

THE IMPACT OF ITEM CHARACTERISTICS ON CONTEXTUALIZED  
PERSONALITY ASSESSMENT

BY

STEPHANIE NICHOLE SEILER

DISSERTATION

Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Psychology  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Fritz Drasgow, Chair  
Associate Professor Daniel Newman  
Professor Brent Roberts  
Professor James Rounds  
Professor Hua-Hua Chang

## ABSTRACT

This dissertation examines the extent to which personality items that were developed to describe personality in general are adequate descriptors of personality within a specific context. This research provides for a deeper understanding of contextualized personality assessment by examining the effects of item context specificity and perceived context relevance on the measurement properties of Big Five Conscientiousness facets within the contexts of work and school. Results indicate that an item's discrimination is related to its perceived context relevance and its level of context specificity. Implications for the improved measurement of personality are discussed.

*To my family, friends, and mentors*

## TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION .....	1
CHAPTER 2: LITERATURE REVIEW .....	4
CHAPTER 3: HYPOTHESES.....	19
CHAPTER 4: METHOD .....	30
CHAPTER 5: RESULTS .....	39
CHAPTER 6: DISCUSSION.....	51
TABLES .....	58
FIGURES .....	103
REFERENCES .....	120
APPENDIX A.....	132
APPENDIX B .....	134
APPENDIX C .....	135
APPENDIX D.....	136

## CHAPTER 1

### INTRODUCTION

There is general agreement that many aspects of behavior in the workplace can be understood in terms of an individual's personality. There is also a high degree of consensus that the Big Five traits commonly labeled Conscientiousness, Extraversion, Agreeableness, Emotional Stability, and Openness to Experience provide a reasonable framework for conducting research in this area. The use of this framework has helped to advance the study of the predictive validity of personality inventories, as well as theoretical research on personality at work.

Prior to a major turn of events in 1991, the notion that personality was not strongly related to job performance had gone "unchallenged for 25 years" (Barrick, Mount, & Judge, 2001; p. 9). As noted frequently in the literature (e.g., Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt, 2007; Roberts, 2006) this longstanding acceptance of the status quo stemmed from the work of a handful of convincing scholars who pointed to the inconsistent and frequently low correlations reported in the literature between personality scales and job performance (e.g., Ghiselli & Barthol, 1953; Guion & Gottier, 1965; Schmitt, Gooding, Noe, & Kirsch, 1984), and was strengthened by behaviorist and social psychological claims that the situation is more powerful than personality in directing behavior (e.g., Mischel, 1968; 1973). In response to these criticisms, several scholars have demonstrated that, when organized in terms of the Big Five, personality scale scores often have meaningful relationships with indices of job performance (e.g., Barrick & Mount, 1991; Hogan & Holland, 2003; Tett, Jackson, & Rothstein, 1991). Barrick et al. (2001) quantitatively summarized fifteen Big Five/job

performance meta-analyses and found overall support for the validity of each trait for some dimensions of performance.<sup>1</sup> They then highlighted the need to redirect research resources to more specific questions regarding these relationships and to further examine lower-level facets of the Big Five dimensions and their relationships to targeted performance dimensions; several research articles in this vein have followed (e.g., Dudley et al., 2006; Griffin & Hesketh, 2004).

The use of the Big Five model for assessment of personality at work is to be expected, given the body of evidence that supports its validity. Although the Big Five model clearly provides a useful framework, several lines of research suggest avenues for further inquiry into how personality assessment functions within and across different test-taking contexts. The application of a Big Five assessment to the work context relies on the assumption that personality structure transfers *as is* into different contexts and that personality only varies in the degree to which the trait is displayed with regularity in a particular context; however, this assumption remains untested in many respects. For example, Gill and Hodgkinson (2007) and Heggstad and Gordon (2008) note that many of the commercial Big Five personality inventories that are commonly used for workplace assessment were developed to measure personality in clinical or non-work contexts; the content of these inventories might therefore be deficient for describing personality at work.

As characterized by Kuncel and Kuncel (1995), it is “farfetched” (p. 189) to assume that respondents view item content in the same way that researchers view item

---

<sup>1</sup> The status quo did little in the way of stopping practitioners from continuing to use personality measures to select employees; thus, researchers had access to adequate data on which to conduct these meta-analyses.

content. The authors use as an example the item, “I often go to parties”; they note that what researchers *want* is for respondents to answer the item in terms of their underlying level of shyness, not some other variable that might predict whether one goes to parties (for example, peer pressure). When responses to an item are further confounded by work contextualization, the item will likely elicit an even greater number of non-trait responses. For example, some respondents might assume that the item is intended to measure whether they prioritize work over leisure time.

In sum, the need to systematically study item characteristics that impact responses to a contextualized personality assessment is often stated, but has not yet been addressed empirically. In the literature review that follows, I review research that provides direction for addressing this question.

## CHAPTER 2

### LITERATURE REVIEW

#### Cross-Cultural Research Framework

The study of contextualized personality assessment can be likened to the more established body of research that investigates the structure and correlates of personality traits across cultures (Schmit, Kihm, & Robie, 2000). In other words, a “context” is similar to a “culture” in that both are part of a larger system wherein comparability and translatability are desired features. In the two sections that follow, I provide a foundation for conceptualizing contextualized personality assessment by summarizing evidence from both *emic* and *etic* approaches to cross-cultural personality research.

#### *Emic Approaches to Examining Personality Trait Structure*

*Emic* research approaches describe a phenomenon as viewed from *within* a particular culture. In the domain of personality research, one such emic approach derives from the lexical hypothesis (Allport & Odbert, 1936). The lexical hypothesis holds that, over time, a meaningful descriptor of personality becomes encoded as a word within a particular language. Put in slightly different terms, a word that is created to describe a person must have arisen from an agreement on the need to describe a unique and discernable characteristic of human nature. If true, then the process of analyzing the relationships amongst these words should lead to the discovery of an overarching personality framework. Evidence suggests that the Big Five are a reasonable representation of these relationships within a number of languages and cultures (Rolland, 2002); alternative lexical solutions also support this structure to some extent (e.g.,



Ashton, Lee, Perugini, Szarota, deVries, DiBlas, Boies & De Raad, 2004; Lee & Ashton, 2008; Saucier & Goldberg, 1996; Waller & Zavala, 1993).

A second emic approach that supports the Big Five structure is the examination of inventories that were not derived from the lexical hypothesis. In general, when these inventories are factor analyzed, all or some of the Big Five factors generally emerge as an adequate structural representation of the items (Rolland, 2002; Gill & Hodgkinson, 2007). For example, in their factor analytic investigation of the 16PF, Birenbaum and Montag (1986) opted for a five factor solution; these factors appear to largely replicate the Big Five.<sup>2</sup> McCrae and Costa have reported similar findings for a number of instruments, including the Eysenck Personality Inventory (McCrae & Costa, 1985) and the Myers-Briggs Type Indicator (McCrae & Costa, 1989). Although slightly different factor structures or additional factors are sometimes extracted within individual inventories, the Big Five traits are viewed as the most replicable; some of the conflicting findings have likely resulted from measurement error due to small samples and/or failure to adequately sample the domain of Big Five trait descriptors.

A similar emic approach can also be used to examine the structure of inventories that were developed to assess work-specific personality with no reference to the Big Five. This research suggests that such inventories—for example, the Occupational Personality Questionnaire (OPQ32)—are interpretable, at least in part, in terms of the Big Five (Bartram, Brown, Fleck, Inceoglu, & Ward, 2006).

---

<sup>2</sup> The results of their study are particularly convincing because the study was not intended to be an investigation of the validity of the Big Five structure. In fact, the Big Five received no mention in their article; as suggested by Digman (1990), the authors might have been largely unaware of the emergence of five-factor personality research.

### *Etic Approach to Examining Personality Trait Structure*

An *etic* approach to studying a phenomenon involves transporting results derived from one culture into a different culture, and subsequently comparing those results. Cross-cultural etic personality research has demonstrated that, when a Big Five measure is translated into the language of another culture, an equivalent five-factor solution is commonly retained within that culture (Rolland, 2002).

The use of the Big Five structure to assess contextualized personality can be viewed as an etic approach. That is, when asked to respond to a general Big Five inventory in terms of one's characteristics in a particular context, the inventory must be "translated" so that it is interpretable in that context. Research has generally supported the Big Five structure in work-contextualized personality assessment (e.g., Robie, Schmit, Ryan, & Zickar, 2000; Smith, Hanges, & Dickson, 2001).<sup>3</sup> For example, Smith et al. (2001) analyzed a subset of items from the Hogan Personality Inventory that were selected explicitly to measure the Big Five. Using confirmatory factor analysis, they concluded that a five-factor structure held across applicant, incumbent, and student populations. Research has also provided some support for the Big Five structure in inventories that have adapted general Big Five factors to include more work-relevant language. For example, the Big Five structure was supported for the Five Factor Model

---

<sup>3</sup> A related line of research suggests that the factor structure and predictive validity of a personality inventory might be influenced by the motivation of the respondents. Schmit and Ryan (1993) examined a Big Five measure of personality and found a more complex factor structure in an applicant group compared with a student sample, suggesting a potential sixth "ideal-employee" factor, whereas Paulhus, Bruce, and Trapnell (1995) found fewer factors in a simulated applicant condition. Any attempt to measure personality at work from applicant data will be confounded by the issue of motivated distortion in applicant populations. To the extent possible, this dissertation focuses on the study of personality outside of the context of selection decisions.

Questionnaire (FFMQ; Gill & Hodgkinson, 2007), an adjectival measure of the Big Five that was developed for use in work settings.

In sum, it is clear that the various methodologies employed to understand personality in general and work-contextualized personality converge to support the Big Five structure. However, although this structure may be adequate, little is known about how non-structural measurement properties might be affected by contextualization. In support of this notion, research on contextualized personality assessment using item response theory (IRT) has indicated that certain items function differentially across contexts. For example, Stark, Chernyshenko, Chan, Lee, and Drasgow (2001) found that a number of personality items within each scale (excluding the Impression Management scale) of the Sixteen Personality Factor Questionnaire (16PF) displayed differential item functioning (DIF) between an applicant and non-applicant sample. They also found differential test functioning (DTF) for 13 of the 15 scales examined. Upon inspection, the researchers did not perceive any item content differences between items displaying DIF and those that did not. Similarly, Zickar and Robie (1999) found DIF for a number of items across three personality scales. Due to test security reasons, they were not able to examine the content of these items.

This primary goal of this dissertation is to identify the characteristics of items that might impact the non-structural measurement properties of a contextualized personality inventory. In the next two sections, I describe research that directly supports this aim.

