01$; RMSEA = .40, CFI = .39, NNFI = .23). A truncated model was estimated for Cognitive Ability First as well (see Figure 7 for path estimates) to increase the sample size ($N = 166$). The fit indices were

[2] NNFI may result in a negative value given the subtraction involved in its calculation: χ^2/df (null model) - χ^2/df (proposed model) ÷ (χ^2/df (null model) -1).

better for this truncated model than any of the previous models ($\chi^2 = 122.56$, $df = 47$, $p < .01$; RMSEA = .10, CFI = .91, NNFI = .88).

Hypothesis 3a indicated that the path estimate between cognitive ability and the criterion would be largest when the Watson-Glaser Form S was administered first. The path estimate between cognitive ability and the criterion was .13 for Cognitive Ability First and .16 for Conscientiousness First; therefore, the data do not support Hypothesis 3a. Hypothesis 3b indicated that the path estimate between Conscientiousness and the criterion would be largest when the NEO-FFI Conscientiousness scale was administered first. The path estimate between Conscientiousness and the criterion is -.03 for Conscientiousness First and .26 for Cognitive Ability First; therefore, the data do not support Hypothesis 3b.

Research Question 1 was evaluated by examining the path estimate between Conscientiousness and the criterion for Conscientiousness First relative to Cognitive Ability First, to determine if the path estimate approached zero for Cognitive Ability First. As seen in the results for Hypothesis 3b, the path estimate between Conscientiousness and the criterion was higher for Cognitive Ability First than Conscientiousness First. Therefore, this study does not offer evidence that incremental validity obtained by combining a Conscientiousness measure with a measure of cognitive ability is lost when the Conscientiousness measure is given after the measure of cognitive ability.

Exploratory Analyses – Procedural Justice Scales

Due to the overall lack of support for the model proposed, exploratory analyses were conducted to determine if any of the subscales from Bauer et al.'s (2001) Selection

Procedural Justice Scale may explain the results. Table 19 displays the means, standard deviations, and coefficient alphas for these scales for both Conscientiousness First and Cognitive Ability First. The coefficient alphas for each of these scales reached acceptable levels (α = .71 to .94). When evaluating the potential impact these variables may have on the model, it is interesting to note the consistently low means for Chance to Perform and Information Known. Each time these scales were administered, means were below 3.0 on a 5.0 scale indicating that participants generally disagreed with statements that these procedural justice conditions were met for the administration of the NEO-FFI Conscientiousness scale and for the Watson-Glaser Form S. These variables may therefore have affected the results, given that Gilliland (1993) proposed perceptions of injustice would have a greater effect on outcomes than varying degrees of justice.

To investigate such potential impact of the Procedural Justice subscales further, correlations were computed between each of the variables (see Tables 20 and 21). Chance to Perform, which had low means, consistently had significant correlations with perceptions of Job Relatedness and Fairness across both groups. Additionally, while there are numerous significant correlations between components of the Procedural Justice subscales, and the variables included in the present study; consistent significant relationships were found for the Treatment and Propriety of Questions subscales. These subscales showed statistically significant relationships with the NEO-FFI Conscientiousness Scale, TAS Initial, TAS Conscientiousness, Perceived Performance Conscientiousness, TAS Cognitive Ability, and Perceived Performance Cognitive Ability across Conscientiousness First and Cognitive Ability First. These significant relationships were found for the perceptions of Treatment and Propriety of Questions when referencing

both the NEO-FFI Conscientiousness Scale and the Watson-Glaser Form S. Additionally, the Propriety of Questions subscale consistently had significant relationships with Expected Performance – Self, Job Relatedness Conscientiousness, Fairness Conscientiousness, Job Relatedness Cognitive Ability, and Fairness Cognitive Ability. These relationships were found for both groups when referencing the NEO-FFI Conscientiousness scale and the Watson-Glaser Form S.

Discussion

Overall, the results of the present study do not support the theory that applicant reactions to the job relatedness of cognitive ability and Conscientiousness assessments impact validity when the assessments are given in the same session, such that the assessment given second has lower validity than had it been given alone (as depicted in Figure 3). Given these results, it is important to examine what other elements might influence or explain the outcomes. For example, finding that a cognitive ability assessment was not viewed as more job related than a personality assessment may be an indication of several factors.

One factor may be the measurement bandwidth of personality. The present study focused the personality measure on one facet, Conscientiousness, based on validity evidence in prior research. Whereas previous studies presented only descriptions of the assessments, they were more inclusive of all factors of personality. Given individuals might see Conscientiousness as the most job related of the various personality dimensions, it may have been the other factors of personality that drove the lower ratings of job relatedness in previous studies. This may have also resulted from the participants reading and responding to each of the items of the NEO-FFI Conscientiousness scale as

opposed to reading overall descriptions of personality measures. As the participants in this study were enrolled in either psychology or business classes, it is also possible that they had encountered discussions of the usefulness of Conscientiousness in selecting the best applicants in some of their courses. This information could have affected their perceptions that an assessment which was clearly labeled as a measure of Conscientiousness was more job-related than had they had no prior knowledge of the value of these assessments in selection. Additionally, participants may have deemed Conscientiousness to be more related to the entry-level managerial position based on the job description they were given. Finally, the unexpected results may be an indication that the NEO-FFI Conscientiousness scale itself is seen as more job related than other measures of personality due to the wording used in some of the items (the word "work" appears in two items and the word "job" appears in a third item). Support for this supposition may be provided by the participants agreeing that they thought about how they behave at work when answering the Conscientiousness items (M = 4.30 and 3.89 SD = 1.09 and 1.35 for Conscientiousness First and Cognitive Ability First respectively).

Another factor that must be considered when evaluating the results is the method of inducing motivation. While it appears that the participants were motivated, as evidenced by means on the TAS ranging from 6.34 to 6.56 on a seven-point scale (see Table 4), it is possible that offering a cash reward for high performance did not induce motivation in the same manner that hiring the higher performers would have. For example, Gilliland (1994) hired the top scorers to complete clerical duties in order to induce motivation. While this situation may explain some of the lack of results, it is not complete given Chan and colleagues (Chan, 1997; Chan et al., 1997; Chan et al., 1998)

offered a cash reward to top performers, similar to the present study. However, in the studies conducted by Chan and colleagues, the reward was given to a higher percentage of participants than in the present study.

Of additional consideration are the high levels of motivation reported and the level of variance in the participants' motivation to perform on the assessments. As described above, mean ratings on the TAS were all above 6.00 on a seven-point scale in the present study. Additionally, the present study demonstrated limited variance in participants' motivation with standard deviations ranging from .59 to .76 (see Table 4). This restriction is in contrast to Chan et al.'s (1997) study in which the mean TAS score was 3.88, though their study also had limited variance with a standard deviation of .73. While it is clear that the participants in the present study reported that they were in fact motivated to do well on the assessments, the source of their motivation remains unclear.

The criterion measures may have also contributed to the lack of support for the hypotheses. While the use of GPA has been employed in the literature as a measure of performance (Schmit & Ryan, 1992), of particular concern here are the subgroup differences in reported GPA, such that Black participants reported significantly lower GPA scores than White or Asian participants. The impact of GPA with respect to the participant's race is supported by the differing validity coefficients in the hypothesized direction, though the differences remained statistically non-significant, when assessing the White participants alone. The criteria measure of Contextual Performance may have also contributed to the lack of support for the hypotheses. This measure of Contextual Performance was adapted from Avis et al.'s (2002) 15-item measure, reduced to 10 items such that it could be completed by a supervisor, instructor, or another person who knew

the participant. While this scale demonstrated acceptable levels of reliability ($\alpha = .90$ for Conscientiousness First and Cognitive Ability First), the changes made by eliminating five items and allowing persons other than a direct supervisor complete the measure may have affected the validity results by confounding the definition of performance to include performance outside of a work environment.

A final consideration of the results is the influence of extraneous variables. While Gilliland (1993) proposed that job relatedness would have the greatest impact on outcomes associated with procedural justice (i.e., test-taking motivation), he also proposed that a violation of a procedural justice rule would be more salient than decreases in the degree of satisfaction that a procedural justice rule was met. It is possible that Chance to Perform, which had much lower means than Job Relatedness, was more salient to the respondents and therefore had greater impact on the outcomes than did the Job Relatedness of the assessments. The consistently high correlations between Chance to Perform and Job Relatedness, as well as Fairness may be an indication of this variable's importance to the results.

Each of the aforementioned factors could explain the results of the present study, resulting in the lack of support for the proposed hypotheses as well as the results found for the research question. Given the possibility that such unknown factors may be influencing the relationships shown in Figure 3, the non-significant effects cannot be taken as conclusive evidence that the proposed impact on criterion-related validity does not exist. Coupled with the existing research that supports many aspects of the model proposed in the present study, the current findings should not be used to conclude that organizations may administer cognitive ability and personality assessments in a single

selection session without adverse results on the validity of the assessments. As such, it is important to conduct additional research to understand the relationships between the variables in the present study to ensure that organizations do not inadvertently administer assessments in a manner that would impact the assessments' validities. One such study would be to investigate the variables when the job relatedness of the assessments is altered such that both assessments should be viewed as more applicable to the job.

Altering Applicant's Reactions – Study Two

Though the model tested in Study 1 was not supported, the aforementioned evidence in the literature indicates that it is important to investigate strategies to offset the potential decrease in criterion-related validity when combining selection assessments in a single testing session. Additionally, there has been a call to move beyond simply describing the outcomes of applicant reactions to investigate how we can improve outcomes by altering applicant reactions (Chan, 2007). There has also been an assertion that we should give organizations concrete examples of how to alter their selection systems based on applicant reactions, including the sequencing of multiple selection assessments (Ryan & Huth, 2008).

Frame-of-Reference Effects for Cognitive Ability Assessments

One method of altering applicant's reactions would be to change their perceptions of job relatedness for the assessment by providing the participants with a contextual frame-of-reference in which the items are altered to be business-related. It stands to reason that by presenting items which use business-related situations, it will be easier for a participant to see how the content of the assessment could be related to the content of the job as well as how the information gained from the assessment could be used to

determine future job performance. Evidence of this effect has been provided for cognitive ability tests. Rynes and Connerely (1993) found that when provided with descriptions of assessments, mean ratings for belief that the company needed to know the information for selection purposes were higher for a business related form of the Watson-Glaser assessment of general ability than for the unaltered form of the Watson-Glaser (M = 4.80 and 3.45 respectively). With regards to personality measures, there is less conclusive evidence for the enhanced job relatedness of personality tests administered with an "at work" frame-of-reference.

Frame-of-Reference Effects for Personality Assessments

Holtz, Ployhart, and Dominguez (2005) found that a frame-of-reference context for a personality measure did not enhance perceptions of job relatedness. The inconsistencies in the results between providing a frame-of-reference for cognitive ability items and providing the same context for personality items may be due to the way in which the context is provided. In cognitive ability assessments the frame-of-reference can be provided by providing a business situation (e.g., asking the participant to calculate a percentage based on increases in profit compared to just providing a mathematical sentence to be solved) whereas the changes to personality assessments are usually made by adding the words "at work" to the item stem or by instructing participants before the start of the assessment to think about the items as they would react in a work environment. Therefore, the "at work" addition to personality items or instructions simply may not be enough to change how the participants view the items. Though a frame-of-reference may not serve to alter the relationships of interest in Gilliland's (1993) model (Figure 2), it does appear as though an at work context may alter

participants' responses to personality assessments in a way that would be beneficial to the part of Gilliland's model assessed in the present paper.

While Holtz et al. (2005) were unable to alter the perceptions of job relatedness on a personality assessment by providing a frame-of-reference, the authors reported the at work context did impact participants' responses on the personality assessment in that there was decreased error variance compared to the unaltered version of the personality assessment. The change in participants' scores on the personality assessments has been supported by other researchers as well. Schmit et al. (1995) proposed that personality assessments lack stronger criterion-related validity coefficients because participants may be thinking about different contexts when answering the items. For example, some participants may respond to items in a manner that reflects their behaviors at work, while others may think of non-work situations when responding. Personality measures that include the words "work" or "job" in only some of the items (e.g., the NEO-FFI Conscientiousness scale) may also result in participants viewing items in different contexts which could be controlled by providing a frame-of-reference. Support for Schmit et al.'s hypothesis was found; when participants were instructed to think about a common frame-of-reference, the error variances were reduced. Additionally, Schmit et al. gathered criterion data, in the form of cumulative college GPA, and found that participants who received general testing instructions and a school specific personality assessment had higher validity coefficients than those who received general testing instructions and the non-contextual personality assessment (r = .41 and .25, respectively). Bing, Whanger, Davison, and VanHook (2004) offer further evidence of the benefit of providing a frame-of-reference to personality assessments when they found

that personality assessments contextualized for school had incremental validity above that of cognitive ability and the non-contextual personality assessment. To ensure that the decreased error variance is attributable to the at work context, it is necessary to check the manipulation given that it has not been shown to alter the perceptions of job relatedness (Holtz et al., 2005).

Though these findings show that constructs other than job relatedness, fairness reactions, and test-taking motivation can alter the validity of a personality assessment, these results are beneficial to improving the hypothesized validity outcomes associated with Figure 2. The increase in validity reported by prior work aligns well with Holtz et al.'s (2005) findings that providing a frame-of-reference for personality assessments did not impact perceptions of job relatedness. Had the manipulation affected perceptions of job relatedness, one would anticipate an increase in test-taking motivation and therefore a decrease in the validity of the personality assessment as evidenced by the arguments set forth by Gilliland (1993) and based on Schmit and Ryan's (1992) study. While other studies have not found this same increase in validity, they also have not reported a substantial decrease in the validity of personality assessments. Hunthausen, Truxillo, Bauer, and Hammer (2003) reported the same trend in validity coefficients with $r = .31$ in the frame-of-reference condition compared to $r = .10$ in the standard condition; however, these differences lacked statistical significance. The lack of statistical significance may have resulted because the researchers controlled for reported test-taking motivation when they found that test-taking motivation was lower in the frame-of-reference condition.

Given that providing a frame-of-reference has been shown to increase perceptions of job relatedness in cognitive ability assessments and increased criterion-related validity

of personality assessments, the potential exists to improve the validity of both types of assessments when given in the same session. In order to maximize this potential, both assessments would need to have a frame-of-reference context and the Conscientiousness assessment would need to be given first. It is expected that the frame-of-reference context for the Conscientiousness measure would result in higher validity coefficients when it is administered first, due to the reduction in error variance by ensuring that all participants are responding based on their behavior in the same context. While the frame-of-reference manipulation on the Conscientiousness assessment should not impact perceptions of job relatedness, fairness, or test-taking motivation, the frame-of-reference manipulation on the cognitive ability assessment could increase those perceptions to overcome any potential negative impact of having lower levels of test-taking motivation after completing the Conscientiousness measure, thereby allowing for a higher validity coefficient for the cognitive ability assessment than if the frame-of-reference context had not been used. It should be noted that Hunthausen et al. (2003) did not find increased validity coefficients for a cognitive ability assessment when providing a frame-of-reference context. However, in their study, they counterbalanced the order in which the personality and cognitive ability assessments were given in an effort to control for the potential effects of cognitive ability. The authors also controlled for cognitive ability by entering it into the first step of their hierarchical regression. The intentional control for potential order effects, therefore, may have affected the reported results.

The reasons for not anticipating benefits from providing frame-of-reference context when the cognitive ability assessment is administered first are as follows: the alteration of the cognitive ability test should result in an increase in test-taking

motivation, the increase in test-taking motivation above the high levels of motivation already expected should remain when the participants are completing the personality assessment, and the elevated levels of test-taking motivation may decrease the potential benefits of the frame-of-reference context for the personality assessment as there have been mixed results as to how the alteration impacts the criterion-related validity of personality assessments. Due to the conflicting results reported in the literature, the following research question is proposed.

Research Question 2: Can the validity of both the cognitive ability and Conscientiousness assessments be optimized by providing a frame-of-reference context and administering the Conscientiousness assessment first?

Methods

Participants

There were 177 participants in this study. One participant completed the assessments out of order and was removed from all analyses. Participants were students enrolled in psychology or business classes at the University of Missouri – St. Louis (N = 160) or psychology classes at the University of North Carolina – Charlotte (N = 16). The participants were recruited along with the participants in Study 1, but participants were not told about the separate studies to ensure random distribution across studies. Participants were not allowed to participate in Study 2 if they participated in Study 1. The majority of the participants were female (N = 105) and Caucasian (N = 106). Participants ranged in age from 18 to 56 (M = 23.64, SD = 5.37). Table 22 contains additional demographic information for participants in Study 2.