## Contextualized Personality Assessment

Whereas item translation in cross-cultural research is performed solely by test developers, translation in contextualized personality assessment is sometimes performed by test developers but is just as often left to the respondents. In other words, personality inventories used in workplace research and practice vary (1) in the extent to which test-takers are instructed to respond only in terms of their characteristics at work (Heggestad & Gordon, 2008; Robie et al., 2000), and (2) in the extent to which the inventory makes use of work-relevant language, as opposed to the general language of inventories that were developed for non-work purposes (Gill & Hodgkinson, 2007; Heggestad & Gordon, 2008). This poses a problem because when test-takers are not given explicit instructions to respond only in terms of work characteristics, they must make their own decisions regarding whether to only think about themselves at work. Similarly, when test-takers are asked to respond to a general Big Five personality inventory in terms of their characteristics at work, they assume the sole responsibility for translating items so that they make sense in that context. Consequently, when it is not clear how an item relates to work, it is likely that many test-takers will adopt an idiosyncratic response strategy (Lievens, De Corte, & Schollaert, 2008; Sackett & Lievens, 2007). As is discussed in the next section, such heterogeneity in response strategy can negatively affect the reliability and validity of the assessment.

### *Traditional Approach*

As labeled by Heller, Watson, Komar, Min, and Perunovic (2007), the *traditional approach* to contextualized personality assessment involves instructing respondents to consider a particular context when responding to a personality inventory. The majority of

research on the criterion validity of work contextualized personality has adopted this traditional approach and has offered some evidence that providing a work *frame-of-reference* (or, “FOR”) for personality items results in stronger correlations between personality scales and job performance (e.g., English, 2001; Pace & Brannick, 2010; Robie, Born, & Schmit, 2001; Schmit, Ryan, Stierwalt, & Powell, 1995). Accordingly, Heggstad and Gordon (2008) recommended that both researchers and practitioners use contextualized assessments. The authors also called for research to address “how much contextualization is required” (p. 321) and to investigate the degree to which contextualized measures demonstrate construct equivalence with general, non-contextualized measures.

Lievens et al. (2008) concluded that an increase in validity for a contextualized assessment results from the joint effects of increased reliability—due to a reduction in within-person inconsistency—and a stronger conceptual overlap with the criterion. More simply, when a general personality inventory is contextualized, respondents should consider each item only in terms of the relevant domain of behaviors. The authors noted that an important area of future investigation is to further examine conceptual overlap, but did not suggest any strategies to address this topic.

Although informative, the study by Lievens et al. (2008) confounds the effects of conceptual overlap with an item’s degree of *context specificity*. For an item such as “I respect others,” adding “at-work” to the item makes the item more context specific, and should also make the item appear more conceptually relevant. However, this does not necessarily hold for all items in a scale. As an extreme example, the item “I am easily

moved to tears *at work*” is fairly specific, but it addresses a characteristic that is not likely to be perceived as a work-relevant attribute.

In the next section, I argue that the influence of an item’s *perceived relevance* in contextualized personality assessment is a logical extension of the effects of *item subtlety* in general personality assessment. I then elaborate further on the concept of item specificity. Finally, I describe my expectations regarding how perceived relevance should interact with item specificity to influence the psychometric properties of an item.

### *Investigating the Joint Effects of Perceived Relevance and Context Specificity Using the Traditional Approach*

The primary goal of this dissertation is to investigate the impact of item characteristics on the measurements properties of a contextualized personality assessment. In an often cited article, Roskam (1985) conjectured that personality items that “can be solved in a variety of ways” (p. 9) will have lower slopes. In other words, an item that is phrased in abstract and general terms should elicit more variability in the decisions respondents make regarding how to answer the item. To the extent that this is true, according to Roskam, such items will be “more distantly related to the trait being measured by the set of items as a whole.” (p. 9).

Empirical evidence testing Roskam’s conjecture has provided mixed results. For example, Zumbo, Pope, Watson, and Hubley (1997) found that abstract items—operationalized as the mean of laypersons’ overall impressions of item abstractness—were actually more discriminating than concrete items. Similarly, Hubley, Woo, and Zumbo (2009) reported a near-zero correlation between laypersons’ ratings of item vagueness and discrimination. However, in support of Roskam’s conjecture, Zickar and

Ury (2002) found that subtle items—those that trained students were least able to assign to the correct personality scale—showed lower item discrimination parameter estimates. Comparable results were reported for the MMPI by Duff (1965) and by Burkhart, Gynther, and Fromuth (1980), who operationalized subtlety as the degree to which an item is indicative of a specific psychological problem.

From the above research, it appears that abstractedness operationalized as a general impression functions differently than abstractedness operationalized as one's ability to judge the meaning of the item in reference to a criterion. According to Johnson (2004), "items possessing several comprehensible meanings can possess validity because respondents can express a personality disposition *by the meaning they see in an item.*" (pg. 294; italics added). A general impression of item abstractedness might not directly impact discrimination because such items still allow for respondents to accurately portray themselves, although in a variety of ways; in fact, this type of abstractedness appears more in line with what I refer to in the next section as *context specificity*. On the other hand, if respondents are unable to infer how an item is related to others in the same scale, or do not perceive that the item is representative of a criterion of interest, this very likely indicates that they will not respond to the item in accordance with other, less abstract items from that scale. The latter conceptualization of item abstractedness seems more in line with Roskam's original reasoning.

*Perceived relevance.* Although the latter notion of item subtlety should also hold true within the study of contextualized personality assessment, I argue that it may be more productive to define a "subtle" item as one that respondents find difficult to perceive as relevant to the context of interest. Specifically, when instructed to respond

using a particular context as a frame-of-reference, items that are rated by laypersons to be irrelevant to that context should elicit a wider range of response strategies. This variability should, in turn, make the item less discriminating than an item in the same scale that is perceived as more relevant.

*Perceived Relevance/Context Specificity Interaction.* It should also hold that the impact of low perceived relevance should be strongest for the most specific items. The concept of specificity has taken on many forms in the study of personality assessment. For example, many have noted the potential for a personality scale to be artificially labeled as a distinct factor due to ‘bloated specifics’ (Cattell, 1972). That is, items that are too similar in terms of content will demonstrate large intercorrelations relative to items that are more diverse in content but are conceptually related. These larger intercorrelations will then result in an artificially narrow factor.

In the domain of contextualized personality researchers, the term *context specific* has been used to refer to (1) an assessment that has been given under a specific frame-of-reference, (e.g., Gordon & Holden, 1998) (2) the correspondingly narrow criteria to which the assessment should correlate (e.g., Heggstad & Gordon, 2008; Heller et al., 2007; Robie et al., 2000; Tett et al., 1991; Robertson & Kinder, 1993; Salgado, 1997) and/or, (3) an assessment that has been developed to measure personality within a particular context (Gill & Hodgkinson, 2007).

Whereas the above conceptualizations of specificity and context specificity refer to items in the aggregate, the term specificity can also be applied to an individual item. Several indices of item specificity have been developed for research purposes. For example, Kaiser and Craig (2004) used the term *behavioral specificity* to refer to the



degree to which an item refers to one particular behavior, as opposed to a category of behaviors. The authors use as an example the item, “Has a pleasant disposition” to typify a broad (less specific) item, whereas they offer the item, “Tries to understand what other people think before making judgments about them” as an example of a very specific item (p. 240). Similarly, Werner and Pervin (1986) described specificity as the extent to which items or entire scales “ask about people’s responses in specific situations” (p. 622).

For the purposes of this dissertation, context specificity refers to the number of contexts to which an item might apply equally well. As noted above, I expect that item specificity will interact with perceived relevance to influence item discrimination. For example, consider again one’s response to the item, “I am easily moved to tears” under a work frame-of-reference. When responding within the context of work, this relatively context-specific item will likely be perceived as very low in work relevance. In other words, the item is more relevant for other contexts than it is for work. Consequently, the item should take on different meanings for different subsets of individuals; some will respond to the item without considering the work frame-of-reference, while others might interpret the item as a colloquial statement that implies a certain level of emotional control at work. However, for an item such as, “I become very emotional when frustrated,” even if the item is judged as less relevant than other items, respondents should more readily be able to apply the item to a work setting because it is less specific to one context. In turn, this latter item should have a higher discrimination estimate than the former item.

Taken together, research on frame-of-reference effects in contextualized personality assessment and research on the impact of item characteristics in general

personality assessment suggests a new avenue for productive research. To my knowledge, this dissertation presents the first investigation into the effects of perceived context relevance and item specificity on contextualized personality assessment.

### *Beyond the Traditional Approach*

A secondary goal of this dissertation is to explore whether items that are low in perceived relevance can be reworded to produce better item discrimination within a specified context.

As already noted, an etic approach to contextualized personality assessment is not necessarily telling of the adequacy of the model for describing personality at work because it ignores those variables that might be unique, or more applicable, to the context of work. If the descriptors that are most relevant to the workplace have been systematically ignored in reaching the simple-structure Big Five result, any structural investigation of a personality inventory that conforms to this model, *even when contextualized*, will likely lead to “robustness of the irrelevant variety” (Maraun, 1997, p. 634).

Additionally, confirming the structure of an inventory does not guarantee that the measurement properties were unaffected. This approach also has the disadvantage of confounding the effects of item specificity and conceptual overlap; responses to items that are perceived as irrelevant to work will likely function differently between a general and work context. The Lievens et al. (2008) study described in the previous section circumvented the issue of using items that would be difficult to interpret under a work frame-of-reference (such as “I talk to a lot of different people at parties”) by replacing those items with items that could apply to a broader number of contexts (such as “I know

how to captivate people”). Although this strategy has the advantage of enabling respondents to respond across contexts, it lacks the rigor of a more systematic method for understanding the interaction between item wording and item responding. The new item might enable comparisons across contexts, but this strategy might result in a loss of information within contexts.

To support this latter notion, Pace and Brannick (2010) created a work-specific, facet-level Openness to Experience measure to parallel the Openness scale of the NEO-PI-R (Costa & McCrae, 1992). The researchers found higher criterion-related validity for the work-specific measure, compared to the NEO PI-R measure.

Whereas the item rewriting procedure utilized by Lievens et al. (2008) was aimed at facilitating comparisons across contexts, the item rewriting procedure utilized by Pace and Brannick (2010) demonstrated higher criterion-related validity within one context. Below, I argue that these two seemingly contradictory goals can be addressed simultaneously. That is, by identifying the item characteristics that influence item responses—such as perceived context relevance and specificity—these characteristics could be referenced in the development of parallel, contextualized items. These items would vary in content across contexts, but could demonstrate equivalence in terms of measuring the underlying trait.