Measures

Job Relatedness. Perceptions of job relatedness were measured by combining Smither et al.'s (1993) scales for face validity and perceived predictive validity with Bauer et al.'s (2001) scales for job relatedness predictive and job relatedness content (see Appendix A). These four scales use a five-point response format ranging from strongly disagree to strongly agree. Coefficient alphas ranged from .88 - .93 in the present study.

Fairness. Perceptions of fairness were measured by Chan et al.'s (1998) three item fairness perceptions scale (see Appendix B). This scale also uses a five-point response format ranging from strongly disagree to strongly agree. Coefficient alphas ranged from .68 - .80 in the present study.

Test-Taking Motivation. Participants' test-taking motivation was measured by the Motivation scale of the TAS (Arvey, Strickland, Drauden, & Martin, 1990; see Appendix C). This 10 item scale assessed the participants' desire to perform well on an assessment using a seven-point disagree - agree response format. Coefficient alphas ranged from .92 - .94 in the present study.

Perceived Performance. Participants' perceived performance on each of the assessments was measured with Sanchez et al.'s (2000) four item scale (see Appendix D). This scale uses a five-point response format ranging from strongly disagree to strongly agree. Coefficient alphas ranged from .92 - .93 in the present study.

Conscientiousness Context Scale. Participants completed a five item scale asking them to indicate the context they were thinking of when responding to the personality items. Using a five-point response format (strongly disagree to strongly agree), participants were asked to indicate the extent to which they were thinking about the

following contexts while responding: at work, at school, at home, with friends, every day (see Appendix E).

Procedural Justice Scales. The following subscales from Bauer et al.'s (2001) Selection Procedural Justice Scale were added for exploratory purposes (see Appendix F): Information Known (extent to which the participant knew what to expect on the assessment), Chance to Perform (extent to which the participant could demonstrate his or her abilities on the assessment), Consistency (extent to which the participant believed the assessment was administered consistently across participants), Treatment (extent to which the participant felt he or she was treated politely and with respect during the assessment), and Propriety of Questions (extent to which the participant believed the content of the assessment was appropriate).

Conscientiousness Measure. The Conscientiousness scale from the NEO-FFI (Costa & McCrae, 1992) was used to assess the participants' Conscientiousness. For this study, the NEO-FFI Conscientiousness scale was altered to an "at work" frame-of-reference by altering the instructions such that the participants were told to think about how they would act and feel at work when answering the questions. The use of the NEO-FFI Conscientiousness scale was consistent with other studies investigating the effects of a frame-of-reference manipulation (Hunthausen et al., 2003; Holtz et al., 2005). Coefficient alpha was .83 in the present study.

Cognitive Ability Measure. The Hogan Business Reasoning Inventory (HBRI) was used to assess cognitive ability in this study. In considering a measure of cognitive ability for this study, it was important to have a measure that had previously demonstrated acceptable levels of validity and used a business related format. It was also

necessary to find a measure that approximated the format of the Watson-Glaser Form S so that comparisons to results from the first study would not be confounded by the length or administration format of the two measures. The HBRI was developed for use in selecting professionals and assesses cognitive ability through verbal, quantitative, and graphic items that address both tactical reasoning and strategic reasoning. This measure has 24 multiple choice items that can be answered in 30 minutes and is considered a power test as opposed to a speed test. The manual for the HBRI reports a correlation of .27 with undergraduate GPA and of .66 with the Watson-Glaser Critical Thinking Appraisal. Coefficient alpha was .69 in the present study.

Task Performance Measure. The participants' self-reported college GPA was used as the task performance criterion measure.

Contextual Performance Measure. A 10 item measure adapted from Avis et al. (2002) was used to assess contextual performance (see Appendix G). This measure was completed by a supervisor, instructor, coworker, or classmate. Coefficient alpha was .94 in the present study.

Measure of Expected Performance. A 10 item measure developed based on the skills, abilities, and activities of an entry level management position (O*NET, 2008) was used to assesses expected performance. This measure was completed by the participant (see Appendix H) and by the person asked to complete the contextual measure of performance (see Appendix I). Coefficient alphas ranged from .85 - .93 in the present study.

Procedure

Participants for this study were recruited at the same time as those for the first study. However, students were only allowed to participate in a single study. In order to ensure that participants participated in only one of the studies, and that the assignment to the study in which they participated was random, participants were not told that there were two separate studies being conducted. Because the participants were not aware of the multiple studies, participants' scores from both studies were considered when awarding the 25 $20 prizes for top scorers.

Upon entering the research session, all participants received informed consent forms. Only after the participant had read and signed both copies of the consent form were they allowed to participate in the study. Participants then received a job description for an entry-level management position, Office Manager (Appendix J). All participants received the same job description and no additional information about the job was provided if requested by a participant in order to control for potential contextual confounds such as amount of information or two-way communication provided (Ryan & Ployhart, 2000). At this point, participants were reminded that the 25 participants with the highest scores would receive $20 and received basic instructions for the session (Appendix K). Participants were then given a test booklet and instructed to answer all items in the test booklet as though they were applying for the job they just read about. The test booklet contained the following in this order: TAS Motivation scale, contextualized NEO-FFI Conscientiousness scale, TAS Motivation scale, Conscientiousness context scale, perceived performance, job relatedness and fairness scales, procedural justice scales, Hogan Business Reasoning Inventory, TAS Motivation

scale, perceived performance, job relatedness and fairness scales, and procedural justice scales. The final two pages of the test booklet asked the participants to report their expected performance, college GPA and to provide the additional demographic information listed in Appendix L. While the research question proposed for this study did not directly address the participants' reactions to the cognitive ability and Conscientiousness assessments, the participants were still asked to provide their reactions to the assessments. By gathering this information, exploratory analyses could be conducted should the validity of the assessments fail to be optimized by the proposed manipulations.

Along with the test booklet, participants received an envelope with the contextual and expected performance measures, instructions for the supervisor and a postage stamped envelope addressed to the researcher for returning the questionnaire. Additionally the participants were given a sheet with the researcher's contact information and instructions reminding them of whom should receive the envelope (see Appendix M). All three study materials (test booklet, envelope and questionnaires, and information sheet) had the same unique numerical identifiers on them. This allowed the test booklet to be matched to the measures of contextual and expected performance, and it also allowed a way to contact the participants receiving a monetary reward without having to collect identifying information (such as name) from the participants.

Results

Preliminary Analyses

There were ten measures used in testing the research question for the present study, and five additional measures given for exploratory purposes. Table 23 contains the

means, standard deviations, and reliability coefficients for the 10 measures used for the present study. Of the 10 measures used for the present study, only two of the measures did not meet acceptable levels of reliability. When the fairness sale was used to measure participants' perceptions of the contextualized Conscientiousness measure the resulting coefficient alpha was .68. Reliability was only slightly higher for the HBRI ($\alpha = .69$). The remaining measures had acceptable levels of reliability with coefficient alpha ranging from .80 (perceptions of fairness for the cognitive ability measure) to .94 (TAS motivation scale for the cognitive ability measure).

Preliminary analyses were also conducted in order to ensure that there were no effects due to the researcher overseeing the study sessions, the university at which the participant was enrolled, or the relationship that the participant had with the individual who completed the measures of Contextual Performance and Expected Performance for the participant. Independent sample *t*-tests revealed that there were not any statistically significant differences based on the researcher (see Table 24) or university affiliation (see Table 25). Additionally, a one-way ANOVA revealed that there were not any statistically significant differences based on the relationship between the participant and the person completing the measures of Contextual Performance and Expected Performance on behalf of the participant (see Table 26 and Table 27). Based on the results of these preliminary analyses, all participant data were analyzed together.

Research Question Analyses

To assess any changes in criterion-related validity, correlations were computed between scores on the contextualized NEO-FFI Conscientiousness scale and each of the criterion measures, and between the HBRI and each of the criterion measures. These

validity coefficients were compared to the validity coefficients for Hypotheses 2a and 2b from Study 1. After the validity coefficients were computed, confidence intervals around the coefficients were calculated to determine if there were statistically significant differences. The confidence intervals overlapped for all of the comparisons indicating that the correlations were not statistically different from one another (see Table 28). Additionally, for the differences that were large enough to potentially offer practical significance (e.g., NEO-FFI Conscientiousness scale with GPA) the differences were not in the anticipated direction. Therefore, the validities for the two assessments were not optimized by using assessments contextualized for an "at work" frame of reference and administering the measure of Conscientiousness first.

Exploratory Analyses

Comparison of Conscientiousness Context. Given the results found for Research Question 2 it is important to look at factors that may be influencing the results. As Holtz et al. (2005) noted, when providing a frame-of-reference context for personality assessments, it is important to evaluate the extent to which the participants were responding based on the given context. To that effect, the means and standard deviations for the Conscientiousness context items were reviewed (see Table 23). The mean is highest for the item asking participants the extent to which they thought about their actions at work when responding to the contextualized NEO-FFI Conscientiousness scale, which is consistent with the frame-of-reference manipulation. However, the means for the remaining four Conscientiousness context items were also considerably high on the five point scale which may indicate that the frame-of-reference provided did not decrease error variance as described by Holtz et al. (2005).

To better understand the extent to which the contextualization of the NEO-FFI Conscientiousness scale resulted in the participants responding to the items based on a work context, one-way ANOVAs were conducted using responses to each of the five items from participants in Conscientiousness First and Cognitive Ability First from Study 1 along with participants from Study 2 (see Tables 29 and 30). The results reveal statistically significant differences in the means across groups for the item asking participants the extent to which they thought about their actions at work: F (2, 520) = 11.01, $p < .01$. Additionally, the mean for this context item was highest for participants in Study 2, though post hoc analyses reveal that the mean from participants in Study 2 (M = 4.43, SD = .88) was only significantly different from those in Cognitive Ability First of Study 1 (M = 3.89, SD = 1.35). Also of note is that the mean was lowest for participants in Study 2 (compared to Conscientiousness First and Cognitive Ability First from Study 1) on the items asking participants if they thought about how they behave at school, at home, or with friends. This may offer further indication that the frame-of-reference context did result in participants focusing more on the "at work" context when responding to the contextualized NEO-FFI Conscientiousness scale, though it did not fully eliminate participants' thoughts of other contexts when responding.

Investigation of Individual Differences. Consistent with Study 1, exploratory analyses were conducted to determine if individual differences may have influenced the results of the present study by acting as extraneous variables. One-way ANOVAs were conducted to determine if there were significant differences on the assessment scores, reaction measures, or criterion variables based on the respondent's reported race. As with Study 1, the race categories were reclassified into White, Black, Asian, and Other.

Respondents who indicated that they were either Hispanic or Native American were recoded as Other due to the low number of respondents in these categories (see Table 22). While most of the variables did not show significant differences between groups, a few variables did as displayed in Tables 31 and 32. Significant differences based on race were found for self-reported GPA, performance on the HBRI, and the contextualized NEO-FFI Conscientiousness scale. As with participants in Study 1, Black participants in the present study had significantly lower self-reported GPAs than did participants who identified themselves as either White or Asian. Participants who identified themselves as White had significantly higher scores on the HBRI than participants who identified themselves as Black or Asian. Finally, participants who reported their race as Asian scored significantly lower on the contextualized NEO-FFI Conscientiousness scale than participants who reported their race as White or Black.

To further investigate any potential individual differences, *t*-tests were conducted to determine if there were significant differences on the assessment scores, reaction measures, or criterion variables based on the participants' gender. Of all the variables, statistically significant differences based on gender were only found for three of the procedural justice subscales. Females had significantly higher means for Treatment Conscientiousness (M = 4.68, SD = .44) than males (M = 4.50, SD = .57) t (171) = -2.46, p = .00, d = -.35. Females also had significantly higher means for Treatment Cognitive ability (M = 4.68, SD = .49) than males (M = 4.53, SD = .56) t (172) = -1.82, p = .05, d = -.29. Finally, females had significantly higher means for Information Known Cognitive Ability (M = 2.75, SD = 1.13) than males (M = 2.71, SD = .93) t (172) = -.24, p = .01, d = -.04.

Based on the initial results of the exploratory analyses into the impact of individual differences (i.e., race), additional analyses were conducted. Given there were some racial differences identified, exploratory analyses were conducted to determine if the validities could be optimized through the use of assessments contextualized with an "at work" frame-of-reference and administering the Conscientiousness measure first when the sample was limited to White participants only. Ideally, the same exploratory analyses would be conducted independently for Black participants and for Asian participants; however, the sample sizes for these races were too small for independent analyses (N = 24 and 21 respectfully). Table 33 displays the validity coefficients and the 95% Confidence Intervals about the coefficients. In each instance, Study 2 had the lowest validity coefficient; however, the results were not significantly different from one another as evidenced by the overlapping confidence intervals.

Procedural Justice Scales. Consistent with Study 1, exploratory analyses were also conducted to determine if any of the subscales from Bauer et al.'s (2001) Selection Procedural Justice Scale may have an impact on the results of the present study. Table 34 displays the means, standard deviations, and coefficient alphas for these scales for Study 2. With the exception of Propriety of Questions Cognitive Ability (α = .67), the reliability of these scales reached acceptable limits (ranging from .74 for Information Known Conscientiousness to .94 for Consistency Cognitive Ability). As observed in the results for Study 1, there were consistently low means for Chance to Perform and Information Known. Each time these scales were administered, means were below 3.0 on a 5.0 scale indicating the participants generally disagreed with statements that these procedural

justice conditions were met for the administration of the contextualized NEO-FFI Conscientiousness scale and for the HBRI.

To investigate further the potential impact of the Procedural Justice subscales on the present study, correlations were computed between each of the variables included in the present study (see Table 35). Focusing on the variables of interest for Study 2 (i.e., contextualized NEO-FFI Conscientiousness scale, HBRI, GPA, and Contextual Performance) participants' perceptions of Treatment during the session were significantly correlated to each of the four variables, with correlation coefficients ranging from .17 (Treatment Cognitive Ability with GPA) to .35 (Treatment Conscientiousness with the contextualized NEO-FFI Conscientiousness scale).

Discussion

The HBRI was chosen for the present study due to its business related context and its reported correlation to the Watson-Glaser Critical Thinking Appraisal. Despite the relatively high reported correlation between the HBRI and the Watson-Glaser Critical Thinking Appraisal, results of the present study may be an indication that the HBRI is not suitable as a contextualized equivalent to the Watson-Glaser Form S. The means presented in Tables 4 and 23 indicate that the HBRI may be a more difficult measure of cognitive ability. The average scores on the Watson-Glaser Form S were 61.83% and 62.08% for Conscientiousness First and Cognitive Ability First respectively, while the average score on the HBRI was 43.88%. This potential difference in difficulty could offer an explanation as to why the HRBI did not demonstrate higher validity coefficients than those found in Study 1. The self-reported GPAs were equivalent across groups, an

indication that all groups possessed the same ability level. Use of a more difficult measure of cognitive ability does not allow for an equal comparison across conditions. The results found in the present study may also have been impacted by the criterion measures used due to racial differences. As found in Study 1, participants who identified themselves as Black reported lower GPAs than participants who reported their race as White or Asian. The racial differences found may indicate that some variable other than performance may influence the criterion measure, thereby impacting the results of the present study. However, the extent to which the use of GPA affected the results of the present study is unclear given that the validity coefficients were not higher than those found in Study 1 when only the results of White participants were analyzed. The measure of Contextual Performance may also have influenced the results of the present study. As with Study 1, the elimination of 5 items from Avis et al.'s (2002) measure and allowing someone other than a direct supervisor to respond to the items may have influenced the validity results found in the present study by confounding the definition of performance to include performance outside of a work environment.

While the present study did not set out to investigate the potential impact of participant reactions, as measured by the procedural justice scales, it stands to reason that perceptions of procedural justice may have influenced the results. Particularly, it seems that participants in the present study felt as though they did not adequately get to demonstrate their abilities as seen by the consistently low means for the Chance to Perform subscale. The significant correlations between Treatment and the variables of interest in the present study may be an indication of its impact on the results of the present study as well.

General Discussion

The present studies strove to extend beyond previous lines of research by exploring Gilliland's application of organizational justice to the field of human resources and evaluating what happens when two different assessments are given in the same session. These goals were addressed in the following ways: 1) determining if criterion-related validity is lower when an assessment is administered after another selection measure than if it had been administered alone, 2) evaluating the manner in which participant reactions impact subsequent test performance by testing a model of a selection system using both a cognitive ability assessment and a measure of Conscientiousness, and 3) assessing if the validities of selection assessments administered in the same session could be maximized by altering the context of the assessments and the order in which they are given.