Pure emic and etic approaches to understanding personality at work are useful, but they preclude a careful analysis of personality in different contexts. A useful alternative to the emic and etic extremes would be to identify those items within the inventory that are judged to be the least descriptive of behaviors and characteristics as displayed in the workplace, and to then contextualize these items using work-relevant

terminology. This approach combines the emic and etic approaches to personality research, a strategy that has frequently been adopted for cross-cultural personality research. That is, statements deemed more descriptive for work (emic) will modify statements that were developed with reference to the general Big Five domain (etic).

By adopting the language that one uses to describe behaviors and characteristics at work, the resulting inventory should be perceived as more relevant to the domain of work and will consequently facilitate accurate item responses. For example, if “I am the life of the party” was rewritten with reference to a work characteristic, it might read “I frequently talk casually with coworkers during break periods,” or, “I enjoy mingling with colleagues at work-related social gatherings.” These rewritten items should reduce the ambiguity involved in responding to the more general item but be equally—if not more—indicative of the trait Extraversion, within the context of work.

#### *Relation to Trait Theory*

Although the focus of the current research is on how item *responses* are influenced by contextualization, it is important to discuss the broader theoretical implications of contextualized personality assessment with regard to the trait theory of personality. Although many would define a personality trait as the summation of behavioral dispositions across time and situations, Roberts (2009) has argued convincingly that cross-situational consistency has never been a tenet of trait theory. It is clear that the situation plays an important role in influencing one’s display of a trait-related behavior (Wood, 2007; Wood & Roberts, 2006; Heggstad & Gordon, 2008). For example, the average level of extraversion displayed across individuals at a party is likely to be higher than the average across individuals attending a lecture. In addition, as

Roberts pointed out, it is a common occurrence to see reference to situations within many of the measures that are commonly used to assess personality traits. It is clear that context must be considered in defining and measuring personality.

Although context is important in conceptualizing personality, and although a contextualized personality measure is often more strongly correlated with behaviors in that context (e.g., Schmit et al., 1995; Wright & Mischel, 1987), Wood (2007) argued that it does not necessarily follow that the contextualized measure is more causal of those behaviors. If a personality construct is to be considered a “trait”, it is more accurate to say, for example, that someone is extraverted at work because he/she is generally extraverted, not because he/she displays extraverted behaviors at work. The latter argument is circular and therefore cannot be used to infer causation. Additionally, if a person is *more* extraverted when attending a party than he/she is at work, the “proper explanatory forces” (Wood, 2007) are (1) the trait itself, and (2) context-specific factors that promote or suppress certain behaviors.

Notwithstanding the conclusions of Wood (2007) and Wood and Roberts (2006), it is also important to acknowledge that responses to personality items are determined *by the respondents*. Although traits and situations are convincingly the true causal forces that dictate behavior, we do not expect that respondents are aware of this principle; accordingly, respondents must make relatively uninformed decisions regarding whether to consider their context-general characteristics, their context-specific characteristics, or both. Since, as asserted by Heggstad & Gordon, 2008, industrial and organizational psychologists “are generally more interested in making predictions about a person’s behavior at work than about his or her behavior in general,” (p. 321) we should be

primarily interested in how personality assessment operates within a context, thus avoiding what they view as “a mismatch between what [a general personality measure] is designed to predict (i.e., aggregated levels of general behavior) and the scope of the criterion we are interested in predicting (i.e., aggregated levels of work behavior)” (p. 321).

## CHAPTER 3

### HYPOTHESES

#### *Background for Hypotheses*

Before I present my hypotheses, it is first necessary to provide some background on the measures that were used in this research.

*Conscientiousness.* For practical reasons, this study focuses on only one of the Big Five personality traits: Conscientiousness. Conscientiousness predicts several aspects of job performance in a wide variety of occupations (Barrick et al., 2001); it thus makes practical sense to focus research efforts on this trait first.

*Facet-level approach.* Although the Big Five structure has support both cross-culturally and across contexts, less is known about the structure of the facets that underlie each trait. This lack of agreement is evident across several major Big Five inventories that include lower-level representations of Conscientiousness (Roberts, Chernyshenko, Stark, & Goldberg, 2005). For example, the NEO-PI-R measures Conscientiousness with six subscales: *dutifulness, achievement striving, competence, order, self-discipline, and deliberation*, whereas the Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008) measures four: *achievement, control, harm avoidance, and traditionalism*.

Roberts et al. (2005), noting that better coverage of a construct can be achieved by combining inventories that were developed under a variety of theoretical frameworks, examined scale intercorrelations across 7 major personality inventories. When they combined these results with the results of Roberts, Bogg, Walton, Chernyshenko, and Stark's (2004) lexical investigation of the facet structure of Conscientiousness, they found that the two approaches shared six facets in common, which they labeled

Industriousness, Order, Self-Control, Responsibility, Traditionalism, and Virtue. Three of these facets were used in this dissertation: Industriousness, Order, and Responsibility. As described in Chernyshenko (2002), Industriousness denotes hard work, ambition, confidence, resourcefulness, competence, and efficiency. Order is described as the ability to plan and organize. Finally, Responsibility is defined as being cooperative, dependable, and willing to use one's time and resources to help others.

The Industriousness and Order facets were chosen for this dissertation because, as noted by Roberts et al. (2005) these facets are present in some form in a number of major Big Five inventories, are frequently found in lexical investigations, and have good discriminant validity (i.e., high loadings on Conscientiousness relative to the other four Big Five factors). The Responsibility facet was chosen because, although Roberts et al. (2005) described the facet as an “interstitial dimension” (p. 132) that has loadings on other Big Five factors, it is frequently seen in some form in several major personality inventories; more importantly for this dissertation, compared with the remaining three facet scales in Chernyshenko (2002), this scale appeared that it would better serve the goals of the current research because it was clear that its items would elicit a wide range of perceived relevance ratings. For example, the item, “When others depend on me, I try to put in an extra effort” would presumably elicit a much higher work relevance rating than the item, “I would gladly spend some of my leisure time trying to improve my community.”

*Definitions of Scales.* The henceforth labeled *Original* scales used in this study are Chernyshenko's (2002) “classical test theory” facet-level measures of Industriousness, Order, and Responsibility facets. Each scale consists of the ten positively



worded and ten negatively worded items from a larger item pool that had the highest item-total correlations. The items are given in Tables 1-3.

Although the work context is the primary focus in this dissertation, I also included a school contextualization condition. Given that the sample used for this research consists of college students, it was possible that a school condition might be more appropriate for this sample. The inclusion of an additional context also served as a replication to test the generalizability of any results that are found. The henceforth labeled *Rewritten Scales* consist of items within each of the Original scales that were rewritten using more work-/school-relevant language. Finally, the *Work-Relevant Scales* and *School-Relevant Scales* (or generically, *Context-Relevant Scales*) are twenty-item facet scales consisting of the Rewritten Scale items and the Original Scale items that were not rewritten. In other words, each of the Context-Relevant Scales contains a combination of Original Scale and Rewritten Scale items. The Rewritten Scale items are given in Tables 4-6, and the development of the Rewritten Scales is described in the Method section below.

#### *Impact of Perceived Relevance*

As discussed in the literature review, research adopting the traditional approach to contextualized personality assessment does not address the question of whether certain items become less interpretable when answered under a particular frame-of-reference (FOR).

Hypothesis 1a states that the effect of perceived relevance on IRT discrimination estimates under frame-of-reference conditions will be similar to the effect of item subtlety (in a non-contextualized condition) found by Zickar and Ury (2002).

*H1a:* Under both Work and School FOR conditions, perceived relevance will be positively correlated with item discrimination.

*Interaction of Perceived Relevance and Item Specificity.*

Hypothesis 1b combines the logic of Hypothesis 1a with the expectation that item specificity will play a large role in the relationship between perceived relevance and discrimination. As discussed in the literature review, respondents should have less difficulty responding to an item that is perceived as less relevant to a context if the item allows room for interpretation (i.e., the item is less specific). Items that are very specific leave little room for interpretation and should in turn result in the most diverse range of response strategies.

*H1b:* Under both Work and School FOR conditions, a perceived relevance by specificity interaction will predict item discrimination values, such that item specificity will only impact the discrimination values of items that are perceived as less relevant.

The relatively small number of variables (20 items per facet; 60 items overall) precludes the use of multiple regression to test for interaction effects. As an alternative, I chose to dichotomize the two predictor variables and to examine the interaction graphically. Figure 1 displays the graphical representation of Hypothesis 1b.

Relevance ratings were dichotomized at the midpoint of the *obtained ratings* within the scale. A median split would define low relevance only in relation to the other items in the scale; therefore, an item could be coded as less relevant even though its rating was relatively high. At the opposite extreme, splitting at the midpoint of the rating scale assumes that, within each facet, raters utilized the entire range of ratings. As this

was not the case in the current research, splitting at the midpoint would have resulted in almost no items being rated as low in relevance. Splitting at the midpoint of the range of obtained ratings functions as a compromise between splitting at the median value and splitting at the midpoint of the full rating scale; it ensures that those items coded as less specific and less relevant are objectively less specific and less relevant, as opposed to in relation to the other items in the scale, but also ensures in most cases that a sufficient number of items will be assigned to each cell.

Because specificity ratings were made by experienced raters who used the entire range of ratings, and because it resulted in relatively equal numbers of items being classified as “low” and “high” specificity, specificity was dichotomized at the midpoint of the actual rating scale.

#### *Rewritten Item Hypotheses*

Extending the logic that of Hypothesis 1a, I expected that items that were rewritten to include context-relevant language would have higher discrimination estimates.

*H2a:* Within both the work and school FOR conditions, the Rewritten items will have larger discrimination values than their Original counterparts.

*H2b:* Within both the work and school FOR conditions, the increase in discrimination of the Rewritten items will be larger for the most specific items.

### *Criterion Validity Hypotheses*

*Effect of frame-of-reference.* Consistent with the results of English (2001), Pace and Brannick (2010), Robie et al. (2001), and Schmit et al. (1995), Hypotheses 3a through 4c state that the correlation between a personality scale and a criterion of interest will be strongest when the correct frame-of-reference is used. Although not the primary focus of this dissertation, these results were included to support the next set of hypotheses by demonstrating that participants understood and followed the frame-of-reference instructions. In other words, these results serve as a manipulation check based on known research results.