The results of the first extension had the potential to be beneficial to practitioners attempting to apply existing research to their practices. There is often a gap between research and practice, and the first study attempted to bridge that gap by demonstrating that when practitioners base decisions on validity results of a study using a single assessment (such as cognitive ability), and then combine that assessment with another assessment in practice (such as Conscientiousness) that the validities are no longer equivalent. Ideally this should not be a problem as organizations should conduct their own validation studies; however, many times organizations do not have the means by which to conduct these analyses nor do they have enough employees to provide the sample size necessary to achieve adequate power for these analyses. While there was not support for the theory addressed in the first study, practitioners should still use caution

when combining assessments in the same session for their selection systems. The evidence provided in the existing literature reviewed for this paper combined with the results of the exploratory analyses are an indication that the results of this study should be interpreted with caution.

Implications

Theoretical. The implications of the second extension to the extant literature provided by the first study are more theoretical in nature. By assessing the model depicted in Figure 3 researchers can gain a greater understanding for how reactions persist over time and how they are combined with reactions to a new assessment to alter performance. While the results did not support the model depicted in Figure 3, as evidenced by the poor fit of the model, the changes in validity coefficients, though they were not statistically significantly different, may be an indication that care should be used in research designs that rely on counterbalancing the measures.

The final extension, addressed in the second study, also has theoretical implications. The theoretical implication of testing the research question in Study 2 is that it had the potential to provide some insight into which outcomes may be altered by manipulating participants' perceptions of procedural justice rules. This addressed Chan's (2007) assertion that knowing how reactions to procedural justice rules impact outcomes is not enough, rather we need to understand what we can do to improve those outcomes given the observed reactions. While the validities were not maximized with the manipulations in Study 2, the results still provide an indication of how manipulations may impact participants' reactions and ultimately the validity of the measures.

Practical. While the results of these studies may not be statistically significant, in a true selection context the differences may prove to be practically significant indicating that organizations should use caution when counterbalancing the assessments administered in their selection systems. The exploratory results of both studies offer another practical implication. In both studies, participants who reported their race as Black had lower scores on the cognitive ability assessment and one of the criterion measures (self-reported GPA). This could be an indication that organizations should use caution when combining a measure of Conscientiousness with a measure of cognitive ability known to have adverse impact. Given that adding a measure of personality to a cognitive ability assessment is a known strategy for reducing overall adverse impact, organizations may need to consider the potential decrease in validity of the assessments when implementing this strategy.

The results of Study 2 provide additional practical implications. Addressing how the outcomes could be improved when combining assessments is also of practical importance as it provides practitioners insight into what may happen when they try to maximize the validity of the assessments they are using. The results of the second study may indicate that contextualizing a measure of Conscientiousness through the instructions alone, may not be enough to decrease the error variance and subsequently increase the criterion-related validity. Additionally, there is some evidence that a contextualized cognitive ability assessment may not have higher criterion related validity if other factors influence the applicants' performance, such as difficulty of the assessment.

Limitations

The first limitation of the present studies was the use of students as participants. It could be argued that despite being instructed to respond as an applicant would, the students were not motivated in the same manner as an actual applicant. The potential lack of motivation by the students was addressed in two ways: by offering a $20 award to the 25 participants with the top scores across both studies (determined by the percentage of correct responses to the measure of cognitive ability they were given) and by measuring their initial levels of test-taking motivation. While the participants in both studies reported high levels of motivation, the source of their motivation remains unknown. The overall high reported measures of motivation and relatively small amount of variance makes it difficult to conclude whether or not the participants were as motivated on these assessments as they would have been had they been applying for a job. However, the use of student participants was not only consistent with previous research in this area (Rynes & Connerley, 1993; Gilliland, 1994; Chan, 1997; Chan et al., 1997; Chan et al., 1998; LaHuis et al., 2003), but it was also necessary due to the hypothesized effects on test-taking motivation and criterion-related validity. It would have been unethical to ask applicants in an actual organizational job search to participate in the present studies as it was hypothesized that their test-taking motivation levels would be altered throughout the testing process. Additionally, it would have been unethical to ask organizations to allow their applicants to participate in the present study as it was hypothesized that the criterion-related validity of the assessments would be altered as well.

The second limitation of these studies was the use of self-reported GPA as the criterion measure. While Kuncel, Credé, and Thomas (2005) voice some concern over the

use of self-reported GPAs due to the tendency of some students to inflate their GPA, the researchers' meta-analysis revealed that self-reported GPA had a strong correlation with GPA provided by the college, $r = .90$. Noftle and Robins (2007) found a similar correlation between these two assessments of GPA ($r = .89$), and their study also found negligible differences between the self-reported and school reported GPA in correlations with personality predictors. The present studies were unable to determine the extent to which the self-reported GPAs were inflated, or otherwise inaccurate; however, the participant's name was never attached to their responses in the studies which may have lowered any desire to inflate their scores. The measure of contextual performance may also be a limitation of the present study if the inclusion of respondents other than a direct supervisor of the participant confounded the definition of performance.

Another limitation of the present studies was the use of the NEO-FFI as the measure of Conscientiousness. Roberts, Chernyshenko, Stark, and Goldberg (2005) found that none of our existing measures of Conscientiousness include all aspects of this construct. Despite the inability of the NEO-FFI to assess all components of Conscientiousness, the reasons for including Conscientiousness in the present studies (its validity, incremental validity over cognitive ability, and potential to reduce adverse impact) were all established using imperfect measures of Conscientiousness. The prevalent use of the NEO-FFI in related studies (e.g., Chan, 1997; Chan et al., 1998; Hunthausen et al., 2003; Holtz et al., 2005) provided additional confidence in its ability to allow for adequate testing of the proposed hypotheses and research questions. However, the results relating to the perceptions of job relatedness for the NEO-FFI Conscientiousness scale may be an indication that use of this scale was in fact a

limitation in the present studies. This limitation may be from the lack of inclusion of all components of Conscientiousness, the limited number of items in the measure, or the language used in the items (i.e., work and job).

The use of two different measures of cognitive ability for Study 1 and Study 2 was another limitation of the present study. The Watson-Glaser Form S assesses cognitive ability through an assessment of inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments. The HBRI uses tactical and strategic reasoning to assess cognitive ability. Though these measures assess cognitive ability in different ways, they both have reported similar criterion-related validities. Gadzella, Hogan, Masten, Stacks, Stephens, and Zascavage (2006) reported a validity coefficient of .30 for an untimed administration of the Watson-Glaser Form S while the HBRI has a validity coefficient of .27 with undergraduate GPA. The lack of similar findings in the present study indicates that the use of the different measures was likely a limitation. This limitation was compounded by the mean percentage ratings for the HBRI being lower than those for the Watson-Glaser Form S, a possible indication that the HBRI is more difficult. While the Watson-Glaser Form S could have been altered to a business related context for use in the second study, it was important to use an instrument with established validity.

Future Research

While the present studies did not offer support for the hypothesized relationships, the results of the exploratory analyses indicate it would be beneficial for researchers to replicate the present studies with modifications made to address the aforementioned limitations. In replicating the present studies, future research should consider alternative

ways to measure the participants' motivation. While the use of the TAS Motivation scale has been supported in the existing literature, an alternative measure of motivation may be necessary to overcome the lack of variance found in the present study.

Future research would also benefit from examining the proposed relationships using alternative assessments. Using a more complete measure of Conscientiousness is warranted given the findings related to perceptions of job relatedness in the present studies that were contrary to those in the literature. A complete measure of Conscientiousness would not only ensure that the construct is fully measured, but would clarify for participants the construct being measured that is necessary in forming perceptions of job relatedness. Additionally, a measure of Conscientiousness that includes more than 12 items may be beneficial as it would give participants more information from which their perceptions could be formed. Future research should also consider contextualizing each item of the Conscientiousness measure as opposed to, or in addition to, contextualizing the instructions for the assessment. Contextualization of the cognitive ability measure should also be considered in future research. Given the results of the present studies, it may be beneficial to contextualize each item of a cognitive ability measure in order to avoid potential differences in length and difficulty. While this would likely result in the use of a contextualized measure of cognitive ability that does not have existing evidence of criterion-related validity, researchers could take some precautions by ensuring that the content validity remains unchanged with the contextualization. In addition to considering alternative measures of Conscientiousness and cognitive ability, future research should also consider different types of selection assessments. For example, future research may consider the impact that a selection

interview may have on the proposed relationships given the customary inclusion of interviews in selection systems.

The criterion measures should also be modified in future research. It would be beneficial for future research to identify criterion measures that do not possess the subgroup differences found with GPA in the present studies. Ideally, future research would also be able to obtain a criterion measure that is more conceptually related to the Conscientiousness and Cognitive Ability components of job performance. This would include ensuring that the measurement of contextual performance is not confounded by the evaluation of performance by someone other than a direct supervisor.

In order to better understand the participants' perceptions of the requirements necessary to complete the job, future research should ask participants to identify the knowledge, skills, and abilities necessary for the job described. This would allow for an investigation as to whether Conscientiousness and Cognitive Ability were perceived as important to the job, without regard to the assessments used to measure these constructs. Having the participants identify their perceived requirements may also serve to ensure that those requirements remain salient throughout the assessment process.

In addition to replicating the present studies, future research should also be conducted to determine if the proposed outcomes of the present studies occur when there is a greater time lag between the administration of the cognitive ability and Conscientiousness measures. Future research should also consider the impact of work experience on the proposed relationships. While the present study did ask participants if they had full- or part-time work experience, future research should investigate the impact of work experience by measuring how long the participant has had each type of work

experience, the type of work in which they gained the experience, and whether or not the participant had supervisory work experience. Finally, future research should be conducted to determine if other procedural justice rules impact perceptions of overall fairness and thus have the proposed outcomes on test-taking motivation and validity. Future research should particularly focus on the role of Chance to Perform given the low means found in the present studies.

While there was an overall lack of support for the proposed relationships in the present studies, the results of these studies do add value to our knowledge of perceptions of procedural justice in several important ways. First, the present studies addressed a need to look at how perceptions of procedural justice (job relatedness) influence outcomes of critical importance to organizations (assessment validity) during a typical selection assessment practice (administering two assessments in the same session). Second, the results of the exploratory analyses offer insight into how the proposed relationships may be impacted when individual differences are found on predictor (cognitive ability) or criteria (GPA) measures. Finally, through identifying the limitations of the present studies, a foundation is provided for future research to investigate further the role of procedural justice on outcomes in a selection context.

References

Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test taking. *Personnel Psychology, 43,* 695-716.

Avis, J. M., Kudisch, J. D., & Fortunato, V. J., (2002). Examining the incremental validity and adverse impact of cognitive ability and Conscientiousness on job performance. *Journal of Business and Psychology, 17,* 87-105.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Bauer, T. N., Truxillo, D. M., Paronto, M. E., Weekley, J. A., & Campion, M. A. (2004). Applicant reactions to different selection technology: Face-to-face, interactive voice response, and computer-assisted telephone screening interviews. *International Journal of Selection and Assessment, 12,* 135-148.

Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicant reactions to selection: Development of the Selection Procedural Justice Scale (SPJS). *Personnel Psychology, 54,* 387-419.

Bell, B. S., Ryan, A. M., & Wiechmann, D. (2004). Justice expectations and applicant perceptions. *International Journal of Selection and Assessment, 12,* 24-38.

Bertua, C., Anderson, N., & Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational Psychology, 78,* 387-409.

Bing, M. N., Whanger, J. C., Davison, H. K., & VanHook, J. B. (2004). Incremental validity of the frame-of-reference effect in personality scale scores: A replication and extension. *Journal of Applied Psychology, 89,* 150-157.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt, & W. C. Borman (Eds.), *Personnel Selection in Organizations* (pp. 71-98). San Francisco: Jossey-Bass.

Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences, 44,* 1596-1603.

Chan, D. (1997). Racial subgroup differences in predictive validity perceptions on personality and cognitive ability tests. *Journal of Applied Psychology, 82,* 311-320.

Chan, D. (2007, April). In S. Schinkel (Chair), *Applicant selection fairness expectations and reactions.* Symposium conducted at the meeting of the Society for Industrial and Organizational Psychology, New York, New York.

Chan, D., & Schmitt, N. (2004). An agenda for future research on applicant reactions to selection procedures: A construct-oriented approach. *International Journal of Selection and Assessment, 12,* 9-23.

Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology, 82,* 300-310.

Chan, D., Schmitt, N., Sacco, J. M., & DeShon, R. P. (1998). Understanding pretest and posttest reactions to cognitive ability and personality tests. *Journal of Applied Psychology, 83,* 471-485.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *86,* 278-321.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, *86,* 425-445.

Cropanzano, R. & Wright, T. A. (2003). Procedural justice and organizational staffing: A tale of two paradigms. *Human Resource Management Review, 13,* 7-39.

Edwards, B. D., & Arthur, W. (2007). An examination of factors contributing to a reduction in subgroup differences on a constructed-response paper-and-pencil test of scholastic achievement. *Journal of Applied Psychology, 92,* 794-801.

Gadzella, B. M., Hogan, L., Masten, W., Stacks, J., Stephens, R., & Zascavage, V. (2006). Reliability and validity of the Watson-Glaser Critical Thinking Appraisal-Forms for different academic groups. *Journal of Instructional Psychology, 33,* 141-143.

Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review, 18,* 694-734.

Gilliland, S. W. (1994). Effects of procedural and distributive justice reactions to a selection system. *Journal of Applied Psychology, 79,* 691-701.

Greenberg, J. (1990). Organizational Justice: Yesterday, today, and tomorrow. *Journal of Management, 16,* 399-432.

Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology, 57*, 639-683.

Holtz, B. C., Ployhart, R. E., & Dominguez, A. (2005). Testing the rules of justice: The effects of frame-of-reference and pre-test validity information on personality test responses and test perceptions. *International Journal of Selection and Assessment, 13,* 75-86.

Hülsheger, U. R., Maier, G. W., & Stumpp, T. (2007). Validity of general mental ability for the prediction of job performance and training success in Germany: A meta-analysis. *International Journal of Selection and Assessment, 15,* 3-18.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternate measures of job performance. *Psychological Bulletin, 96*, 72-98.

Hunthausen, J. M., Truxilla, D. M., Bauer, T. N., & Hammer, L. B. (2003). A field study of frame-of-reference effects on personality test validity. *Journal of Applied Psychology, 88,* 545-551.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*, 869-879.

Jöreskog, K. G., & Sörbom, D. (2004). LISREL 8.72 for Windows [Computer Software]. Skokie, IL: Scientific Software International, Inc.

Kuncel, N R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75,* 63-82.

LaHuis, D. M., Perreault, N. E., & Ferguson, M. W. (2003). The effect of legitimizing explanations on applicants' perceptions of selection assessment fairness. *Journal of Applied Social Psychology, 33,* 2198-2215.

LePine, J. A., & Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: Evidence of differential relationships with Big Five personality characteristics and cognitive ability. *Journal of Applied Psychology, 86,* 326-336.

Macan, T. H., Avedon, M. J., Paese, M., & Smith, D. E. (1994). The effects of applicant reactions' to cognitive ability tests and an assessment center. *Personnel Psychology, 47,* 715-738.

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,* 683-729.

Moscoso, S., & Salgado, J. F. (2004). Fairness reactions to personnel selection techniques in Spain and Portugal. *International Journal of Selection and Assessment, 12*, 187-196.

Noftle, N. R., & Robbins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93,* 116-130.

O'Neill, T. A., Goffin, R. D., & Gellatly, I. R. (2010). Test-taking motivation and personality test validity. *Journal of Personnel Psychology, 9,* 117-125.

O*NET. (2008), *Summary report for:43-1011.00 – First-line Supervisors/Managers of office and administrative support workers.* Retrieved October 27, 2008 from http://online.onetcenter.org/link/summary/43-1011.00.

Ployhart, R. E., & Ryan, A. M. (1998). Applicants' reactions to the fairness of selection procedures: The effects of positive rule violations and time of measurement. *Journal of Applied Psychology, 83,* 3-16.

Potosky, D., Bobko, P., & Roth, P. L. (2005). Forming composites of cognitive ability and alternative measures to predict job performance and reduce adverse impact: Corrected estimates and realistic expectations. *International Journal of Selection and Assessment, 13,* 304-315.