Two job performance measures, developed by Williams and Anderson (1991), were used in this study. The first criterion, *in-role behavior*, refers to the performance of required work behaviors, or what others have labeled *task performance*. The second criterion, *organizational citizenship behaviors – individuals*, measures the extent to which an employee engages in behaviors that primarily benefit another individual within the organization.

Meta-analytic research provides evidence that Conscientiousness is important for both task performance (Barrick & Mount, 1991; Barrick et al., 2001) and organizational citizenship behaviors (Organ, 1994; Organ & Ryan, 1995; Podsakoff, MacKenzie, Paine, & Bachrach, 2000). However, no meta-analytic evidence is available for the facets used in this dissertation. As such, I relied on both a rational approach and the reported results from Chernyshenko (2002) to formulate these hypotheses.

The criterion *in-role behavior* can logically be expected to relate to all three of the facets measured in this dissertation. Given that that Conscientiousness is a robust

predictor of task performance, it is logical to assume that these facets will all be positively correlated with this criterion. Again, these relationships should be stronger under Work FOR instructions.

*H3a:* For the work criterion *in-role behavior*, the Original Industriousness scale taken under a work FOR will correlate more strongly than the Original Industriousness scale taken under a school FOR.

*H3b:* For the work criterion *in-role behavior*, the Original Order scale taken under a work FOR will correlate more strongly than the Original Order scale taken under a school FOR.

*H3c:* For the work criterion *in-role behavior*, the Original Responsibility scale taken under a work FOR will correlate more strongly than the Original Responsibility scale taken under a school FOR.

Chernyshenko (2002) hypothesized that the Responsibility facet would have the strongest relationship with organizational citizenship behaviors. Although Industriousness, Order, and Responsibility were all correlated with this criterion, Chernyshenko concluded that the Industriousness and Order relationships were primarily due to their shared variance with Responsibility. Therefore, for the *organizational citizenship behavior criterion*, I chose to focus only on the Responsibility facet.

*H3d:* For the work criterion *organizational citizenship behavior*, the Original Responsibility scale taken under a

work FOR will correlate more strongly than the Original Responsibility scale taken under a school FOR.

One school performance criterion, college GPA, was measured in this study. Of the three facets investigated in this dissertation, Chernyshenko (2002) found that only Industriousness and Responsibility were significantly correlated with college GPA ( $r = .39$  and  $.18$ , respectively). However, the  $.10$  correlation between Order and college GPA (corresponding to  $p = .07$ ) suggests that the magnitude of this relationship might increase to a statistically significant value under a school FOR condition. As such, I included all three facets in my hypotheses.

*H4a:* For the school criterion *college GPA*, the Original Industriousness scale taken under a school FOR will correlate more strongly than the Original Industriousness scale taken under a work FOR.

*H4b:* For the school criterion *college GPA*, the Original Order scale taken under a school FOR will correlate more strongly than the Original Order scale taken under a work FOR.

*H4c:* For the school criterion *college GPA*, the Original Responsibility scale taken under a school FOR will correlate more strongly than the Original Responsibility scale taken under a work FOR.

*Effect of item rewording.*

Hypotheses 5a through 6b address the effects of item wording on criterion validity. Building on the rationale given for Hypotheses 3a-4c, I expected that, compared with the Original scales, scales that make use of context-relevant expressions (i.e., the Context-Relevant scales) will result in stronger correlations with the criteria described above.

*Work FOR.* Within the Work FOR condition, the following hypotheses address the impact of the Work-Relevant scales on criterion validity.

*H5a:* Within the Work FOR condition, compared with the Original Industriousness scale the Work-Relevant Industriousness scale will be more strongly correlated with in-role behavior.

*H5b:* Within the Work FOR condition, compared with the Original Order scale the Work-Relevant Order scale will be more strongly correlated with in-role behavior.

*H5c:* Within the Work FOR condition, compared with the Original Responsibility scale the Work-Relevant Responsibility scale will be more strongly correlated with in-role behavior.

*H5d:* Within the Work FOR condition, compared with the Original Responsibility scale the Work-Relevant Responsibility scale will be more strongly correlated with organizational citizenship behavior.

*School FOR.* Within the School FOR condition, the following hypotheses address the impact of the School-Relevant scales on criterion validity.

*H6a:* Within the School FOR condition, compared with the Original Industriousness scale the School-Relevant Industriousness scale will be more strongly correlated with college GPA.

*H6b.* Within the School FOR condition, compared with the Original Order scale the School-Relevant Order scale will be more strongly correlated with college GPA.

*H6c.* Within the School FOR condition, compared with the Original Responsibility scale the School-Relevant Responsibility scale will be more strongly correlated with college GPA.

Tests of Hypotheses 3a through 6b were conducted using the program DEPCOR (Silver, Hittner, & May, 2006), which implements Williams' (1959) test for the significance of the difference between two dependent correlations. Williams' method has been recommended by both Neill and Dunn (1975) and Krishnamoorthy and Xia (2007), based on the results of Monte Carlo simulations comparing various tests of dependent correlations. The criterion used to infer statistical significance was a one-tailed  $p < .05$ .

*Applicant reactions.* The literature on applicant reactions suggests that, relative to other types of assessment, personality tests elicit more negative applicant reactions. Among the consequences of negative applicant reactions are withdrawal from the application process (Chan, Schmitt, Jennings, Clause, & Delbridge, 1998; Rynes &



Connerly, 1993) and a lowered intention to recommend the organization to others (Chan et al., 1998; Gilliland, 1994; Ployhart & Ryan, 1998; Rynes & Connerly, 1993).

Assessments that are perceived as more job-related typically elicit more positive reactions (Rosse, Miller, & Stecher, 1994).

Although there is limited agreement regarding the measurement of applicant reactions, several researchers have utilized measures of the perceived *predictive validity*, *face validity*, and *fairness* of the assessment (e.g., Glode, 2002; Ryan & Sackett, 1987; Smither, Reilly, Millsap, Pearlman, & Stoffey, 1993). All three components of reactions should result directly from one's interpretation and affective response to the content of the assessment (Chan & Schmitt, 2004). Therefore, I expected that the work-relevant inventory would elicit higher mean ratings of predictive validity, face validity, and fairness.

*H7a:* Mean predictive validity, face validity, and fairness ratings will be higher for the Work-Relevant inventory.

The literature also suggests that the perception of the *personal impact* of the assessment may be an important component of reactions to a personality test (Seiler, 2009). One's perception of the personal impact of the assessment is an overall judgment of the invasiveness of the assessment and one's level of discomfort in responding to the items. A test that is perceived as more job-related has been shown to elicit lower ratings of invasiveness and discomfort (Harland, Rauzi, & Biasotto, 1995; Rosse et al., 1994; Ryan & Sackett, 1987).

*H7b:* The mean personal impact rating (when reverse-scored) will be higher for the Work-Relevant inventory.

## CHAPTER 4

### METHOD

#### *Item Relevance Ratings: Work Context*

*Participants.* Participants were a convenience sample of twenty-three volunteers recruited through personal contacts. All participants had between four and over forty years of work experience.

*Instructions and procedures.* The instructions were delivered online as a series of presentation slides.<sup>4</sup> Participants used their own computers to view the online instructions and to enter their ratings and suggested rewrites into a spreadsheet that was provided to them; they then e-mailed their completed spreadsheets to me. Participants were invited to contact me if they had questions about the procedures or instructions. No participants contacted me, but upon viewing the data it was clear that two participants had not understood or followed the instructions (i.e., one participant entered “yes” or “no” for each item instead of using the 0-3 rating scale; one participant’s pattern of ratings was clearly very different from the majority and the participant’s suggested rewrites also indicated a lack of understanding). These two participants’ data were removed from all subsequent analyses.

*Relevance ratings.* Participants rated each Original scale item in terms of their perception of its relevance to work on a scale from 0 (does not describe a work behavior/characteristic at all) to 3 (describes a work behavior/characteristic extremely

---

<sup>4</sup> Prior to sending the study materials to all participants, a subset of three participants volunteered to view the instructions and provided feedback on portions of the instructions that were not clear. I revised the instructions based on their feedback and then confirmed with these volunteers that the instructions were clear.

well). A summary of the instructions for this task are given in Appendix A (*Part 1*).

*Suggest item rewrites.* Participants were also invited to suggest rewrites to items that they rated as less than a 3 in work relevance. These suggestions were referenced during the creation of the Work-Relevant scales that are described below. A summary of the instructions for this task are also given in Appendix A (*Part 2*).

#### *Item Relevance Ratings: School Context*

*Participants.* Twenty participants (who are a subset of the participants described in the frame-of-reference manipulation study described next) provided ratings of the perceived school relevance of each Original scale item.

*Instructions and procedures.* To ensure that participants understood the instructions and were highly attentive during the rating task, I limited the study sessions to no more than five participants at a time. I personally delivered instructions to participants, using a series of PowerPoint slides similar to those given to participants in the work-relevance rating task. Participants were invited to ask questions for clarification at any time. Only one participant indicated that he did not fully understand the instructions; after I answered his questions, this participant confirmed that he understood the task.

*Relevance ratings.* Participants rated each item in terms of their perception of its relevance to school on a scale from 0 (does not describe a school behavior/characteristic at all) to 3 (describes a school behavior/characteristic extremely well). Instructions for this task were similar to those summarized in Appendix A (*Part 1*).

### *Item Specificity Ratings*

Four doctoral I/O psychology students—myself and three others—collaborated to rate the degree of contextual specificity of each of the Original Scale items. Specificity ratings were on the following scale: 0: not at all specific to any particular context(s); 1: could be applied to several contexts equally well; 2: applies mostly to two or three contexts; 3: very specific to one particular context. We each first independently rated the specificity of the item; next, we discussed our ratings and resolved any discrepancies in ratings so that our final ratings differed by no more than one point. Final specificity ratings were calculated as the average of the four ratings. Mean relevance and specificity ratings for each item are given in Tables 1-3.

### *Development of Rewritten Scales and Context-Relevant Scales*

To create the Rewritten scales, several items from each of the three Original Scales were rewritten to use more work-/school-relevant language. Due to practical constraints in the timing of data collection, the method by which items were selected to be rewritten was constrained by unknowns that, if known, would have allowed for a more systematic scale development process; therefore, results based on these rewritten items are considered preliminary and exploratory.

At the time that the Work-Rewritten scales were created, work relevance ratings were known but item specificity ratings had not yet been determined. In short, to create these scales I engaged in informal discussions with two graduate student volunteers to determine which items might merit rewriting; during these discussions, we considered both the item's average relevance rating and our own informal ratings of item specificity. We opted to rewrite items that were either 1) low in relevance, or 2) at least moderate in

relevance but low in specificity. During the process of item rewriting, we referenced the suggested rewrites provided by participants in the work-relevance rating task; edited versions of many of these suggested items were included in the resulting Work-Rewritten scales. The number of items composing each of the Rewritten Scales varies by facet.

A similar process was implemented to create the School-Rewritten Scales. However, because the school relevance ratings were collected while the frame-of-reference study was already underway, we instead used our own informal judgments of each item's relevance to school. In other words, both specificity and perceived relevance were unknowns. Tables 4-6 present both the Work-Rewritten and School-Rewritten scale items, alongside their Original scale counterparts.

#### *Frame-of-Reference Manipulation*

*Participants.* Participants were students enrolled in psychology courses at a large Midwestern university; all participants were given course credit or extra credit for their participation. All participants ( $N = 475$ ) responded under a work frame-of-reference to the Original Scale and Work-Relevant Scale items. Three-hundred ninety-four (394) of the same participants also responded under a school frame-of-reference to the Original scale and School-Relevant scale items.

Complete demographic information was available for the majority of participants. Sixty-one percent (61%) of the sample is female, and the average age is 19.8 ( $SD = 1.46$ ). The majority of participants are White (62%), followed by Asian or Pacific Islander (24%), Hispanic (6%), and Black (4%); the remaining four percent of participants are Native American, reported more than one ethnicity, or did not report their ethnicity.

Participants also reported the extent and nature of their work experience. Only six percent (6%) of participants reported that they had no work experience, and only sixteen percent (16%) reported less than one year of work experience. The remaining participants (78%) reported having between one and two years of experience (31%), between three and five years of experience (36%), or over five years of experience (11%). Of those participants who reported having at least some work experience, ninety-four percent (96%) described having held at least one “real” job, whereas the remaining participants described work experiences such as tutoring, volunteering, babysitting, or helping with a family business or family farm.

*Frame-of-reference instructions.* Several steps were taken to ensure that the FOR instructions were highly salient to participants.<sup>5</sup> The FOR instructions were presented as full-page cover sheets and included pictures relevant to the context (i.e., the Work FOR instruction page included a picture of a man using a copier and a woman punching a time clock; the School FOR instruction page included a picture of a student reading a book and a teacher lecturing). In addition, the words “Work Context” or “School Context” were included in large font in the margins of each page of the personality questionnaire

---

<sup>5</sup> Although some researchers (e.g., Robie et al., 2000) have used an “at work” tag at the end of every personality statement to contextualize the assessment, I opted to instead use frame-of-reference instructions and to make the frame-of-reference as salient as possible. Adding “at work” to every item could easily increase participant fatigue. This strategy also serves to isolate the effects of perceived relevance on discrimination from the possible effect that an item might become less discriminating simply because it “sounds strange”, for example, “I see no reason for belonging to clubs or community organizations *at work*.” In addition, I wanted to retain the original item in its unaltered form so that the assessment resembled as much as possible how a personality assessment is typically presented to an applicant or incumbent in a workplace setting. A similar procedure to that used in this dissertation was implemented by Hunthausen, Truxillo, Bauer, & Hammer (2003) and by Lievens et al., (2008), who found that a maximum of only ten percent of participants likely did not use the instructed frame-of-reference when responding (I suspect that a disproportionate number of this ten percent were those in the study’s *general* condition).

packet. Before beginning the study, participants first received a brief verbal introduction to alert them to the frame-of-reference instructions. They were told that they would respond to personality items in both a work and a school condition and were given some time to locate the two instruction pages. Participants were asked to read the instructions carefully, and to keep the instructions in mind as they responded to each item. They were informed that many of the questions would be the same across the two conditions.

The written FOR instructions were as follows:

Work FOR:

Imagine that you are at work, and your employer has asked you to respond to the questionnaire that follows. You are uncertain as to why your employer has asked you to do this, but you were told that it is required. You can think about yourself in terms of a job you currently have, one you held in the past, or a combination of jobs you have held. Your employer gave you the following instructions: “On the following pages, there are statements and words describing people's characteristics and behaviors. Please consider how these statements describe you when you think about yourself **as an employee.**”

School FOR:

Imagine that your advisor, who is doing research on the characteristics of college students, has asked you to respond to the questionnaire that follows. Your advisor gave you the following instructions: “On the following pages, there are statements and words describing people's

characteristics and behaviors. Please consider how these statements describe you when you think about yourself **as a student**.”

The order of presentation of the Work/School FOR instructions was counterbalanced. Similarly, within each FOR condition, half of the participants responded to the Original scales first and half responded to the Rewritten scales first.

*Job performance.* Self-reported job performance was measured using Williams and Anderson’s (1991) *in-role behavior* and *organizational citizenship behavior* scales. Both of these factor-analytically derived scales consist of 7 items; the items are given Appendix D. Responses to each item were made on a scale from 1 (Strongly Disagree) to 5 (Strongly Agree).

*Test Reactions.* Based on the results of previous research (Seiler & Kuncel, 2005; 2007), I have adapted items from several measures of test reactions in order to measure reactions to a personality assessment (Seiler, 2009). The scales composing a 28-item version of the questionnaire are described below. The items are given in Appendix E.

1. Predictive Validity (6 items). Measures the perception of the actual utility of an assessment. A representative item is, “I am confident that the assessment can predict how well an applicant will perform on the job.” Items were adapted from Smither et al. (1993) and Glode (2002).
2. Face Validity (5 items). Measures the belief that the content of a personality test is related to behaviors that are required in a job. A representative item is, “The actual content of the personality instrument is clearly related to job tasks.” Items were adapted from Smither et al. (1993).



3. Fairness Perceptions (8 items). Measures the evaluation of the appropriateness of using the assessment to select employees. A representative item is, “I feel that using the test to select applicants for a job is unfair.” Items were adapted from Smither et al. (1993), Glode (2002), Rynes and Connerly (1993), and Ryan and Sackett (1987).
4. Personal Impact (9 items). Represents the perception of the invasiveness of the assessment and the subject’s level of comfort in responding to the items in the assessment. Representative items are “This type of assessment asks too much personal information” and “I would be comfortable taking this test in a work setting.” Items were adapted from Harland et al. (1995), Rosse et al. (1994), and Ryan and Sackett (1987).

*Participants.* Participants were forty-four students enrolled in psychology courses at a large Midwestern university; all participants were given course credit or extra credit for their participation. Twenty-three of the participants completed the test reactions measure in reference to their responses to the Original Scales under a work FOR. The other twenty-one participants completed the measure after responding to the Work-Relevant Scales under a work FOR. Participants were asked verbally to indicate whether they had work experience; only one participant indicated having no work experience.

#### *IRT Parameter Estimation*

The 2-parameter logistic (2PL) model has been used with some success to model personality data at the item level (e.g., Reise & Waller, 1990; Zumbo et al., 1997). This model estimates a difficulty ( $b$ ) and discrimination ( $a$ ) parameter for each item in a scale

using dichotomous (0 = incorrect; 1 = correct) response data. Data that are polytomous are artificially dichotomized.<sup>6</sup>

2PL item parameters were estimated using the BILOG-MG (Version 3.0) program. Priors for estimation of item slopes were removed, and slope estimates were set to the logistic metric. (i.e., all normal ogive slope estimates are multiplied by a scaling factor of approximately 1.7). All other options were set to their defaults. As the 2PL model requires dichotomous data, the raw subject response data were transformed such that Strongly Disagree, Disagree, and Neutral<sup>7</sup> responses were recoded as 0, and Agree and Strongly Agree were recoded as 1. All negatively worded items were reverse-scored.

---

<sup>6</sup> The use of less restrictive models such as graded response models or ideal point models is, in my opinion, outside of the scope of this dissertation. The graded response model would be most relevant for an investigation concerning the difficulty parameter, not the discrimination parameter (i.e., a graded response model provides  $m-1$  estimates of item difficulty but only one estimate of item discrimination; all else being equal, discrimination estimates for the 2PL model and graded response model are equivalent). An ideal point model would be most relevant for examining items that could be expected to display unfolding (i.e., items that individuals having “moderate” levels of a trait are more likely to endorse than individuals either high or low on the trait). Given that this dissertation made use of scales that were developed using classical test theory methodology, the 2PL model was expected to fit the data adequately (cf. Chernyshenko, 2002; Chernyshenko et al., 2007; Drasgow et al., 2010).

<sup>7</sup> Huang, Church, and Katigbak (1997) followed the same scoring convention, citing evidence that the “?” response to the Job-Descriptive Index is more indicative of a negative response.

## CHAPTER 5

### RESULTS

#### *Test of Unidimensionality*

Most IRT models, including the 2PL model, make the assumption of scale unidimensionality. For each of the six Conscientiousness facet scales (the three Work FOR Original scales and the three School FOR Original scales), a parallel analysis was conducted on the raw (non-dichotomized) data to test this assumption. On examining the scree plots, the second eigenvalue of each scale was found to be approximately equal to the second eigenvalue from the simulated data; these results suggest that all scales are sufficiently unidimensional for the purpose of IRT modeling in the current study. Figures 2-7 depict the results of the parallel analyses.

#### *Tests of Model Fit*

As a test of model/data fit at the item level, Drasgow, Levine, Tsien, Williams, and Mead (1995) recommended the use of  $\chi^2/\text{df}$  ratios for item doubles and triples. Whereas  $\chi^2/\text{df}$  ratios calculated for single items can be affected by the compensatory effects of local item misfits, ratios for item doubles and triples are more sensitive to this effect. As such, large  $\chi^2/\text{df}$  ratios for item doubles and triples are useful for testing misfitting items due to violations of unidimensionality and local independence. For descriptive purposes, Tables 7 and 8 present frequency tables and summary statistics of the  $\chi^2/\text{df}$  ratios for all item singles, pairs, and triples within each 20-item scale. As can be seen, ratios for item singles were all less than 1 for all scales, as were a large percentage of the ratios for item pairs and triples. The mean  $\chi^2/\text{df}$  ratios for singles,

doubles, and triples within each scale were all well below the 3.0 criterion deemed by Drasgow et al. (1995) to indicate misfit<sup>8</sup>.