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58,* 103-139.

Rosse, J. G., Miller, J. L., & Stecher, M. D. (1994) A field study of job applicants' reactions to personality and cognitive ability testing. *Journal of Applied Psychology, 79,* 987-992.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review, 16,* 155-180.

Ryan, A. M., & Huth, M. (2008). Not much more than platitudes? A critical look at the utility of applicant reactions research. *Human Resource Management Review, 18,* 119-132.

Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decision: A critical review and agenda for the future. *Journal of Management, 26,* 565-606.

Rynes, S. L., & Connerley, M. L. (1993). Applicant reactions to alternative selection procedures. *Journal of Business and Psychology, 7*, 261-277.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology, 82*, 30-43.

Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology, 76,* 323-346.

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., & Fruyt, F. D. (2003). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology, 56*, 573-605.

Sanchez, R. J., Truxillo, D. M., & Bauer, T. N. (2000). Development and examination of an expectancy-based measure of test-taking motivation. *Journal of Applied Psychology, 85,* 739-750.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86,* 162-173.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in Personnel Psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262-274.

Schmit, M. J., & Ryan, A. M. (1992). Test-taking dispositions: A missing link? *Journal of Applied Psychology, 77,* 629-637.

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology, 80,* 607-620.

Schmitt, N., & Chan, D. (1999). The status of research on applicant reactions to selection tests and its implications for managers. *International Journal of Management Reviews, 1,* 45-62.

Smither, J.W., Reilly, R.R., Millsap, R.E., Pearlman, K., & Stoffey, R.W. (1993). Applicant reactions to selection procedures. *Personnel Psychology, 46,* 49-76.

Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology, 81*, 134-141.

Tett, R. P., Jackson, D. N., & Cardy, R. L. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44,* 703-742.

Truxillo, D. M., Bauer, T. N., & Sanchez R. J., (2001). Multiple dimensions of procedural justice: Longitudinal effects on selection system fairness and test-taking self-efficacy. *International Journal of Selection and Assessment, 9,* 336-349.

Truxillo, D. M., Steiner, D. D., & Gilliland, S. W. (2004). The importance of organizational justice in personnel selection: Defining when selection fairness really matters. *International Journal of Selection and Assessment, 12,* 39-53.

University of Missouri – St. Louis. (2008). *Student profile Fall 2007.* Retrieved October 27, 2008 from http://www.umsl.edu/services/academic/student/sprofile.html.

Figure 1

Gilliland's Model of Procedural Justice in the Workplace (1993, p. 700)

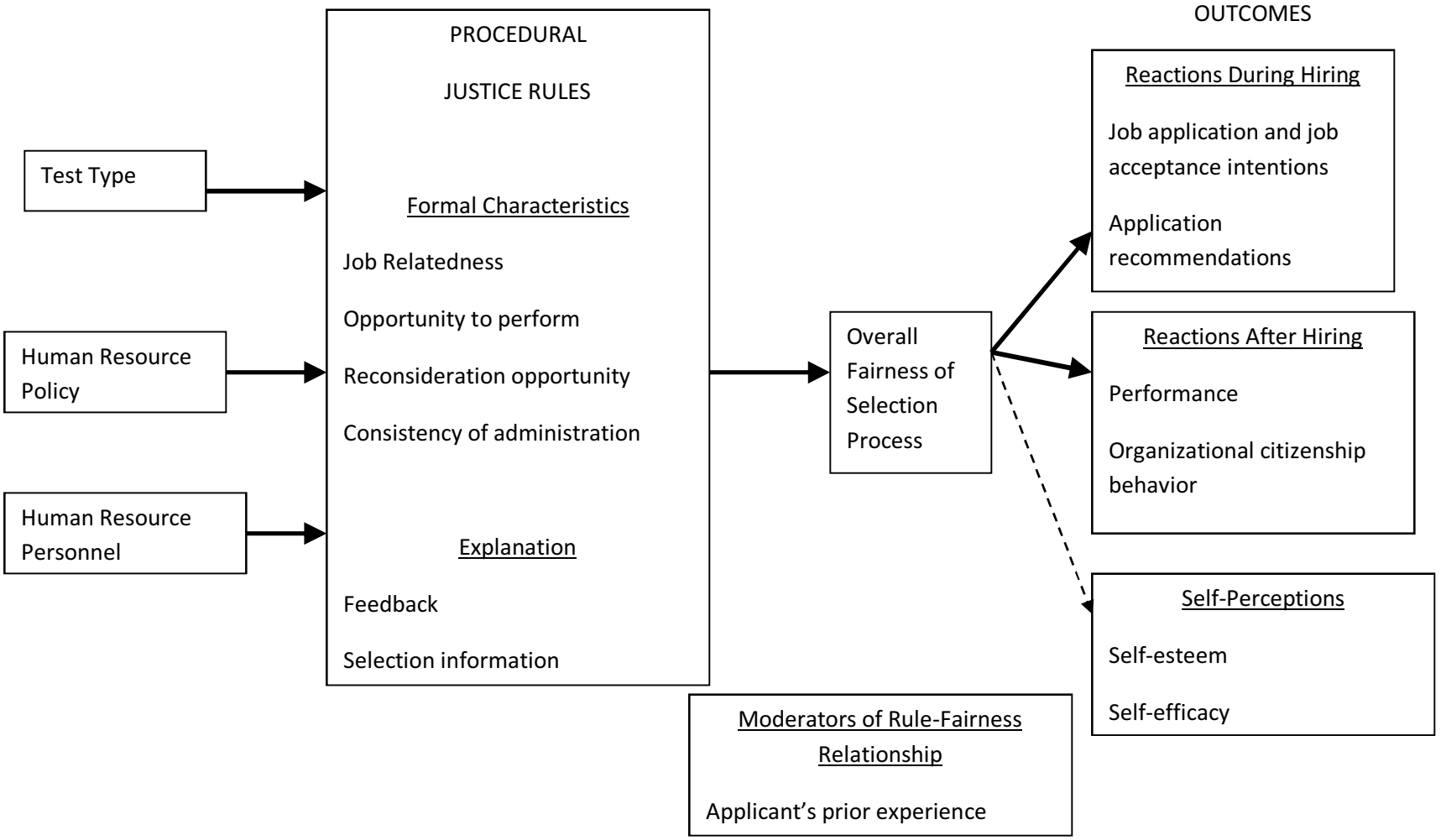


Figure 2

Aspect of Gilliland's (1993) Model and an Extension Tested in the Present Study

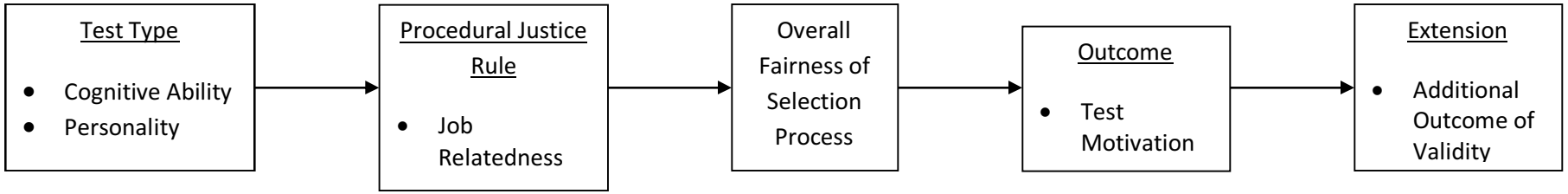


Figure 3

Model of Participant Reactions to Multiple Assessments in the Same Session

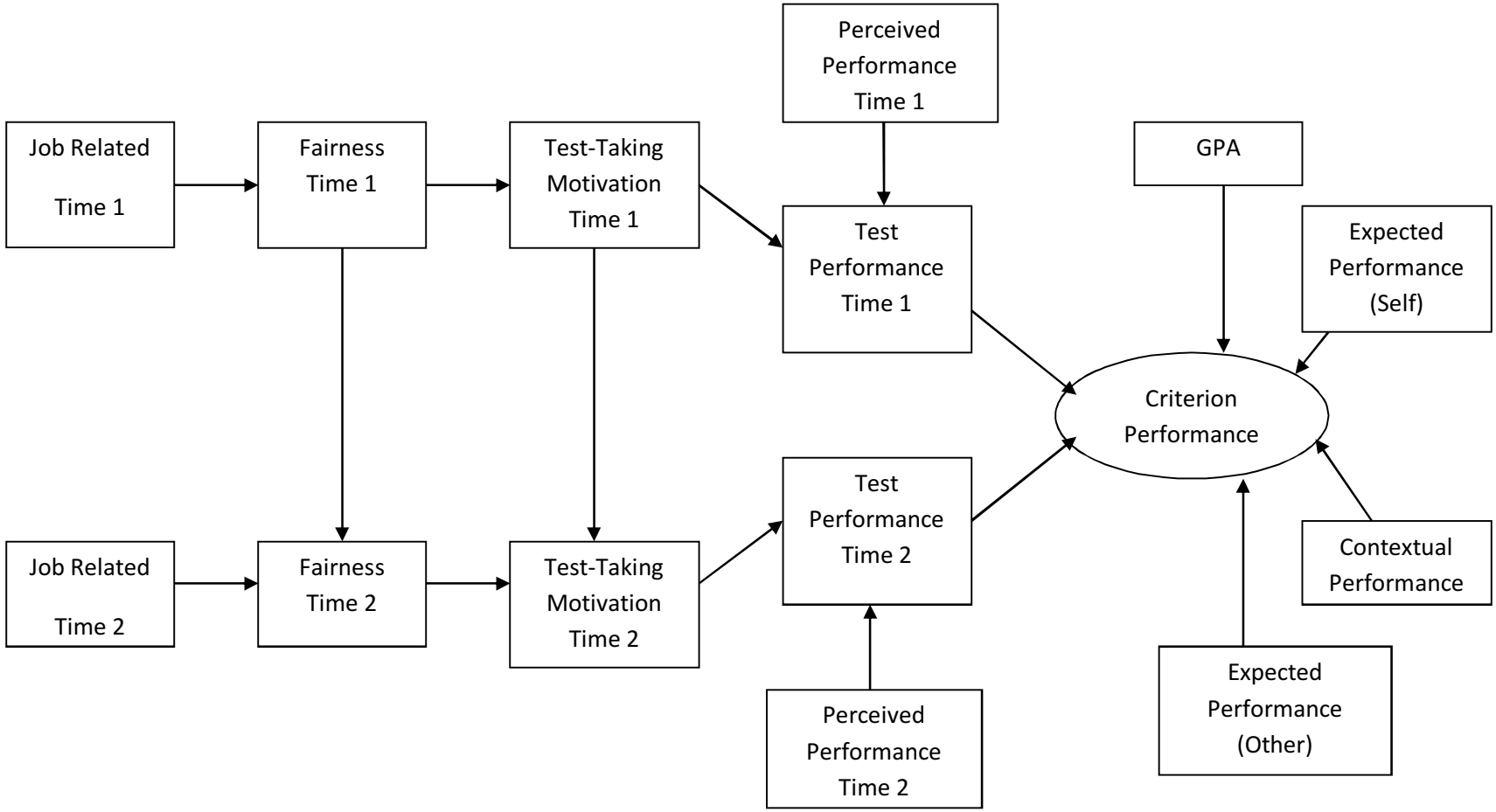
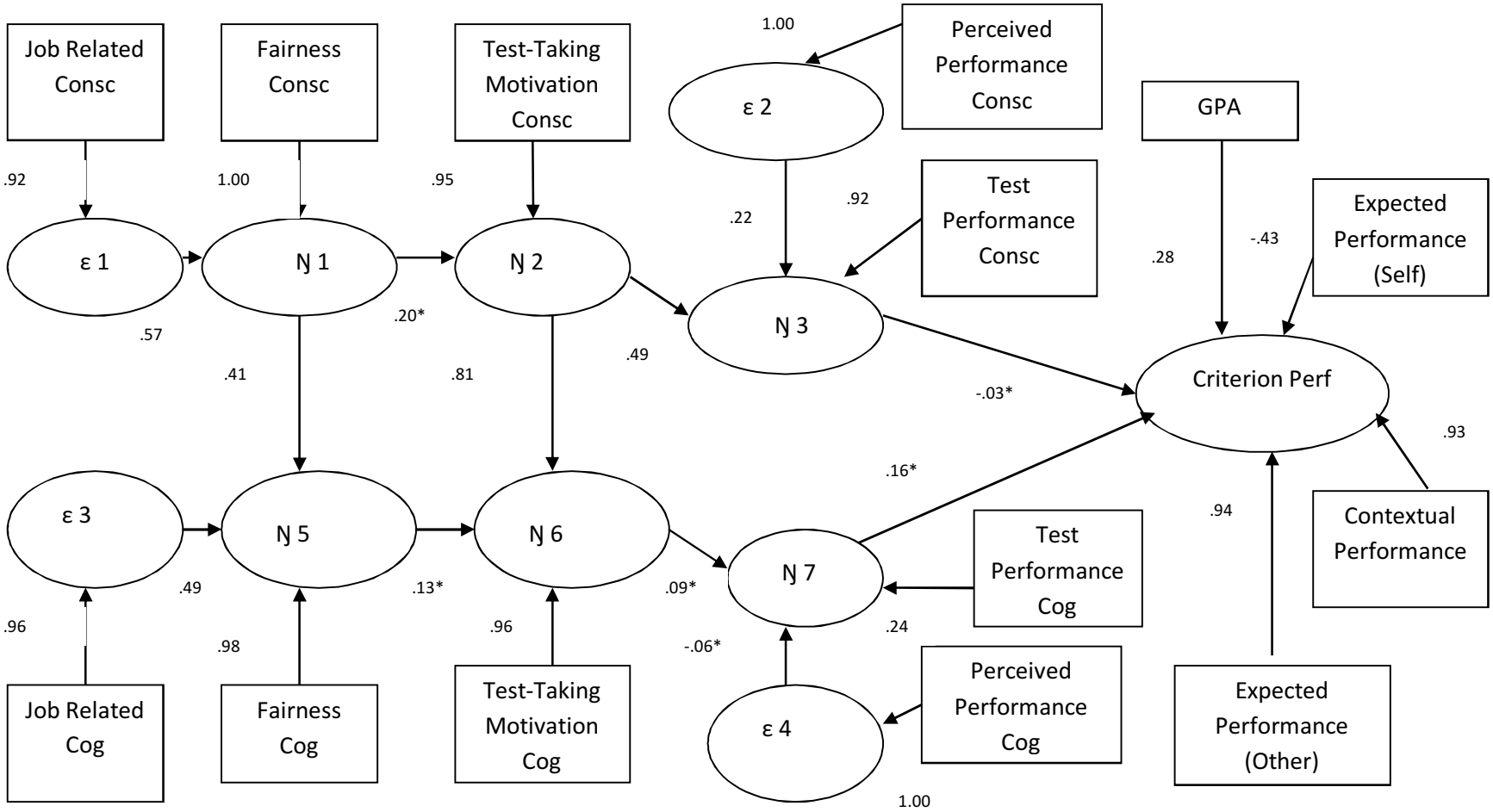


Figure 4

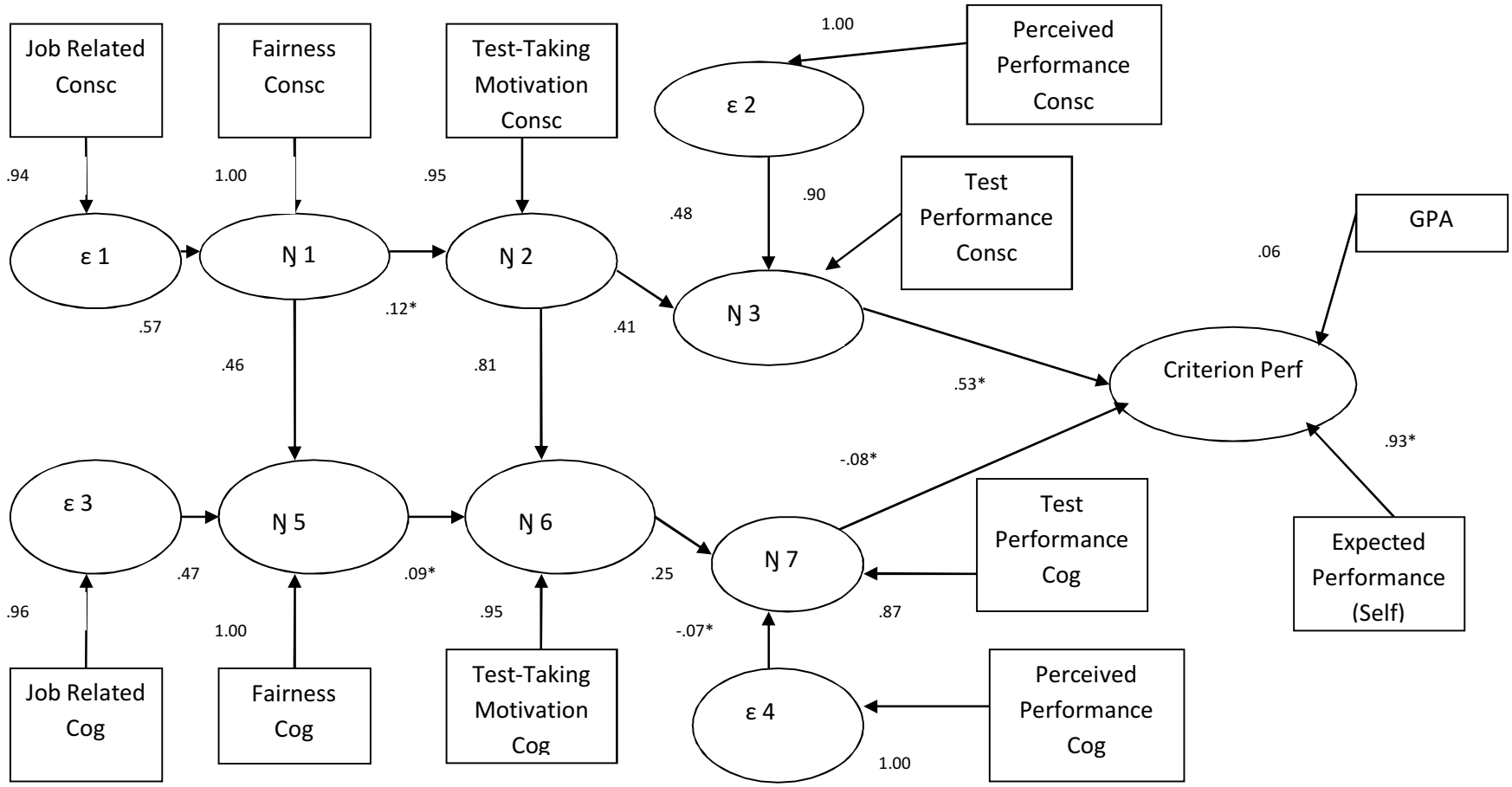
Path Estimates for the Model of Conscientiousness First Participant Reactions to Multiple Assessments in the Same Session



* $p < .05$

Figure 5

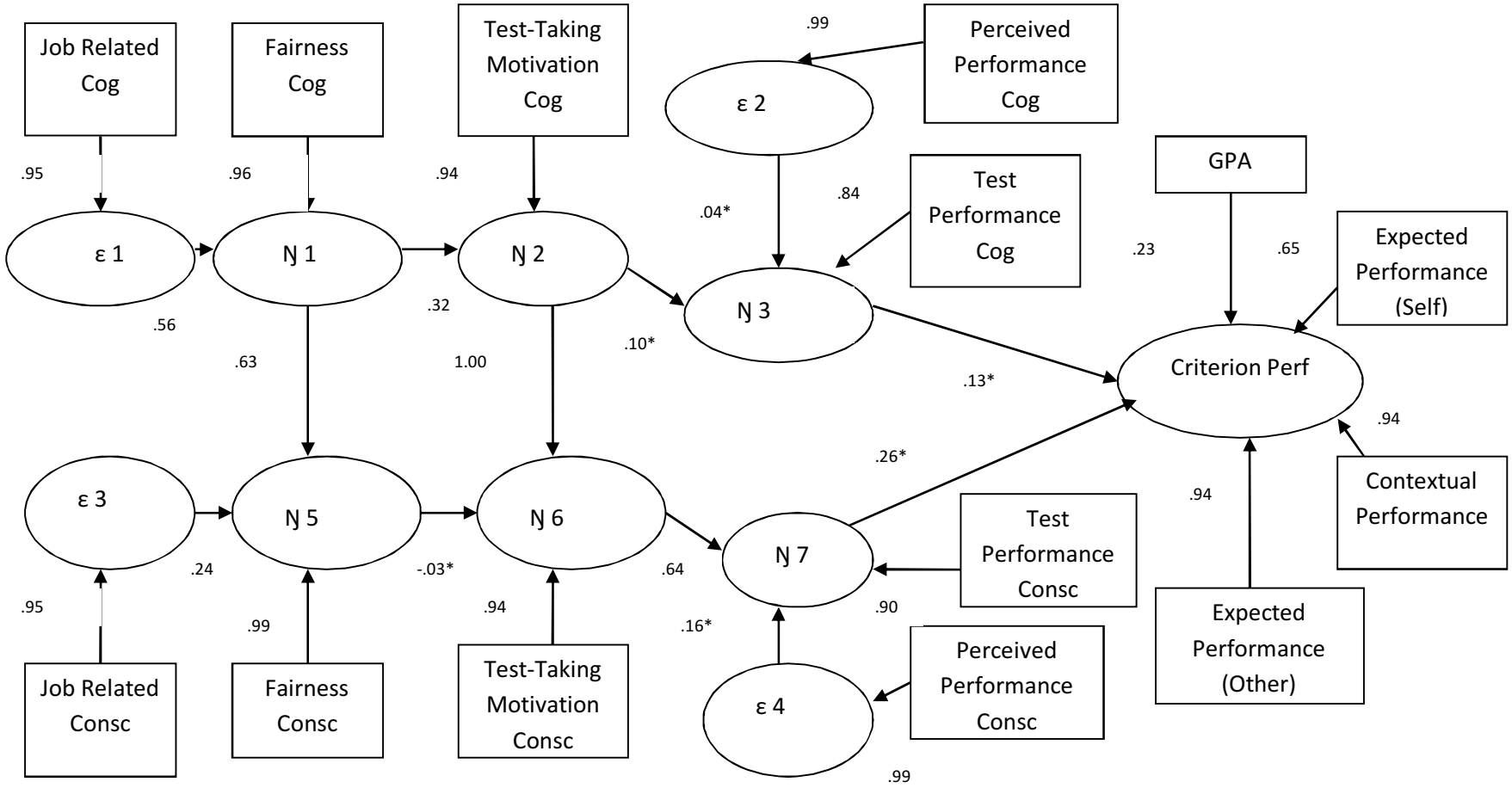
Path Estimates for the Truncated Model of Conscientiousness First Participant Reactions to Multiple Assessments in the Same Session



* $p < .05$

Figure 6

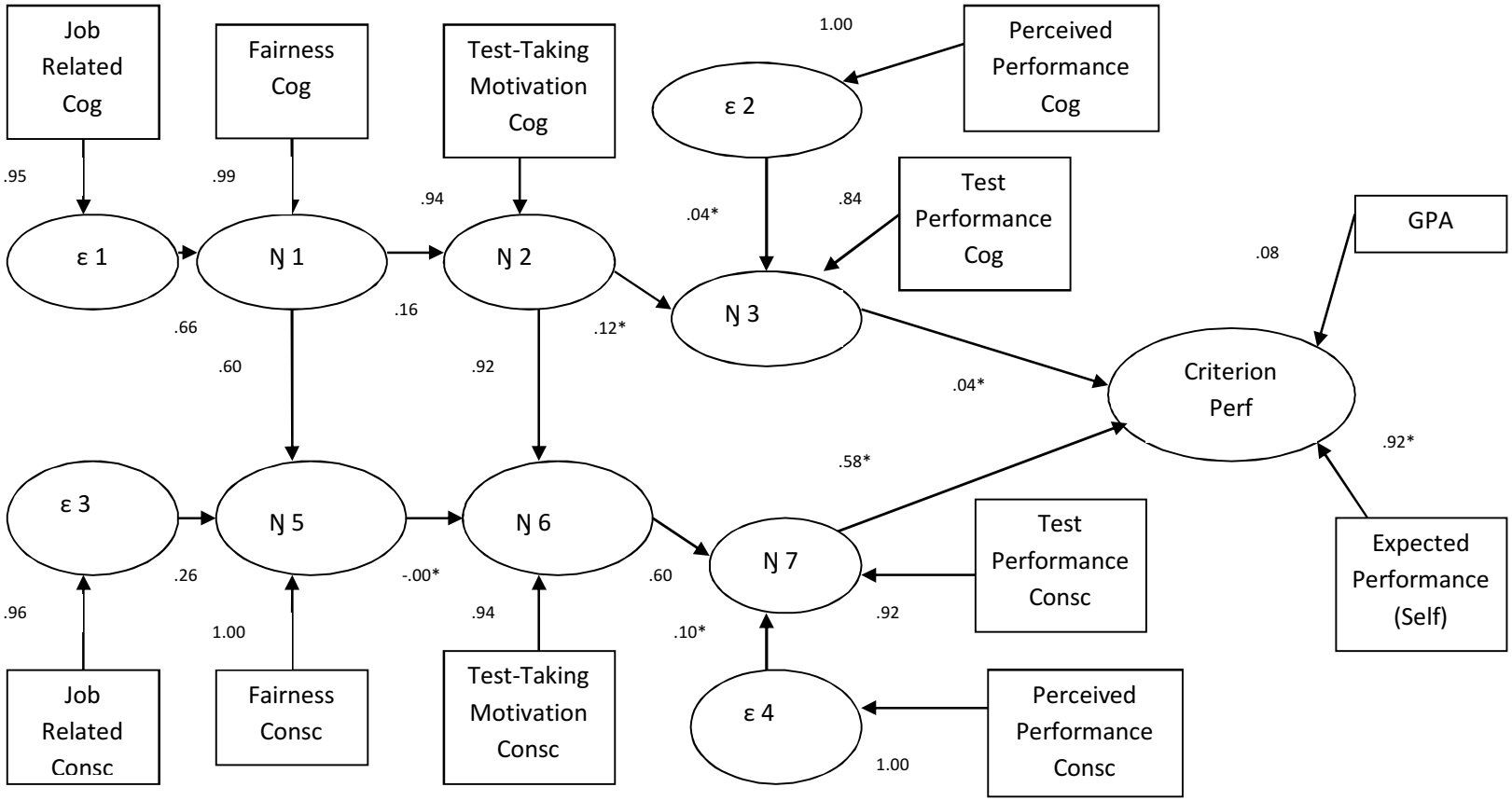
Path Estimates for the Model of Cognitive Ability First Participant Reactions to Multiple Assessments in the Same Session



* $p < .05$

Figure 7

Path Estimates for the Truncated Model of Cognitive Ability First Participant Reactions to Multiple Assessments in the Same Session



* $p < .05$

Table 1

Comparison of Face Validity and Job Relatedness Scales

Face Validity Scale[1]	Job-Relatedness Content Scale[2]
It would be obvious to anyone that the examination is related to the job.	**It would be clear to anyone that this test is related to the [*insert job title*] job.**
The actual content of the examination was clearly related to the job.	**The content of the test was clearly related to the [*insert job title*] job.**
I did not understand what the examination had to do with the job.	
I could not see any relationship between the examination and what is required on the job.	
There was no real connection between the examination I went through and the job.	

1. Smither et al. (1993, p.75)

2. Bauer et al. (2001, p.418-419)

Table 2

Correlations Between Perceptions of Job Relatedness and Fairness (Chan et al., 1998)

	Cognitive Ability	Personality
Pre-Test		
Face Validity	.40	.37
Predictive Validity	.59	.59
Post-Test		
Face Validity	.42	.42
Predictive Validity	.64	.62

Table 3

Participant Demographics by Group

	Gender			Race						
	Female	Male	Not Reported	Caucasian	African American	Hispanic	Asian	Native American	Other	Not Reported
Conscientiousness First	114	58	0	103	40	4	22	1	2	0
Cognitive Ability First	110	65	0	99	42	1	29	0	3	1

Table 4

Means, Standard Deviations, and Reliability Coefficients by Group

	Conscientiousness First				Cognitive Ability First			
Scale	*N*	*M*	*SD*	*α*	*N*	*M*	*SD*	*α*
TAS Initial†	169	6.56	.59	.88	173	6.44	.68	.86
TAS Conscientiousness†	171	6.50	.66	.90	173	6.36	.76	.90
TAS Cognitive Ability†	172	6.44	.74	.91	174	6.34	.74	.89
Job Relatedness Conscientiousness	172	3.34	.67	.88	174	2.84	.82	.92
Job Relatedness Cognitive Ability	172	2.76	.80	.92	175	2.77	.79	.91
Fairness Conscientiousness	172	3.27	.94	.67	174	3.20	.97	.81
Fairness Cognitive Ability	172	3.20	1.06	.83	175	2.95	.94	.73
Perceived Performance Conscientiousness	171	4.40	.61	.85	175	4.08	.79	.93
Perceived Performance Cognitive Ability	172	4.15	.77	.89	175	4.05	.73	.89
NEO-FFI Conscientiousness Scale	172	40.16	5.87	.85	175	37.04	6.74	.84
Watson-Glaser Form S	172	24.73	5.68	.75	175	24.83	5.21	.71
Expected Performance – Self	172	4.30	.54	.87	175	4.19	.56	.86
Expected Performance – Other	83	4.14	.62	.92	83	4.19	.60	.90
Contextual Performance	84	4.31	.58	.90	91	4.38	.54	.90
GPA	164	3.17	.47	NA	171	3.14	.51	NA

Table 4 continued

	Conscientiousness First				Cognitive Ability First			
Scale	*N*	*M*	*SD*	*α*	*N*	*M*	*SD*	*α*
Conscientiousness Context								
At work	172	4.30	1.09	NA	175	3.89	1.35	NA
At school	172	4.39	.96	NA	175	4.21	1.12	NA
At home	172	4.25	1.01	NA	175	3.92	1.27	NA
With friends	172	3.84	1.26	NA	175	3.65	1.30	NA
Everyday	171	4.42	.87	NA	175	4.11	1.11	NA

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 5

Researcher Differences for All Variables by Group

	Researcher 1		Researcher 2			
Scale	*M*	*SD*	*M*	*SD*	*df*	*t*
TAS Initial†						
Conscientiousness First	6.55	.59	6.60	.63	167	-.29
Cognitive Ability First	6.43	.68	6.54	.79	171	-.49
TAS Conscientiousness†						
Conscientiousness First	6.49	.66	6.54	.72	169	-.24
Cognitive Ability First	6.36	.74	6.31	1.09	171	.21
TAS Cognitive Ability†						
Conscientiousness First	6.45	.75	6.59	.55	170	-.65
Cognitive Ability First	6.33	.74	6.51	.74	172	-.73
Job Relatedness Conscientiousness						
Conscientiousness First	3.35	.67	3.23	.65	170	.65
Cognitive Ability First	2.83	.83	2.93	.67	172	-.36
Job Relatedness Cognitive Ability						
Conscientiousness First	2.77	.80	2.62	.85	170	.67
Cognitive Ability First	2.76	.79	2.95	.75	173	-.73
Fairness Conscientiousness						
Conscientiousness First	3.27	.94	3.33	.97	170	-.25
Cognitive Ability First	3.18	.98	3.47	.72	172	-.91
Fairness Cognitive Ability						
Conscientiousness First	3.14	1.07	3.83	.65	170	-2.37*
Cognitive Ability First	2.93	.95	3.13	.69	173	-.66
Perceived Performance Conscientiousness						
Conscientiousness First	4.40	.61	4.43	.67	169	-.18
Cognitive Ability First	4.07	.80	4.40	.47	173	-1.30
Perceived Performance Cognitive Ability						
Conscientiousness First	4.14	.78	4.36	.53	170	-1.03
Cognitive Ability First	4.03	.73	4.48	.58	173	-1.90
NEO-FFI Conscientiousness Scale						
Conscientiousness First	40.15	5.85	40.29	6.32	170	-.08
Cognitive Ability First	37.07	6.64	36.50	8.68	173	.26
Watson-Glaser Form S						
Conscientiousness First	24.79	5.60	24.07	6.71	170	.45
Cognitive Ability First	24.81	5.11	52.20	7.04	173	-.23