#### *Impact of Relevance on Discrimination Estimates*

Hypothesis 1a, which stated that perceived relevance would be positively correlated with item discrimination, was partially supported. In the Work FOR condition, item relevance was positively correlated with item discrimination for the Industriousness ( $r = .45, p = .05$ ) and Responsibility ( $r = .53, p = .02$ ) scales but not for the Order scale ( $r = .09, p = .70$ ). As can be seen in Table 1, the relevance rating for all but one Industriousness item (Item 5, “Getting average grades is enough for me.”) was uniformly high; and yet, even with that item removed the correlation between relevance and discrimination was still in the anticipated direction (although not statistically significant;  $r = .31, p = .19$ ). The correlation for overall Conscientiousness (the mean of the three facet scales) was  $r = .33 (p = .01)$ .

In the School FOR condition, item relevance was significantly positively correlated with item discrimination for the Responsibility scale only ( $r = .51, p = .02$ ). The correlations were in the anticipated direction, but not statistically significant, for Industriousness ( $r = .09, p = .71$ ) and overall Conscientiousness ( $r = 0.21, p = .11$ ). Counter to the stated hypothesis, the correlation was negative, but not significant, for the Order facet ( $r = -0.15, p = .53$ ). Summary statistics for both the Work FOR and School

---

<sup>8</sup> Due to the fact that certain items within each scale are clearly very similar in content, I suspected that many specific pairs and triples would have large  $\chi^2/df$  ratios. Analysis confirmed that  $\chi^2/df$  were in fact very large for many of the suspect item pairs (see Tables 9 and 10), and were never as large for item pairs that are seemingly much less similar in content. However, as it was not my intent to develop a scale with optimal psychometric properties, but rather to investigate the impact of item characteristics on item responses, I opted to retain all items for subsequent analysis.

FOR conditions are given in Table 11.

Overall, the correlation results are relatively consistent across both the Work and School FOR conditions. Specifically, in both conditions, the highest correlation was found for the Responsibility facet. This makes sense, given that the Responsibility facet has the most variability in relevance ratings in both conditions. Although the correlation for Industriousness was significant in the Work FOR but not the School FOR, this again may be due to the relatively low variability in relevance ratings in the School FOR condition. The correlations for the Order facet were the most discrepant between the two conditions; however, given that both values were non-significant, it is difficult to make any inferences regarding these relationships.

Figures 8 and 9 graphically depict the relationship between relevance and discrimination for all three facets within the Work and School FOR conditions, respectively; each facet is represented with a unique symbol to highlight the differential relationships at the facet level, such that it can be seen that item relevance/discrimination relationships were partially impacted by the range of relevance ratings within the facet and were not uniformly linear across facets. This limitation is discussed further in the Discussion section.

#### *Impact of Specificity on Discrimination Estimates*

Hypothesis 1b stated that item specificity would only impact the discrimination values of items that are perceived as less relevant. Again, due to the relatively small number of items available for the interaction analyses, the results that follow should be considered exploratory. However, the results are suggestive of the relationship between

relevance, specificity, and discrimination. The exploratory interaction results for the Work FOR and School FOR conditions are described separately below.

*Work FOR interaction results.* Given the lack of variability in relevance ratings for Industriousness within the Work FOR condition (using a midpoint split, only Item 5 was coded as *less relevant*), this interaction was not considered. As can be seen in Figure 10, the graphical depiction of relevance and specificity for the Work FOR Order facet does not support the hypothesized interaction. Figure 11 depicts the results for the Work FOR Responsibility facet. This graph suggests a possible interaction between perceived relevance and specificity that is not of the hypothesized form; items that were more specific and more relevant had slightly higher discrimination values than their less specific counterparts. Finally, Figure 12 depicts a possible interaction for overall Work FOR Conscientiousness. This graph appears to indicate that item relevance impacted discrimination estimates for only those items coded as more specific.

*School FOR interaction results.* Figures 13-16 show the graphical depiction of relevance and specificity for School FOR Industriousness, Order, and Responsibility, and overall Conscientiousness. None of the graphs are suggestive of a relevance/specificity interaction.

#### *Rewritten Item Results*

*Scale Analyses.* The convergent and discriminant validity of the Rewritten scales were investigated by examining scale intercorrelations and exploratory factor analysis (EFA) results, using the primary sample data. Taken together, the pattern of intercorrelations and factor analytic results indicate that the Rewritten scales demonstrate moderate to strong levels of convergent and discriminant validity.

As can be seen in Table 12, the correlation for each Rewritten scale was highest with its Original scale counterpart, providing support for convergent validity. Although an examination of discriminant validity proves difficult when examining three highly intercorrelated facets belonging to the same trait, some support for discriminant validity can be found in that the correlations amongst the Rewritten scales are similar in magnitude to the correlations amongst the Original scales. In other words, the Rewritten scales are distinguishable from one another to a similar degree as the Original scales are distinguishable from one another.

The EFA procedure served to test whether Rewritten scale items share common variance with their corresponding Original scale items, as opposed to with other Rewritten items as part of a “work-relevant” factor. For each facet and for both FOR conditions, the following sets of items were analyzed together in a two factor solution using maximum likelihood factor analysis with direct oblimin rotation: items from the Rewritten scale, its Original scale counterpart, and items from the Original scale having the lowest correlation with the Rewritten scale. I chose to include only the items from the lowest correlated Rewritten scale due to my concern that, because all of the scales are highly intercorrelated, including the third, more strongly correlated Rewritten scale would yield uninterpretable patterns of factor loadings or might even fail to converge.<sup>9</sup>

Tables 13-18 display the EFA results. Across all analyses, the pattern of factor loadings provides moderate to strong support to make the case that Rewritten scale items

---

<sup>9</sup> This concern was legitimate: three-factor solutions involving the Original scale items and all of the Rewritten items resulted in a number of Original and Rewritten items splitting their loadings equally across two or three factors in an uninterpretable pattern. Solutions were equally uninterpretable across a number of alternative rotation methods.

are measures of their respective facets, as opposed to measures of a “context-relevant” construct. For example, for Work FOR Industriousness, all Original and Rewritten scale items had higher loadings on the first factor, and smaller, frequently near-zero loadings on the second factor. For those items that did have nonzero loadings on the second factor, there is no apparent similarity in content amongst those items or in relation to the comparison facet, Order. Three of the Rewritten Order items, which were expected to load highly on the second factor, did have higher loadings on the first factor. However, these items also had nonzero loadings on the second factor. On inspecting the content of these three items, there are no apparent similarities in content amongst the items or in relation to Industriousness. Similar results were found for the remaining five scales (Work-Rewritten Order and Responsibility; School-Rewritten Industriousness, Order, and Responsibility).

*Item discrimination analyses.* Discrimination values for the Rewritten Scale items were obtained by running a separate 2PL model to estimate each item individually. For example, thirteen Original Industriousness Scale items were rewritten for the Work FOR condition; therefore, thirteen individual 2PL models—each estimating one of these Rewritten items—were run. In other words, parameter estimates for each Rewritten item were estimated with reference to the item’s Original Scale, not with reference to its Rewritten Scale (as would be the case if a single 2PL model was run on only the rewritten items) or a combination of Original Scale and Rewritten Scale items. As the aim was to compare the discrimination estimate of each rewritten item with its Original Scale counterpart, this procedure serves to ensure that both estimates are on the same metric. The procedure was implemented as follows:



1. For each Original Scale item that was rewritten, the subject response vector for that item was replaced with the response vector of its rewritten counterpart. The remaining nineteen response vectors remained unchanged.
2. A 2PL model was run with the nineteen Original Scale parameter estimates constrained to be equal to their initial estimates. Only the rewritten item's parameter estimates were unconstrained.

Hypothesis 2a, which stated that, within both the work and school FOR conditions, the Rewritten items would have larger discrimination values than their Original counterparts, was not supported. Across all Work-Rewritten Industriousness items, the average change in discrimination (calculated as the Rewritten scale item discrimination minus the Original scale item discrimination) actually favored the Original scale items (average change = -0.34). Similarly, the Original Order items had, on average, higher discrimination values than their work-rewritten counterparts (average change = -0.22). For Responsibility, the average change in discrimination was effectively zero (average change = -.01). These results are presented in Tables 19-21.

Similar results were found for the School-Rewritten items. The average change in discrimination favored the Original scale Industriousness, Order, and Responsibility items (average change = -0.44, -0.42, and -0.14, respectively). These results are presented in Tables 22-24.

I also examined the change in discrimination for only those Original items that received low relevance ratings. Again, because just one item in the Original Industriousness scale received a low work relevance rating, that scale was omitted from further analysis. Of the six items in the Original Order scale that received low work

relevance ratings and were rewritten, only two of the six corresponding rewritten items had higher discrimination values. Similarly, of the seven items in the Original Responsibility scale that received low work relevance ratings and were rewritten, only one of the seven corresponding rewritten items had a higher discrimination value.

In the School FOR condition, similar results were found. Of the seven items in the Original Industriousness scale that received low school relevance ratings and were rewritten, only one of the corresponding rewritten items had a higher discrimination value. Of the fourteen Original Order items that received low school relevance ratings and were rewritten, only four of the corresponding rewritten items had higher discrimination values. Finally, of the seven Responsibility items that received low school relevance ratings and were rewritten, only one of the corresponding rewritten items had a higher discrimination value.

In light of the null results for Hypothesis 2a, Hypothesis 2b, which stated that the increase in discrimination will be largest for the most specific items, could not be tested as stated. However, partial support is found for the Work FOR Order and Responsibility scales in that an increase in discrimination was not *largest* for the most specific items—the increase was instead present *only for those items*. Across all Work-Rewritten Order items, the average change in discrimination for items coded as Less Specific was -.32, whereas the average difference in discrimination for items coded as More Specific was .21. Using the non-dichotomized specificity data, this corresponds to a correlation of .45 between item specificity and change in discrimination. For the Responsibility scale, across all Rewritten items the average difference in discrimination for items coded as Less Specific was -.18, whereas the average difference in discrimination for items coded

as More Specific was .16. Using the non-dichotomized specificity data, this corresponds to a correlation of  $r = .60$  between item specificity and change in discrimination.

Clearly, these results are purely exploratory and should be interpreted with caution, particularly in light of the small number of items that contributed to the results and the inconsistent results across scales. In retrospect, however, the results do potentially make some sense. Although not hypothesized, it appears that item specificity had at least a moderate negative correlation with item discrimination ( $r = -0.18$  in the Work FOR;  $r = -0.17$  in the School FOR). It is quite possible that this effect is present because the more specific an item, the more likely it is that the item will be inapplicable to some respondents (cf., Gordon & Holden, 1998); this could in turn evoke random responding—particularly for those individuals who attend to the literal content of items (cf., Carter, Lake, & Zickar, 2010). By default, when an item is rewritten to appear more relevant to a context, it will become more specific. Items that are low in perceived relevance but broad will become substantially more specific when rewritten, relative to items that are low in perceived relevance but specific. Consequently, any discrimination gain achieved by rewriting a broad item to become more context-relevant might be offset by a loss in discrimination due to item specificity. When the original item is already fairly specific, this same offset will not occur. This is an important area for future research.