Table 5 continued

	Researcher 1		Researcher 2			
Scale	*M*	*SD*	*M*	*SD*	*df*	*t*
Expected Performance – Self						
Conscientiousness First	4.30	.54	4.27	.56	170	.20
Cognitive Ability First	4.20	.53	4.05	1.00	173	.83
Expected Performance – Other						
Conscientiousness First	4.13	.63	4.28	.57	81	-.63
Cognitive Ability First	4.19	.60	4.27	.75	81	-.23
Contextual Performance						
Conscientiousness First	4.29	.59	4.53	.52	82	-1.10
Cognitive Ability First	4.37	.55	4.73	.25	89	-1.15
GPA						
Conscientiousness First	3.18	.45	3.08	.64	162	.70
Cognitive Ability First	3.16	.50	2.81	.57	169	2.09*

* $p < .05$

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 6

University Differences for All Variables by Group

	UMSL		UNCC			
Scale	*M*	*SD*	*M*	*SD*	*df*	*t*
TAS Initial†						
Conscientiousness First	6.56	.59	6.52	.59	167	.22
Cognitive Ability First	6.44	.68	6.42	.77	171	.07
TAS Conscientiousness†						
Conscientiousness First	6.49	.67	6.56	.59	169	-.30
Cognitive Ability First	6.36	.76	6.39	.73	171	-.17
TAS Cognitive Ability†						
Conscientiousness First	6.47	.74	6.40	.75	170	.29
Cognitive Ability First	6.35	.73	6.34	.88	172	.03
Job Relatedness Conscientiousness						
Conscientiousness First	3.34	.67	3.36	.63	170	-.12
Cognitive Ability First	2.84	.82	2.80	.91	172	.19
Job Relatedness Cognitive Ability						
Conscientiousness First	2.79	.80	2.36	.75	170	1.71
Cognitive Ability First	2.80	.79	2.43	.70	173	1.63
Fairness Conscientiousness						
Conscientiousness First	3.29	.92	2.97	1.11	170	1.11
Cognitive Ability First	3.20	1.00	3.19	.56	172	.00
Fairness Cognitive Ability						
Conscientiousness First	3.22	1.03	2.82	1.47	170	1.23
Cognitive Ability First	2.96	.95	2.77	.80	173	.70
Perceived Performance Conscientiousness						
Conscientiousness First	4.42	.60	4.16	.71	169	1.36
Cognitive Ability First	4.10	.80	3.92	.67	173	.76
Perceived Performance Cognitive Ability						
Conscientiousness First	4.21	.69	3.39	1.35	170	3.54**
Cognitive Ability First	4.07	.73	3.89	.75	173	.87
NEO-FFI Conscientiousness Scale						
Conscientiousness First	40.17	5.80	40.09	7.13	170	.04
Cognitive Ability First	36.91	6.79	38.69	6.13	173	-.92
Watson-Glaser Form S						
Conscientiousness First	24.72	5.74	24.91	4.99	170	-.11
Cognitive Ability First	24.81	5.18	25.00	5.80	173	-.12

Table 6 continued

	UMSL		UNCC			
Scale	*M*	*SD*	*M*	*SD*	*df*	*t*
Expected Performance – Self						
Conscientiousness First	4.31	.53	4.10	.59	170	1.27
Cognitive Ability First	4.20	.55	4.11	.70	173	.57
Expected Performance – Other						
Conscientiousness First	4.14	.63	4.23	.15	81	-.26
Cognitive Ability First	4.21	.60	3.78	.70	81	1.42
Contextual Performance						
Conscientiousness First	4.29	.58	4.87	.15	82	-1.71
Cognitive Ability First	4.40	.53	4.02	.76	89	1.53
GPA						
Conscientiousness First	3.17	.47	3.22	.46	162	-.39
Cognitive Ability First	3.13	.52	3.21	.42	169	-.56

** $p < .01$

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 7

Means and Standard Deviations for Relationship to Participant by Expected Performance and Contextual Performance

	Conscientiousness First		Cognitive Ability First	
	M	*SD*	*M*	*SD*
Expected Performance – Other				
Supervisor	4.05	.62	4.26	.62
Instructor	3.96	.97	3.51	.54
Co-Worker	4.37	.39	4.26	.61
Classmate	4.60	.45	4.20	.44
Other	4.50	.50	4.45	.07
Contextual Performance				
Supervisor	4.24	.61	4.49	.46
Instructor	4.30	.67	3.91	.69
Co-Worker	4.49	.37	4.36	.58
Classmate	4.57	.41	4.40	.42
Other	4.60	.69	4.20	1.04

Table 8

ANOVA Results for Expected Performance and Contextual Performance as a Function of Relationship to the Participant

	Conscientiousness First					Cognitive Ability First				
	df	*SS*	*MS*	*F*	η^2	*df*	*SS*	*MS*	*F*	η^2
Expected Performance – Other										
Between Groups	4	2.78	.70	1.87	.09	4	3.60	.90	2.67*	.12
Within Groups	78	28.96	.37			78	26.28	.34		
Contextual Performance										
Between Groups	4	1.33	.33	.98	.05	4	2.91	.73	2.64*	.11
Within Groups	79	26.69	.38			86	23.73	.28		

* $p < .05$

Table 9

Correlations between TAS – Motivation Scale Given at Three Times

	1.	2.	3.
1. TAS Initial	-	.64***	.64***
2. TAS Conscientiousness	.87***	-	.83***
3. TAS Cognitive Ability	.71***	.74***	-

*** $p < .001$

Conscientiousness First is presented below the diagonal and Cognitive Ability First is presented above the diagonal.

Table 10

Means and Standard Deviations for TAS Motivation Scale

	Conscientiousness First		Cognitive Ability First	
	M	*SD*	*M*	*SD*
TAS Conscientiousness	6.50	.66	6.36	.76
TAS Cognitive Ability	6.46	.74	6.34	.74

Table 11

ANOVA Results for TAS Motivation as a Function of Assessment Order

	df	*SS*	*MS*	*F*	η^2
TAS Conscientiousness					
Between Groups	1	1.64	1.64	3.26	.01
Within Groups	342	172.21	.50		
TAS Cognitive Ability					
Between Groups	1	1.22	1.22	2.25	.01
Within Groups	344	186.17	.54		

Table 12

Correlations Between Assessment Scores and Criteria Measures by Assessment Order

	Conscientiousness First			Cognitive Ability First		
	r	95% CI		*r*	95% CI	
NEO-FFI Conscientiousness Scale						
GPA	.16	.00	.30	.15	.00	.29
Contextual Performance	.09	-.13	.30	.30	.10	.48
Watson-Glaser Form S						
GPA	.25	.10	.38	.34	.20	.47
Contextual Performance	.24	.03	.43	.07	-.14	.27

Table 13
Means and Standard Deviations for Conscientiousness First Based on Race

	M	*SD*
GPA		
White	3.22	.46
Black	2.91	.43
Asian	3.47	.35
Other	3.11	.48
Watson-Glaser Form S		
White	26.21	5.63
Black	21.28	4.92
Asian	23.73	4.85
Other	25.86	4.26
NEO-FFI Conscientiousness Scale		
White	40.79	5.42
Black	41.00	6.07
Asian	36.41	6.14
Other	38.00	6.43
TAS Initial†		
White	6.56	.56
Black	6.74	.43
Asian	6.18	.83
Other	6.63	.50
TAS Conscientiousness†		
White	6.47	.65
Black	6.68	.55
Asian	6.21	.85
Other	6.64	.42
Perceived Performance Cognitive Ability		
White	4.22	.71
Black	4.29	.69
Asian	3.89	.75
Other	3.25	1.26
Expected Performance – Self		
White	4.36	.46
Black	4.39	.59
Asian	3.93	.61
Other	4.03	.54

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 14

Conscientiousness First ANOVA Results for Study Variables as a Function of Race

	df	*SS*	*MS*	*F*	η^2
GPA					
Between Groups	3	4.51	1.50	7.70***	.13
Within Groups	160	31.28	.20		
Watson-Glaser Critical Thinking Appraisal – Form S					
Between Groups	3	735.20	245.07	8.61***	.13
Within Groups	168	4782.50	28.47		
NEO-FFI: Conscientiousness					
Between Groups	3	410.82	136.94	4.20**	.07
Within Groups	168	5474.62	32.59		
TAS Initial†					
Between Groups	3	4.51	1.50	4.62**	.08
Within Groups	165	53.63	.33		
TAS Conscientiousness†					
Between Groups	3	3.35	1.12	2.64*	.05
Within Groups	167	70.64	.42		
Perceived Performance Cognitive Ability					
Between Groups	3	8.51	2.84	5.16**	.08
Within Groups	168	92.29	.55		
Expected Performance – Self					
Between Groups	3	4.28	1.43	5.29**	.09
Within Groups	168	45.32	.27		

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

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 15

Means and Standard Deviations for Cognitive Ability First Based on Race

	M	*SD*
GPA		
White	3.20	.50
Black	2.86	.53
Asian	3.30	.37
Other	3.18	.54
Watson-Glaser Form S		
White	26.24	5.18
Black	22.12	4.73
Asian	23.93	4.25
Other	24.50	6.95
Expected Performance - Other		
White	4.21	.58
Black	3.85	.55
Asian	4.33	.65
Other	4.95	.07
Contextual Performance		
White	4.44	.50
Black	4.05	.60
Asian	4.42	.56
Other	4.85	.21

Table 16

Cognitive Ability First ANOVA Results for Study Variables as a Function of Race

	df	*SS*	*MS*	*F*	η^2
GPA					
Between Groups	3	4.31	1.44	5.94***	.10
Within Groups	166	40.16	.24		
Watson-Glaser Critical Thinking Appraisal – Form S					
Between Groups	3	530.03	176.68	7.15***	.11
Within Groups	170	4199.45	24.70		
Expected Performance – Other					
Between Groups	3	3.19	1.06	3.16*	.11
Within Groups	78	26.18	.34		
Contextual Performance					
Between Groups	3	2.41	.81	2.87*	.09
Within Groups	86	24.13	.28		

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

Table 17

ANOVA Results for TAS Motivation as a Function of Assessment Order for White Participants

	df	*SS*	*MS*	*F*	η^2
TAS Conscientiousness					
Between Groups	1	1.08	1.08	2.10	.01
Within Groups	198	101.71	.51		
TAS Cognitive Ability					
Between Groups	1	.95	.95	1.93	.01
Within Groups	199	97.82	.49		

Table 18

Correlations between Assessment Scores and Criteria Measures by Assessment Order for White Participants

	Conscientiousness First			Cognitive Ability First		
	r	95% CI		*r*	95% CI	
NEO-FFI Conscientiousness Scale						
GPA	.26**	.07	.44	.23*	.03	.41
Contextual Performance	.18	-.10	.43	.52***	.30	.69
Watson-Glaser Form S						
GPA	.22*	.03	.40	.30**	.11	.47
Contextual Performance	.19	-.08	.44	-.01	-.27	.25

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

Table 19

Means, Standard Deviations, and Coefficient Alphas for Exploratory Procedural Justice Subscales

	Conscientiousness First				Cognitive Ability First			
Scale	*N*	*M*	*SD*	*α*	*N*	*M*	*SD*	*α*
Information Known Conscientiousness	172	2.98	1.04	.77	173	2.76	1.12	.87
Chance to Perform Conscientiousness	172	2.13	.98	.90	174	2.21	1.03	.90
Consistency Conscientiousness	172	4.44	.74	.85	175	4.50	.77	.88
Treatment Conscientiousness	172	4.68	.46	.86	175	4.72	.40	.81
Propriety of Questions Conscientiousness	171	4.46	.66	.83	175	4.38	.68	.71
Information Known Cognitive Ability	172	2.91	1.11	.85	175	2.52	1.00	.72
Chance to Perform Cognitive Ability	171	2.21	.94	.91	175	2.04	.95	.88
Consistency Cognitive Ability	172	4.42	.80	.94	175	4.43	.77	.84
Treatment Cognitive Ability	172	4.65	.51	.86	174	4.67	.44	.74
Propriety of Questions Cognitive Ability	170	4.35	.71	.72	174	4.25	.74	.72

Table 20

Correlations between Exploratory Procedural Justice Scales and Study Variables – Conscientiousness First

	1.	2.	3.	4.	5.	6.
1. GPA						
2. Contextual Performance	.32**					
3. Expected Performance - Self	.05	-.11				
4. Expected Performance - Other	.13	.82**	-.11			
5. Watson-Glaser Form S	.25**	.24*	.00	.18		
6. NEO-FFI Conscientiousness Scale	.16*	.09	.44**	-.04	.13	
7. TAS Initial†	-.06	.00	.29**	-.04	-.01	.50**
8. TAS Conscientiousness†	-.03	.04	.31**	-.03	.06	.51**
9. Perceived Performance Conscientiousness	.01	-.05	.34**	.09	.02	.52**
10. Job Relatedness Conscientiousness	.14	.12	.12	.09	.06	.20**
11. Fairness Conscientiousness	.04	.01	.17*	-.04	-.07	.16*
12. TAS Cognitive Ability†	-.01	.04	.22**	-.06	.18*	.52**
13. Perceived Performance Cognitive Ability	-.06	-.03	.25**	-.04	.07	.45**
14. Job Relatedness Cognitive Ability	.00	.01	.11	.01	.01	.12
15. Fairness Cognitive Ability	.06	-.03	.05	-.06	.03	.14
16. Information Known Conscientiousness	-.13	.00	.03	-.03	-.13	.04
17. Chance to Perform Conscientiousness	-.02	.09	.10	.02	-.22**	-.08
18. Consistency Conscientiousness	.03	-.12	.30**	-.07	.16*	.29**
19. Treatment Conscientiousness	.01	.08	.42**	.11	.18*	.47**
20. Propriety of Questions Conscientiousness	-.01	-.01	.23**	.02	.02	.26**
21. Information Known Cognitive Ability	-.09	.03	-.08	-.01	-.14	-.01
22. Chance to Perform Cognitive Ability	.07	-.08	-.03	-.06	-.01	-.04
23. Consistency Cognitive Ability	.09	-.09	.27**	-.03	.24**	.38**
24. Treatment Cognitive Ability	-.03	-.06	.41**	.04	.13	.49**
25. Propriety of Questions Cognitive Ability	-.05	.02	.26**	-.02	.00	.26**

Table 20 continued

	7.	8.	9.	10.	11.	12.
8. TAS Conscientiousness	.87**					
9. Perceived Performance Conscientiousness	.45**	.44**				
10. Job Relatedness Conscientiousness	0.15	.24**	.23**			
11. Fairness Conscientiousness	.08	.10	.16*	.53**		
12. TAS Cognitive Ability	.71**	.74**	.44**	.22**	.12	
13. Perceived Performance Cognitive Ability	.29**	.31**	.53**	.16*	.28**	.55**
14. Job Relatedness Cognitive Ability	.10	.14	.16*	.53**	.39**	.15*
15. Fairness Cognitive Ability	.08	.11	.10	.34**	.61**	.17*
16. Information Known Conscientiousness	.03	.07	.09	.25**	.16*	-.04
17. Chance to Perform Conscientiousness	-.02	.08	-.03	.33**	.32**	.00
18. Consistency Conscientiousness	.32**	.27**	.28**	.12	.11	.30**
19. Treatment Conscientiousness	.35**	.36**	.41**	.24**	.11	.34**
20. Propriety of Questions Conscientiousness	.30**	.29**	.21**	.31**	.32**	.25**
21. Information Known Cognitive Ability	-.03	.01	.02	.19*	.06	-.03
22. Chance to Perform Cognitive Ability	-.03	-.02	.00	.30**	.27**	.06
23. Consistency Cognitive Ability	.27**	.27**	.35**	.24**	.26**	.27**
24. Treatment Cognitive Ability	.45**	.44**	.36**	.15	.06	.42**
25. Propriety of Questions Cognitive Ability	.22**	.24**	.23**	.24**	.29**	.34**

Table 20 continued

	13.	14.	15.	16.	17.	18.
14. Job Relatedness Cognitive Ability	.14					
15. Fairness Cognitive Ability	.20**	.63**				
16. Information Known Conscientiousness	-.03	.23**	.11			
17. Chance to Perform Conscientiousness	.00	.38**	.24**	.20**		
18. Consistency Conscientiousness	.12	.06	.15	-.03	-.21**	
19. Treatment Conscientiousness	.24**	.14	.04	.05	-.19*	.48**
20. Propriety of Questions Conscientiousness	.17*	.18*	.29**	.01	-.04	.49**
21. Information Known Cognitive Ability	.02	.27**	.14	.73**	.10	-.01
22. Chance to Perform Cognitive Ability	.06	.55**	.41**	.02	.51**	-.08
23. Consistency Cognitive Ability	.26**	.22**	.28**	.08	-.10	.62**
24. Treatment Cognitive Ability	.26**	.10	.06	.01	-.21**	.44**
25. Propriety of Questions Cognitive Ability	.28**	.37**	.34**	.05	.00	.39**

	19.	20.	21.	22.	23.	24.
20. Propriety of Questions Conscientiousness	.49**					
21. Information Known Cognitive Ability	.02	.03				
22. Chance to Perform Cognitive Ability	-.16*	.01	.09			
23. Consistency Cognitive Ability	.45**	.42**	.10	.05		
24. Treatment Cognitive Ability	.85**	.44**	.02	-.17*	.52**	
25. Propriety of Questions Cognitive Ability	.47**	.68**	.13	.15*	.51**	.49**

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

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 21

Correlations between Exploratory Procedural Justice Scales and Study Variables – Cognitive Ability First

	1.	2.	3.	4.	5.	6.
1. GPA						
2. Contextual Performance	.23*					
3. Expected Performance - Self	.07	.26*				
4. Expected Performance - Other	.19	.77**	.26*			
5. Watson-Glaser Form S	.34**	.07	.10	.08		
6. NEO-FFI Conscientiousness Scale	.15*	.30**	.51**	.19	.14	
7. TAS Initial†	-.00	.19	.18*	.21	.09	.30**
8. TAS Conscientiousness†	.08	.14	.33**	.11	.07	.55**
9. Perceived Performance Conscientiousness	.06	.06	.33**	-.01	.16*	.42**
10. Job Relatedness Conscientiousness	.05	.05	.19*	.20	.17*	.16*
11. Fairness Conscientiousness	.18*	.03	.10	.01	.24**	.25**
12. TAS Cognitive Ability†	.09	.10	.30**	.08	.09	.48**
13. Perceived Performance Cognitive Ability	.02	.05	.38**	.02	.13	.32**
14. Job Relatedness Cognitive Ability	.09	.01	.14	.08	.06	.08
15. Fairness Cognitive Ability	.14	-.02	.10	.01	.23**	.19*
16. Information Known Conscientiousness	-.01	.06	.15*	.11	.05	.18*
17. Chance to Perform Conscientiousness	.15	-.09	.18*	-.05	-.05	.14
18. Consistency Conscientiousness	.04	.04	.15*	-.18	.15*	.09
19. Treatment Conscientiousness	.13	.13	.22**	.09	.08	.18*
20. Propriety of Questions Conscientiousness	.03	.20	.24**	.22*	.17*	.25**
21. Information Known Cognitive Ability	.02	-.05	.07	.00	.05	.14
22. Chance to Perform Cognitive Ability	.15*	-.10	.16*	-.10	-.06	.10
23. Consistency Cognitive Ability	.07	-.11	.08	-.21	.21**	.10
24. Treatment Cognitive Ability	.14	.15	.09	.08	.15*	.17*
25. Propriety of Questions Cognitive Ability	-.04	.10	.21**	.08	.15	.19*

Table 21 continued

	7.	8.	9.	10.	11.	12.
8. TAS Conscientiousness†	.64**					
9. Perceived Performance Conscientiousness	.19*	.35**				
10. Job Relatedness Conscientiousness	.06	.15	.19*			
11. Fairness Conscientiousness	.10	.13	.20**	.54**		
12. TAS Cognitive Ability†	.64**	.83**	.35**	.12	.16*	
13. Perceived Performance Cognitive Ability	.26**	.30**	.74**	.14	.18*	.39**
14. Job Relatedness Cognitive Ability	.03	.15	.16*	.65**	.43**	.14
15. Fairness Cognitive Ability	.10	.12	.23**	.46**	.70**	.17*
16. Information Known Conscientiousness	.20**	.18*	.09	.31**	.19*	.14
17. Chance to Perform Conscientiousness	-.01	.07	.18*	.60**	.54**	.00
18. Consistency Conscientiousness	.07	.04	.04	.01	.15	.03
19. Treatment Conscientiousness	.21**	.23**	.26**	.10	.05	.21**
20. Propriety of Questions Conscientiousness	.27**	.29**	.36**	.21**	.23**	.20**
21. Information Known Cognitive Ability	.12	.16*	.05	.26**	.21**	.13
22. Chance to Perform Cognitive Ability	-.02	.00	.22**	.42**	.45**	.02
23. Consistency Cognitive Ability	.07	.08	.09	.10	.17*	.13
24. Treatment Cognitive Ability	.26**	.18*	.23**	.10	.09	.20**
25. Propriety of Questions Cognitive Ability	.26**	.26**	.28**	.16*	.27**	.23**

Table 21 continued

	13.	14.	15.	16.	17.	18.
14. Job Relatedness Cognitive Ability	.24**					
15. Fairness Cognitive Ability	.29**	.62**				
16. Information Known Conscientiousness	.11	.21**	.17*			
17. Chance to Perform Conscientiousness	.21**	.54**	.44**	.23**		
18. Consistency Conscientiousness	.08	-.14	-.06	-.06	-.04	
19. Treatment Conscientiousness	.25**	.07	.11	.06	.03	.34**
20. Propriety of Questions Conscientiousness	.30**	.19*	.25**	.18*	.16*	.21**
21. Information Known Cognitive Ability	.12	.39**	.33**	.74**	.25**	-.13
22. Chance to Perform Cognitive Ability	.33**	.60**	.53**	.12	.79**	-.06
23. Consistency Cognitive Ability	.11	-.04	.03	.01	.01	.75**
24. Treatment Cognitive Ability	.28**	.09	.15*	.05	.00	.37**
25. Propriety of Questions Cognitive Ability	.33**	.31**	.37**	.13	.22**	.12

	19.	20.	21.	22.	23.	24.
20. Propriety of Questions Conscientiousness	.47**					
21. Information Known Cognitive Ability	.07	.16*				
22. Chance to Perform Cognitive Ability	-.01	.12	.22**			
23. Consistency Cognitive Ability	.29**	.18*	-.04	.04		
24. Treatment Cognitive Ability	.79**	.37**	.07	.03	.39**	
25. Propriety of Questions Cognitive Ability	.27**	.71**	.16*	.26**	.15	.32**

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

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 22

Participant Demographics for Study 2

Gender			Race						
Female	Male	Not Reported	Caucasian	African American	Hispanic	Asian	Native American	Other	Not Reported
105	69	5	106	24	5	21	1	6	8

Table 23

Means, Standard Deviations, and Reliability Coefficients for Study 2

	N	*M*	SD	*α*
TAS Initial†	174	6.54	.65	.92
TAS Conscientiousness†	174	6.55	.64	.92
TAS Cognitive Ability†	174	6.43	.76	.94
Job Relatedness Conscientiousness	175	3.28	.69	.88
Job Relatedness Cognitive Ability	175	3.15	.81	.93
Fairness Conscientiousness	176	3.23	.95	.68
Fairness Cognitive Ability	176	3.40	.92	.80
Perceived Performance Conscientiousness	176	4.39	.68	.92
Perceived Performance Cognitive Ability	176	4.14	.78	.93
NEO-FFI Conscientiousness Scale Contextualized	176	40.57	5.49	.83
HBRI	176	10.53	3.04	.69
Expected Performance – Self	175	4.38	.49	.85
Expected Performance – Other	87	4.25	.66	.93
Contextual Performance	89	4.43	.62	.94
GPA	172	3.12	.47	NA
Conscientiousness Context				
At work	176	4.43	.88	NA
At school	176	4.12	1.19	NA
At home	176	3.76	1.26	NA
With friends	176	3.43	1.43	NA
Everyday	176	4.23	.99	NA

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 24

Researcher Differences for All Variables for Study 2

	Researcher 1		Researcher 2			
Scale	*M*	*SD*	*M*	*SD*	*df*	*t*
TAS Initial†	6.55	.64	6.23	1.18	172	.98
TAS Conscientiousness†	6.56	.63	6.13	1.10	172	1.14
TAS Cognitive Ability†	6.44	.75	5.67	.90	172	1.77
Job Relatedness Conscientiousness	3.28	.69	3.04	.71	173	.71
Job Relatedness Cognitive Ability	3.15	.81	3.18	.81	173	-.08
Fairness Conscientiousness	3.22	.95	3.67	.90	174	-.93
Fairness Cognitive Ability	3.41	.93	3.08	.57	174	.70
Perceived Performance Conscientiousness	4.40	.68	4.13	1.01	174	.78
Perceived Performance Cognitive Ability	4.14	.78	3.94	.83	174	.52
NEO-FFI Conscientiousness Scale Contextualized	40.60	5.47	39.25	7.23	174	.49
HBRI	10.56	3.06	9.25	1.50	174	.86
Expected Performance – Self	4.39	.49	3.83	.45	173	2.30*
Expected Performance – Other	4.27	.66	3.70	.76	85	1.49
Contextual Performance	4.44	.62	4.07	.48	87	1.02
GPA	3.12	.48	3.16	.35	170	-.19

* $p < .05$

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 25

University Differences for All Variables for Study 2

	UMSL		UNCC			
Scale	*M*	*SD*	*M*	*SD*	*df*	*t*
TAS Initial†	6.57	.64	6.27	.69	172	1.68
TAS Conscientiousness†	6.57	.63	6.31	.74	172	1.57
TAS Cognitive Ability†	6.46	.74	6.14	.89	172	1.61
Job Relatedness Conscientiousness	3.26	.67	3.42	.80	173	-.90
Job Relatedness Cognitive Ability	3.14	.83	3.22	.58	173	-.41
Fairness Conscientiousness	3.24	.95	3.08	1.01	174	.64
Fairness Cognitive Ability	3.39	.93	3.48	.88	174	-.35
Perceived Performance Conscientiousness	4.42	.67	4.08	.75	174	1.93
Perceived Performance Cognitive Ability	4.16	.77	3.91	.83	174	1.25
NEO-FFI Conscientiousness Scale Contextualized	40.55	5.51	40.75	5.45	174	-.14
HBRI	10.57	2.98	10.19	3.69	174	.48
Expected Performance – Self	4.40	.50	4.16	.42	173	1.85
Expected Performance – Other	4.24	.67	4.50	.51	85	-.94
Contextual Performance	4.40	.63	4.75	.37	87	-1.34
GPA	3.12	.47	3.12	.51	170	-.03

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Table 26

Means and Standard Deviations for Relationship to Participant and Expected Performance and Contextual Performance for Study 2

	M	*SD*
Expected Performance – Other		
Supervisor	4.30	.62
Instructor	4.30	NA
Co-Worker	4.24	.90
Classmate	4.16	.73
Other	3.93	.96
Contextual Performance		
Supervisor	4.48	.56
Instructor	4.90	NA
Co-Worker	4.52	.55
Classmate	4.24	.77
Other	4.13	.96

Table 27

ANOVA Results for Expected Performance and Contextual Performance as a Function of Relationship to the Participant for Study 2

	df	*SS*	*MS*	*F*	η^2
Expected Performance – Other					
Between Groups	4	.73	.18	.40	.02
Within Groups	82	36.95	.451		
Contextual Performance					
Between Groups	4	1.31	.33	.85	.04
Within Groups	84	32.23	.38		

Table 28

Correlations between Assessment Scores and Criteria Measures by Assessment Order

	Conscientiousness First†			Cognitive Ability First†			Study 2††		
	r	95% CI		*r*	95% CI		*r*	95% CI	
Conscientiousness Assessment									
GPA	.16	.00	.30	.15	.00	.29	-.01	-.16	.14
Contextual Performance	.09	-.13	.30	.30	.10	.48	.34	.14	.51
Cognitive Ability Assessment									
GPA	.25	.10	.38	.34	.20	.47	.18	.03	.32
Contextual Performance	.24	.03	.43	.07	-.14	.27	.11	-.10	.31

† Conscientiousness assessment was the NEO-FFI Conscientiousness Scale; Cognitive Ability assessment was the Watson-Glaser Critical Thinking Appraisal – Form S †† Conscientiousness Assessment was the Contextualized NEO-FFI Conscientiousness Scale; Cognitive Ability assessment was the HBRI

Table 29

Conscientiousness Context Items Means and Standard Deviations for Study 1 and Study 2 Participants

	Study 1 Conscientiousness First		Study 1 Cognitive Ability First		Study 2	
Conscientiousness Context	*M*	*SD*	*M*	*SD*	*M*	*SD*
At work	4.30	1.09	3.89	1.35	4.43	.88
At school	4.39	.96	4.21	1.12	4.12	1.19
At home	4.25	1.01	3.92	1.27	3.76	1.26
With friends	3.84	1.26	3.65	1.29	3.43	1.43
Everyday	4.42	.87	4.11	1.11	4.23	.99

Table 30

ANOVA Results for Conscientiousness Context Items as a Function of Group for Study 1 and Study 2

	df	*SS*	*MS*	*F*	η^2
At work					
Between Groups	2	27.89	13.94	11.01	.04
Within Groups	520	658.40	1.27		
At school					
Between Groups	2	6.55	3.28	2.73	.01
Within Groups	520	624.57	1.20		
At home					
Between Groups	2	22.00	11.00	7.82	.03
Within Groups	520	731.62	1.41		
With friends					
Between Groups	2	14.34	7.17	4.05	.02
Within Groups	520	920.36	1.77		
Everyday					
Between Groups	2	8.26	4.13	4.17	.02
Within Groups	519	513.37	1.00		

Table 31

Means and Standard Deviations for Study 2 Based on Race

	M	*SD*
GPA		
White	3.14	.41
Black	2.84	.60
Asian	3.32	.46
Other	3.09	.45
HBRI		
White	11.58	2.50
Black	7.96	2.50
Asian	9.45	3.72
Other	9.58	2.07
NEO-FFI: Conscientiousness Contextualized		
White	40.99	5.19
Black	42.13	3.80
Asian	37.72	6.54
Other	41.75	5.59

Table 32

Study 2 One-way ANOVA Results for Variables as a Function of Race

	df	*SS*	*MS*	*F*	η^2
GPA					
Between Groups	3	3.02	1.01	4.89**	.08
Within Groups	164	33.68	.21		
HBRI					
Between Groups	3	321.05	107.02	14.52***	.21
Within Groups	167	1230.78	7.37		
NEO-FFI: Conscientiousness Contextualized					
Between Groups	3	327.29	109.10	3.88*	.07
Within Groups	167	4701.66	28.15		

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

Table 33

Correlations between Assessment Scores and Criteria Measures by Assessment Order for White Participants

	Conscientiousness First†			Cognitive Ability First†			Study 2††		
	r	95% CI		*r*	95% CI		*r*	95% CI	
Conscientiousness Assessment									
GPA	.26**	.07	.44	.23*	.03	.41	.11	-.09	.29
Contextual Performance	.18	-.10	.43	.52***	.30	.69	.15	-.13	.40
Cognitive Ability Assessment									
GPA	.22*	.03	.40	.30**	.11	.47	.05	-.14	.24
Contextual Performance	.19	-.08	.44	-.01	-.27	.25	-.07	-.33	.21

† Conscientiousness assessment was the NEO-FFI Conscientiousness Scale; Cognitive Ability assessment was the Watson-Glaser Critical Thinking Appraisal – Form S †† Conscientiousness Assessment was the Contextualized NEO-FFI Conscientiousness Scale; Cognitive Ability assessment was the HBRI

Table 34

Means, Standard Deviations, and Coefficient Alphas for Exploratory Procedural Justice Subscales for Study 2

Scale	*N*	*M*	*SD*	*α*
Information Known Conscientiousness	176	2.90	.98	.74
Chance to Perform Conscientiousness	174	2.19	1.10	.93
Consistency Conscientiousness	176	4.40	.76	.87
Treatment Conscientiousness	175	4.61	.50	.85
Propriety of Questions Conscientiousness	174	4.40	.64	.75
Information Known Cognitive Ability	176	2.72	1.06	.84
Chance to Perform Cognitive Ability	176	2.46	1.08	.93
Consistency Cognitive Ability	176	4.47	.70	.94
Treatment Cognitive Ability	176	4.62	.52	.87
Propriety of Questions Cognitive Ability	176	4.49	.60	.67

Table 35

Correlations between Exploratory Procedural Justice Scales and Study Variables – Conscientiousness First