Figure 17 provides an illustrative example. The first figure shows the relatively specific Original item, “Although I think charitable causes are important, I rarely become involved in charitable activities,” in relation to its rewritten counterpart, “Although I understand that coworkers sometimes need assistance with a task, I see no reason for helping them more than is required.” As can be seen, the rewritten item has a higher

discrimination estimate. In contrast, the relatively broad Original item, “I frequently do not put things in their proper place” has a higher discrimination estimate than its rewritten counterpart, “I frequently do not put work materials back where they belong.”

### *Criterion Validity Results*

*Effect of frame-of-reference.* Correlation values that were used for the Work FOR and School FOR criterion validity analyses are summarized in Table 12.

Hypothesis 3a, which stated that the Work FOR Original Industriousness/in-role behavior correlation would be larger than the corresponding correlation in the School FOR condition, was supported ( $r = .52$  and  $.41$ , respectively;  $t = 3.27$ , one-tailed  $p = .00$ ).

Hypothesis 3b, which stated that the Work FOR Original Order/in-role behavior correlation would be larger than the corresponding correlation in the School FOR condition, was supported ( $r = .33$  and  $.28$ , respectively;  $t = 1.95$ , one-tailed  $p = .03$ ).

Hypothesis 3c, which stated that the Work FOR Original Responsibility/in-role behavior correlation would be larger than the corresponding School FOR correlation was supported ( $r = .50$  and  $.45$ , respectively;  $t = 1.71$ , one-tailed  $p = .04$ ).

Hypothesis 3d, which stated that the Work FOR Original Responsibility/OCBI correlation would be larger than the corresponding correlation in the School FOR condition, was supported ( $r = .48$  and  $.40$ , respectively;  $t = 2.69$ , one-tailed  $p = .00$ ).

Hypotheses 4a, which stated that the School FOR Original Industriousness/college GPA correlation would be larger than the corresponding Work FOR correlation, was supported ( $r = .22$  and  $.12$ , respectively;  $t = 2.26$ , one-tailed  $p = .01$ ).

Hypothesis 4b, which stated that the School FOR Original Order/college GPA correlation would be larger than the corresponding Work FOR correlation, was not

supported; the School FOR Original Order/college GPA correlation ( $r = .07$ ) is equal to the corresponding Work FOR correlation ( $r = .07$ ).

Hypothesis 4c, which stated that the School FOR Original Responsibility/college GPA correlation would be larger than the corresponding Work FOR correlation, was partially supported. The School FOR Original Responsibility/college GPA correlation is not significantly larger than the corresponding Work FOR correlation ( $r = .19$  and  $.13$ , respectively;  $t = 1.56$ , one-tailed  $p = .06$ ), but the difference is in the anticipated direction.

*Effect of item rewriting.* The values for the tests of Hypotheses 5a-6d, which collectively stated that the Work-Relevant scales would have higher correlations with relevant criteria, are given in Table 12. None of these hypotheses were supported.

In retrospect, given the small or null effects found in the tests of Hypotheses 3a-4c, it is not surprising that the Work-Relevant scale hypotheses received no support. That is, if the effects are small or null when comparing measures across two distinct contexts, it would be surprising to find larger effects when comparing measures within the same context. Despite the null results, however, it can be seen that criterion validity coefficients did not *decrease* for the Context-Relevant scales. Combined with the test reactions results described next, this provides evidence that context-relevant item development warrants future research.

#### *Test Reactions*

The test reactions results are given in Table 25. Hypothesis 7a, which stated that the Work-Rewritten scales would elicit more positive applicant reactions than the Original scales, was partially supported. The mean predictive validity, face validity, and

fairness ratings for the Work-Relevant Scales were higher than the mean ratings for the Original Scales (Cohen's  $d = .35, .53, \text{ and } .32$  respectively). The predictive validity mean difference ( $t(42) = 1.13$ , one-tailed  $p = .13$ ) and fairness mean difference ( $t(40) = 1.07$ , one-tailed  $p = .15$ ) were not significant. The face validity mean difference was significant ( $t(42) = 1.74$ , one-tailed  $p = .04$ ). The mean difference for the average of these three reactions scales combined was significant (Cohen's  $d = .62$ ;  $t(42) = 2.03$ , one-tailed  $p = .03$ ).

The result for Hypothesis 7b, which stated that the perception of personal impact would be more positive (when reverse-scored) for the Work-Relevant scales, was not statistically significant ( $t(40) = 1.51$ , one-tailed  $p = .07$ ). However, the mean difference in Personal Impact was in the anticipated direction (Cohen's  $d = .47$ ).

Given the small sample size used to test hypotheses 7a-7b, I caution against using significance values to infer the presence or absence of any relationships. Overall, the effect sizes do indicate that the Work-Rewritten scales elicit more positive applicant reactions.

## CHAPTER 6

### DISCUSSION

Over the past 20 years, a large body of research has amassed to address the question of how, and under what conditions, personality is related to behavior in the workplace. Seminal research at the beginning of this time period undoubtedly fueled an increase in the use of personality assessments by industrial and organizational psychologists and other human resources professionals (Rothstein & Goffin, 2006), but despite significant progress in the ensuing years, there is still considerable uncertainty among some regarding the utility of personality assessment and the meaning of personality in relation to job performance—in particular when measured via self-report. If one looks only at the early meta-analytic summaries of personality/job performance relationships, it is certainly possible to conclude that researchers have not proved its merits, given the many commentaries pointing to limitations in these early studies.

Looking beyond these early meta-analyses, however, one is left with an even stronger case for the utility of personality assessment. More recent research has shown, for example, that adopting a facet-level approach to measurement, examining only the relationships that reflect theoretically meaningful trait/performance relationships, and inclusion of moderator variables such as organizational characteristics and job functions, all serve to strengthen the observed relationships between personality and work behavior. Even more recently, research has highlighted the potential to further improve personality assessment by modifying the wording of items to facilitate accurate interpretation by respondents. This emerging research supports the notion that item characteristics and contextualization interact with individual differences in response processes to affect the

measurement of personality. This dissertation contributes to this area of emerging research by addressing the question of whether inventories that were developed to assess personality in general function well in measuring personality in a specific context.

Although inconsistent across facets and FOR conditions (work and school), taken together the results suggest that the perception of a personality item's relevance to the context in which it is being endorsed can have a small to moderate effect on its discrimination. In the work FOR condition, there was some indication that discrimination is also impacted by the context specificity of the item. This result supports the conclusion that items perceived as irrelevant to work do not provide as much useful information for measuring a trait in the context of work, and that this effect is perhaps limited to items that are initially more context-specific. Given this result, I recommend that researchers take care to consider how individual items will be interpreted when respondents are given frame-of-reference instructions, or when they could be expected to impose a frame-of-reference based on the purpose of the assessment.

This research also appears to indicate that changing the wording of an item to more closely reflect the context of interest does not necessarily improve—and may in many cases harm—item discrimination. This result is particularly curious given that the perceived relevance of the Original Scale items was positively correlated with their discrimination values. However, the result does make sense in light of the unanticipated relationship between specificity and discrimination, and there is some indication that low-relevance items that were initially more specific showed improvement in discrimination when rewritten. This result provides additional insight into the research of Lievens et al. (2008), who justified replacing low-relevance items with broader items to facilitate



responding across contexts, and Pace and Brannick (2010), who found that items written to closely reflect the context of work had higher criterion-related validity. Specifically, the current findings suggest that the perceived relevance and specificity of an item should be considered together in determining whether an item is likely to produce desirable measurement properties, or whether one should consider removing or rewriting the item. Although speculative, this research suggests that a well-constructed inventory of broad statements should function well for assessing personality both within a specific context and across contexts. However, if the use of more specific items is desired (for example, to provide richer information on the trait as it is displayed within a particular context) researchers will then need to ensure that the items are perceived as relevant to the context of interest, while also ensuring that the items are not assessing overly narrow behaviors. If the researcher desires to retain context-specific items while also comparing a trait across contexts, this may require that the actual wording of these items vary across contexts.

Whereas the results for the rewritten items were inconclusive in terms of their potential for improving the measurement of contextualized personality, the test reactions and criterion validity results do provide a case for further exploration into the utility of adopting a context-driven approach to item development, in particular for use within selection settings. Specifically, reactions to the Work-Relevant inventory were more positive than reactions to the Original inventory, and the criterion validity results indicate that the Work-Relevant scales were at least as strongly correlated with relevant criteria. Even if it is deemed inappropriate to rely exclusively on context-relevant (and thus, more specific) items, it could be that these items are an excellent supplement to a general

personality inventory. For example, Rosse, Miller, and Stecher (1994) found that reactions to a personality inventory were more positive when taken in conjunction with an ability test. It appeared that the inclusion of a test that was perceived as more relevant offset some of the negative reactions to the personality inventory.

### *Limitations and Future Research*

This research has some important limitations. First, the small number of items per scale used in this study limits the confidence one can have in the main effects that were found. The number of items used also precluded the ability to conduct statistical tests of the differences in slopes for the dichotomized variables, or to conduct multiple regression analyses to test for interactions in the non-dichotomized data. Although the ideal research design to test for both main effects and interactions would involve administering a very large, unidimensional set of items that are balanced in terms of perceived relevance and specificity, administering such a large scale to each participant would not be practical. As an alternative, future research could make use of a nested design, whereby subsets of items, some overlapping, are administered to different groups of participants. There are potentially rich pools of such data already available within several I/O consulting organizations and academic research programs.

A second limitation is that the process for creating the Rewritten items was a highly unstructured process. In retrospect, given that there was limited information on the characteristics of each item at the time that the Rewritten items were developed, it would have been best to create a Rewritten version of every Original item instead of choosing to rewrite only a subset of items. Ideally, the Rewritten items would have then gone through a number of rounds of development in order to ensure that each item was judged by

subject matter experts to measure a characteristic as similar as possible to its Original counterpart and that the item did not become too complex or more specific than was necessary for contextualizing the item.

Related to the second limitation, perceived relevance ratings and specificity ratings were not collected for the Rewritten items; as such, it was not possible to examine precisely the degree of change in relevance and specificity for each item. This information would be useful both to confirm that the Rewritten items were actually perceived as more relevant (and, to what extent) and for understanding more precisely the conditions under which changes to perceived relevance and specificity might contribute to a gain or loss in discrimination.