	1.	2.	3.	4.	5.
1. GPA					
2. Contextual Performance	.18				
3. Expected Performance - Self	-.11	.21			
4. Expected Performance - Other	.14	.81**	.26*		
5. HBRI	.18*	.11	.16*	.06	
6. NEO-FFI Conscientiousness Scale Contextualized	-.01	.34**	.36**	.22*	.00
7. TAS Initial†	.01	.50**	.49**	.43**	.05
8. TAS Conscientiousness†	-.03	.47**	.46**	.39**	.06
9. Perceived Performance Conscientiousness	-.05	.35**	.40**	.22*	.02
10. Job Relatedness Conscientiousness	.06	.23*	.26**	.16	-.07
11. Fairness Conscientiousness	-.14	.11	.19*	-.01	-.15*
12. TAS Cognitive Ability†	.04	.44**	.52**	.35**	.20**
13. Perceived Performance Cognitive Ability	.06	.12	.38**	.08	.18*
14. Job Relatedness Cognitive Ability	.06	.18	.32**	.17	.05
15. Fairness Cognitive Ability	.00	.13	.30**	.02	.05
16. Information Known Conscientiousness	-.09	.14	-.01	.11	-.14
17. Chance to Perform Conscientiousness	-.14	-.03	-.06	.04	-.36**
18. Consistency Conscientiousness	.01	.20	.24**	.13	.14
19. Treatment Conscientiousness	.18*	.14	.16*	.05	.21**
20. Propriety of Questions Conscientiousness	.02	.02	.19*	-.06	.13
21. Information Known Cognitive Ability	-.05	.14	.05	.19	-.13
22. Chance to Perform Cognitive Ability	-.01	.10	.15*	.14	-.07
23. Consistency Cognitive Ability	.19*	.29**	.20**	.21	.17*
24. Treatment Cognitive Ability	.17*	.22*	.25**	.13	.24**
25. Propriety of Questions Cognitive Ability	.04	.05	.21**	-.01	.07

Table 35 Continued

	6.	7.	8.	9.	10.
7. TAS Initial†	.50**				
8. TAS Conscientiousness†	.48**	.88**			
9. Perceived Performance Conscientiousness	.59**	.52**	.54**		
10. Job Relatedness Conscientiousness	.44**	.30**	.31**	.48**	
11. Fairness Conscientiousness	.30**	.22**	.20**	.39**	.56**
12. TAS Cognitive Ability†	.42**	.71**	.73**	.53**	.31**
13. Perceived Performance Cognitive Ability	.25**	.32**	.29**	.40**	.24**
14. Job Relatedness Cognitive Ability	.29**	.24**	.21**	.31**	.64**
15. Fairness Cognitive Ability	.21**	.19*	.17*	.32**	.42**
16. Information Known Conscientiousness	.27**	.00	.01	.15*	.20**
17. Chance to Perform Conscientiousness	.06	-.03	.00	.09	.40**
18. Consistency Conscientiousness	.31**	.26**	.22**	.26**	.20**
19. Treatment Conscientiousness	.35**	.30**	.30**	.25**	.12
20. Propriety of Questions Conscientiousness	.44**	.29**	.31**	.36**	.38**
21. Information Known Cognitive Ability	.12	.02	.01	.09	.14
22. Chance to Perform Cognitive Ability	.13	.17*	.15*	.16*	.35**
23. Consistency Cognitive Ability	.30**	.28**	.28**	.26**	.27**
24. Treatment Cognitive Ability	.34**	.35**	.34**	.29**	.15
25. Propriety of Questions Cognitive Ability	.30**	.30**	.30**	.29**	.36**

Table 35 Continued

	11.	12.	13.	14.	15.
12. TAS Cognitive Ability†	.19*				
13. Perceived Performance Cognitive Ability	.17*	.53**			
14. Job Relatedness Cognitive Ability	.42**	.30**	.36**		
15. Fairness Cognitive Ability	.60**	.23**	.28**	.64**	
16. Information Known Conscientiousness	.15*	.01	.14	.17*	.04
17. Chance to Perform Conscientiousness	.41**	-.03	.05	.35**	.28**
18. Consistency Conscientiousness	.10	.29**	.19*	.15	.06
19. Treatment Conscientiousness	.05	.25**	.14	.12	.03
20. Propriety of Questions Conscientiousness	.28**	.25**	.10	.29**	.21**
21. Information Known Cognitive Ability	.12	.02	.15	.18*	.18*
22. Chance to Perform Cognitive Ability	.26**	.20**	.31**	.67**	.47**
23. Consistency Cognitive Ability	.15*	.34**	.24**	.22**	.13
24. Treatment Cognitive Ability	-.02	.35**	.23**	.19*	.05
25. Propriety of Questions Cognitive Ability	.20**	.33**	.25**	.42**	.34**

	16.	17.	18.	19.
17. Chance to Perform Conscientiousness	.15*			
18. Consistency Conscientiousness	.08	-.10		
19. Treatment Conscientiousness	.02	-.14	.62**	
20. Propriety of Questions Conscientiousness	.08	.05	.38**	.56**
21. Information Known Cognitive Ability	.69**	.16*	.01	-.05
22. Chance to Perform Cognitive Ability	.10	.47**	-.05	-.07
23. Consistency Cognitive Ability	.07	-.05	.78**	.67**
24. Treatment Cognitive Ability	-.02	-.14	.61**	.78**
25. Propriety of Questions Cognitive Ability	.02	.07	.37**	.47**

	20.	21.	22.	23.	24.
21. Information Known Cognitive Ability	.00				
22. Chance to Perform Cognitive Ability	.05	.13			
23. Consistency Cognitive Ability	.38**	.05	.10		
24. Treatment Cognitive Ability	.47**	.01	.06	.74**	
25. Propriety of Questions Cognitive Ability	.67**	.02	.24**	.45**	.57**

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

†All reaction measures use a five-point scale with the exception of the TAS which uses a seven-point scale.

Appendix A

Job Relatedness Scale

Directions: Please answer each of the following statements on your answer sheet using the following scale:

Response anchors: 1 = strongly disagree, 3 = neither agree nor disagree, 5 = strongly agree

Items:

1. I did not understand what the examination had to do with the job.
2. I could not see any relationship between the examination and what is required on the job.
3. It would be obvious to anyone that the examination is related to the job.
4. The actual content of the examination was clearly related to the job.
5. There was no real connection between the examination I went through and the job.
6. Failing to pass the examination clearly indicates that you can't do the job.
7. I am confident that the examination can predict how well an applicant will perform on the job.
8. My performance on the examination was a good indicator of my ability to do the job.
9. Applicants who perform well on this type of examination are more likely to perform well on the job than applicants who perform poorly.
10. The employer can tell a lot about the applicant's ability to do the job from the results of the examination.
11. Doing well on this test means a person can do the entry level manager job well.
12. A person who scored well on this test will be a good entry level manager.
13. It would be clear to anyone that this test is related to the entry level manager job.
14. The content of the test was clearly related to the entry level manager job.

Appendix B

Fairness Scale

Directions: Please answer each of the following statements on your answer sheet using the following scale:

Response anchors: 1 = strongly disagree, 3 = neither agree nor disagree, 5 = strongly agree

Items:

1. I feel that using the test to select applicants for the job is fair.
2. The use of the test would allow screening every applicant fairly and giving them the same opportunity to compete for the job.
3. Using the test would cut down on favoritism that can sometimes be a problem when applicants are selected for jobs.

Appendix C

Test Attitude Survey (TAS) Motivation Scale

Directions: Please answer each of the following statements on your answer sheet using the following scale:

Response anchors: 1 = disagree, 4 = neither agree nor disagree, 7 = agree

Items:

1. Doing well on this test is important to me.
2. I wanted to do well on this test.
3. I tried my best on this test
4. I tried to do the very best I could on this test.
5. While taking this test, I concentrated and tried to do well.
6. I want to be among the top scorers on this test.
7. I pushed myself to work hard on this test.
8. I was extremely motivated to do well on this test.
9. I just didn't care how I did on this test.
10. I didn't put much effort into this test.

Appendix D

Perceived Performance Scale

Directions: Please answer each of the following statements on your answer sheet using the following scale:

Response anchors: 1 = strongly disagree, 3 = neither agree nor disagree, 5 = strongly agree

Items:

1. I expect that I did well enough on the test I took today to stay in the selection process.
2. I believe I did well on the test I took today.
3. I believe that I will get a good score on the test I took today.
4. I believe that I passed the test that I took today.

Appendix E

Conscientiousness Context Scale

Directions: Please answer each of the following statements on your answer sheet using the following scale:

Response anchors: 1 = strongly disagree, 3 = neither agree nor disagree, 5 = strongly agree

Items:

1. I was thinking about how I behave at work when answering these questions.
2. I was thinking about how I behave at school when answering these questions.
3. I was thinking about how I behave at home when answering these questions.
4. I was thinking about how I behave with friends when answering these questions.
5. I was thinking about how I behave everyday when answering these questions.

Appendix F

Procedural Justice Scales

Directions: Please answer each of the following statements on your answer sheet using the following scale:

Response anchors: 1 = strongly disagree, 3 = neither agree nor disagree, 5 = strongly agree

Information known:

1. I understood in advance what the testing process would be like.
2. I knew what to expect on the test.
3. I had ample information about what the format of the test would be.

Chance to perform:

4. I could really show my skills and abilities through this test.
5. This test allowed me to show what my job skills are.
6. This test gives applicants the opportunity to show what they can really do.
7. I was able to show what I can do on this test.

Consistency:

8. The test was administered to all applicants the same way.
9. There were no differences in the way the test was administered to different applicants.
10. Test administrators made no distinction in how they treated applicants.

Treatment:

11. I was treated politely during the testing process.
12. The test administrators were considerate during the test.
13. The test administrators treated applicants with respect during today's testing process.
14. The testing staff put me at ease when I took the test.
15. I was satisfied with my treatment at the test site.

Propriety of Questions:

16. The content of the test did not appear to be prejudiced.
17. The test itself did not seem too personal or private.
18. The content of the test seemed appropriate.

Appendix G

Contextual Measure of Performance

The person that has asked you to complete this questionnaire has volunteered to participate in a research study at the University of Missouri – St. Louis. This research is being conducted as part of a doctoral dissertation.

The participants in this study have been asked to give the following questionnaire to a supervisor, though they may have obtained permission to give this questionnaire to someone else. Please answer the following questions based on your observations of the person that asked you to complete this form.

You will notice that this form does not contain the participant's name. The identifying number on this form is not tied to the participant's name either, so please answer the questions below as honestly as possible. Additionally, you have been provided a postage paid envelope addressed to the primary researcher for this study, assuring that the person you are rating will not see your responses.

Should you have any questions about this questionnaire or the research being conducted, please contact the primary researcher.

How long have you known the person you are rating? _____

What is your relationship to the person you are rating? _____ Supervisor

_____ Instructor

_____ Co-worker

_____ Classmate

_____ Other (please indicate _____)

Appendix G - Continued

	1 Does NOT meet expectations	2	3 Meets Expectations	4	5 Exceeds Expectations
Demonstrates responsibility	1	2	3	4	5
Takes personal pride in quality of work	1	2	3	4	5
Uses work/class time efficiently	1	2	3	4	5
Accepts changes in work/class schedule	1	2	3	4	5
Reports to work/class on time	1	2	3	4	5
Produces at a high level of output	1	2	3	4	5
Performs job/class assignments correctly as scheduled	1	2	3	4	5
Follows through on commitments	1	2	3	4	5
Quickly adjusts to changes in policies	1	2	3	4	5
Completes work with little supervision	1	2	3	4	5

Appendix H

Expected Performance Scale

Directions: Please answer each of the following items based on how you believe that you would perform in a work environment.

Response anchors: 1 = would not meet expectations, 3 = would meet expectations, 5 = would exceed expectations

Items:

1. Communicate with others effectively.
2. Use logic and reasoning when trying to solve problems considering relative costs and benefits.
3. Manage your/their time well along with the time of others.
4. Ability to motivate and direct the people around you/them.
5. Ability to adjust your/their actions in relation to others' actions.
6. Ability to generate multiple ideas about a topic.
7. Use of mathematic principles to solve problems.
8. Encourage and build trust, respect, and cooperation among others.
9. Ability to use computers and standard computer applications (MS Word, Excel, etc).
10. Develop specific goals and ways to accomplish them.

Appendix I

Measure of Expected Performance

Please answer each of the following items based on how you believe that you/the participant would perform in a work environment.

	1 Would NOT Meet Expectations	2	3 Would Meet Expectations	4	5 Would Exceed Expectations
Communicates with others effectively.	1	2	3	4	5
Uses logic and reasoning when trying to solve problems	1	2	3	4	5
Manages time well along with the time of others.	1	2	3	4	5
Ability to motivate and direct people.	1	2	3	4	5
Ability to adjust actions in relation to others' actions.	1	2	3	4	5
Ability to generate multiple ideas about a topic.	1	2	3	4	5
Use of mathematic principles to solve problems.	1	2	3	4	5
Encourages and builds trust, respect and cooperation among others.	1	2	3	4	5
Ability to use computers and standard computer applications (MS Word, Excel, etc).	1	2	3	4	5
Develops specific goals and ways to accomplish them.	1	2	3	4	5

Appendix J

Job Description

Below is a job description for the position of Office Manager. Please respond to all items in your booklet today as you would if you were applying for this job.

Office Manager

Responsibilities

- Supervise all aspects of the office to ensure that quality standards and deadlines are met.
- Ensure that proper procedures are followed and correct any errors or problems found.
- Provide guidance to employees for handling difficult and complex problems.
- Implement corporate and departmental policies and procedures.
- Train employees in their job duties and company policies.
- Evaluate employees' job performance.
- Discuss job performance problems with employees to identify causes and issues and to work on resolving the problems.

Job Requirements

- Must be able to communicate clearly and effectively with all levels of the organization.
- Must be proficient with standard computer applications (MS Word, Excel, etc).
- Must be able to manage one's own time and the time of others.
- Must be able to gather and analyze information to make decisions and solve problems.
- Must have an understanding of business and management principles.
- Must have an Associate's Degree and two years experience, or a Bachelor's Degree.

Adapted from O*NET.

Appendix K

Participant Instructions

You will be taking two different selection assessments today. These assessments are similar to ones that you may encounter if you were applying to the position of Office Manager that you just read about. Please answer your assessments as if you were applying for this job, keeping in mind that the top 25 performers will receive $20.

In addition to the two assessments, you will be asked to provide your opinion about the assessments and some information about yourself. These portions of your test booklet will not be used to determine the top 25 performers, so please respond to them as honestly as you can.

Once you begin answering your test booklet, you will need complete all of the items in order. Please do not skip around, or go back to change your responses to either the test items or your opinions about the items. None of the sections in your test booklet will be timed, but you should be able to complete all portions in two hours or less. I do ask that you work through the booklet carefully as you will not be allowed to go back and change your responses.

When you have completed all portions of your booklet, you are free to leave. Keep in mind, that you need to perform to the best of your ability as only the top 25 performers will receive $20. To find out if you are one of the top performers, you will need to periodically check the website listed on your take home sheet for your unique identification number. I will not be providing feedback about your performance outside of posting the identifying numbers of the top 25 performers.

Thank you. You may begin.

Appendix L

Demographic Information Sheet

Organizations often collect information about their applicants in order to make sure the tests they use are fair to all applicants. The information below will be used to determine if the tests you completed today are fair for everyone. Your name should not be placed anywhere on this sheet. Please answer each question to the best of your ability.

Age _____

Race _____

Gender _____

Year in college (please circle): Freshman Sophomore Junior Senior Post Baccalaureate

College Grade Point Average (GPA) _____

Are you a transfer student? Yes No

Have you held a full time job (more than 35 hours a week)? _____

Have you held a part time job? _____

Have you taken tests similar to the Conscientiousness measure given today when applying for a job? Yes No

Have you taken tests similar to the cognitive ability measure given today when applying for a job? Yes No

Appendix M

Participant Handout

Thank you for participating in my research study. Please remember to give the 2 page questionnaire and self-addressed envelope to your supervisor. If you do not currently have a work supervisor, or you are unable to find a suitable supervisor to complete the form, please email me at the address below.

Please keep this sheet so that you will be able to collect your $20 reward should you be one of the top 25 performers in my study. The web address listed below will be used to identify the top performers, so please check it regularly. If your unique identifier code (found in the upper right hand corner of this page) is listed, please bring this form with you in order to claim your reward.

Website for participants: http://dgrambowresearchpage.blogspot.com

If you should have any questions about the study, please email me at **dana.grambow@umsl.edu**. Please allow 24 hours for me to respond to all emails.