A third limitation of the current study is that results from a sample of students may not generalize to full-time workers. As work is less central to a student's life, the process of accessing thoughts about work is likely to be a more difficult task for students. If this is the case, any effects that *are* found in the current research are likely to be underestimates of the effects that would be found with a sample with a more nuanced understanding of how their traits vary across contexts. However, is also possible that there are age differences in the impact that item characteristics have on self-reports and that the results of the current study are *most* relevant to those with less job experience. This is an important area for future investigation, as it has been shown that giving an equivocal response to a personality item might be an indication of immaturity (Hernández, Drasgow, & González-Romá, 2004). Future research could explore whether personality stability and change over time impact the effects of contextualization.

A related concern that might be raised by I/O practitioners is the degree to which any of the findings from this research would generalize to the assessment of job applicants or to an assessment used to provide developmental feedback to incumbents. Although not the focus of this dissertation, some will likely argue that it would be more valuable for the field to investigate the impact of perceived relevance and item specificity using a sample of applicants or incumbents. This concern, although legitimate, is in no way at odds with the results that were found. Instead, the results suggest that it would be worthwhile to attempt to replicate and extend the findings using applicant and incumbent data. It would be particularly valuable to investigate how perceived relevance and specificity might impact response strategy (e.g., one's decision to lie, exaggerate, selectively sample from their own past behaviors, etc., in order to present a favorable impression). Despite concerns that obvious (i.e., less subtle) items will elicit faking behavior, research generally supports the opposite conclusion—assessments perceived as job-related are more likely to promote an accurate response strategy (e.g., Ingerick, McFarland, Vasilopoulos, & Cucina, 2004; Seiler & Kuncel, 2005; 2007). It could also be that applicants and/or incumbents are less willing to falsify their responses to very specific, work-relevant items out of concern that the employer could easily verify their answers.

A final limitation that warrants mention is that only facets of Conscientiousness were examined. Although this is a practical place to begin, it does seem that Conscientiousness might not be as prone to problems due to perceived irrelevance, compared with other traits (Pace & Brannick, 2010). For example, items from certain facets of Openness to Experience, such as the Aesthetics and Fantasy facets of the NEO

PI-R, would likely receive much lower relevance ratings overall than any of the facets of Conscientiousness included in this dissertation. Given the results of the current research and the discussion by Pace and Brannick (2010), it could instead be that the item content itself is simply inappropriate for assessing these facets as they are displayed in the workplace. Outside the study by Pace and Brannick (2010), this possibility has received little to no attention within the literature that attempts to explain why Openness to Experience is often not strongly related to performance in most jobs (for example, see the review of the relevant literature by Griffin & Hesketh, 2004). More research is needed to investigate the potential of work-relevant item content for improving the observed relationships between personality and theoretically relevant job performance criteria.

### *Conclusion*

Recent research has helped to clarify the way in which we interpret the meaning of personality assessment scores in relation to context-specific criteria such as job performance. Of particular interest is whether personality items that are appropriate for measuring personality in general are the most useful or sufficient descriptors for measuring personality in the context of work. Despite the progress that has been made over the past 20 years, there are many promising avenues for further improvement in which we acknowledge that respondents do not interpret or respond to personality items how we as researchers would like them to. I expect that a personality assessment that “looks like” it is relevant for the context in which it is being used will not only help facilitate accurate responding, but will also garner much-needed support from many of those—scientists and stakeholders included—who view personality assessment with suspicion.

## TABLES

Table 1. Original Industriousness Items and Item Characteristics

	Item	Specificity Rating	Work FOR			School FOR		
			2PL Estimate a	b	Relevance Rating	2PL Estimate a	b	Relevance Rating
1	I lower my standards as tasks get more difficult. (R)	1.75	1.26	-0.91	2.5	1.34	-0.42	2.20
2	I often need a push to get started. (R) *	1.50	1.15	-0.38	2.3	0.96	0.05	2.25
3	I do not work as hard as the majority of people around me. (R)	2.00	1.53	-1.50	2.6	2.01	-0.84	2.15
4	Setting goals and achieving them is not very important to me. (R)	1.00	1.52	-1.79	2.3	1.48	-1.86	2.15
5	Getting average grades is enough for me. (R)	3.00	0.78	-2.21	1.3	1.08	-1.76	2.50
6	I exert medium effort on most tasks I do. (R)	2.00	1.76	-0.70	2.5	1.94	-0.64	2.15
7	I do what is required, but rarely anything more. (R)	1.50	1.72	-0.89	2.5	2.11	-0.42	2.35
8	I can never guarantee completion of even the simplest projects. (R)	1.50	0.95	-3.45	2.5	1.41	-2.61	1.90
9	I have high standards and work toward them.	1.75	2.52	-1.01	2.9	3.38	-0.96	2.55
10	I handle even difficult tasks smoothly.	1.25	1.31	-0.65	2.9	1.09	-0.44	2.10
11	I go above and beyond what is required.	1.50	2.69	-0.36	2.8	2.88	-0.16	2.20
12	I strive to improve myself.	0.50	2.11	-1.56	2.5	1.69	-1.73	2.00
13	I am a very competent person.	0.75	1.25	-1.62	2.8	1.10	-1.71	1.90
14	I make every effort to do more than what is expected of me.	1.25	2.17	-0.53	2.8	2.42	-0.39	1.85
15	I demand the highest quality in everything I do.	0.50	2.97	-0.62	2.7	2.84	-0.55	2.10

Table 1, continued

Item		Specificity Rating	Work FOR			School FOR		
			2PL Estimate		Relevance Rating	2PL Estimate		Relevance Rating
			a	b		a	b	
16	I am effective at my work.	2.50	1.66	-1.80	2.8	1.28	-1.88	2.30
17	I set extremely high standards for myself.	0.25	2.53	-0.58	2.5	2.40	-0.54	2.30
18	I try to be the best at anything I do.	0.00	1.88	-0.99	2.6	2.03	-0.81	2.05
19	I often contribute little to group projects. (R)	2.00	0.91	-2.39	2.6	0.96	-2.37	2.20
20	I invest little effort in my work. (R)	2.00	1.45	-2.16	2.6	2.32	-1.63	1.80

*Notes.* (R) denotes a reverse-scored item.

\* In order to conform to the assumptions of the 2PL model, the qualifier “about half the time,” which was used in the original item, was replaced with the more definitive “often.”

Table 2. Original Order Items and Item Characteristics

	Item	Specificity Rating	Work FOR			School FOR		
			2PL Estimate a	b	Relevance Rating	2PL Estimate a	b	Relevance Rating
1	I frequently forget to put things back in their proper place. (R)	1.25	1.16	-1.49	1.6	1.75	-0.82	0.80
2	I prefer keeping my options open and rarely plan in advance. (R)	0.50	1.01	-0.55	1.7	1.11	-0.49	1.10
3	I seldom make detailed "to do" lists. (R)	0.75	1.07	-0.67	1.9	1.12	-0.86	1.65
4	I do not like work spaces that are too clean and tidy. (R)	2.50	2.02	-1.25	2.3	1.54	-1.28	1.20
5	For me, being organized is unimportant. (R)	0.50	2.91	-1.07	2.1	3.60	-1.01	1.30
6	Being neat is not exactly my strength. (R)	0.25	2.75	-0.50	1.8	2.99	-0.28	1.40
7	I frequently do not put things in their proper place. (R) *	0.75	1.95	-1.01	1.8	2.53	-0.60	1.05
8	As long as I have a little bit of clear space on my desk, I am happy to do my work. (R)	2.00	1.30	0.18	2.1	1.38	0.21	2.05
9	I have a daily routine and stick to it.	0.25	0.81	0.10	2.4	1.11	0.27	1.65
10	I need a neat environment in order to work well.	1.75	2.49	-0.18	2.3	2.31	-0.13	2.10
11	I dislike doing things without proper planning.	0.50	1.21	-0.15	2.2	1.32	-0.09	1.70
12	I prefer to do things in a logical order.	0.50	1.05	-1.60	2.4	1.22	-1.51	2.20
13	Organization is a key component of most things I do.	0.75	2.76	-0.45	2.8	3.59	-0.35	2.10
14	I hardly ever lose or misplace things.	0.25	0.89	-0.24	1.9	1.14	0.02	1.15
15	I become annoyed when things around me are disorganized.	0.25	1.81	-0.33	2.1	1.82	-0.44	1.75



Table 2, continued

	Item	Specificity Rating	Work FOR			School FOR		
			2PL Estimate a	b	Relevance Rating	2PL Estimate a	b	Relevance Rating
16	I keep detailed notes of important meetings and lectures.	2.00	0.56	-0.91	2.4	0.81	-1.03	2.80
17	I hate when people are sloppy.	0.25	1.09	-0.68	1.4	1.09	-0.21	1.65
18	Every item in my room and on my desk has its own designated place.	2.50	1.80	-0.11	2.1	1.82	-0.17	1.35
19	Most of the time my belongings are in complete disarray. (R)	2.00	1.82	-0.96	1.0	1.88	-0.82	1.10
20	I feel comfortable even in very disorganized settings. (R)	0.50	1.73	-0.11	1.9	1.73	-0.07	1.20

*Notes.* (R) denotes a reverse-scored item.

\* In order to conform to the assumptions of the 2PL model, the qualifier “half of the time,” which was used in the original item, was replaced with the more definitive “frequently.”

Table 3. Original Responsibility Items and Item Characteristics

	Item	Specificity Rating	Work FOR			School FOR		
			2PL Estimate a	2PL Estimate b	Relevance Rating	2PL Estimate a	2PL Estimate b	Relevance Rating
1	I only think of myself and rarely of others. (R)	0.00	1.25	-1.95	1.8	1.45	-1.80	0.90
2	I am not the most responsible person among my friends. (R)	1.50	1.14	-1.25	1.5	1.20	-1.03	1.20
3	Sometimes it is too much of a bother to do exactly what is promised. (R)	1.00	1.36	-1.00	2.1	1.29	-0.90	1.15
4	If I am running late to an appointment, I may decide not to go at all. (R)	2.75	1.08	-2.16	2.0	0.82	-2.13	1.15
5	Although I think charitable causes are important, I rarely become involved in charitable activities. (R)	3.00	0.58	-0.37	1.0	0.73	-0.36	0.60
6	I am usually not the most responsible group member, but I will not shirk on my duties either. (R)	2.00	1.06	-0.34	2.3	1.03	-0.28	2.30
7	I have a reputation for being late for almost every meeting or event. (R)	0.75	1.03	-1.96	2.6	0.32	-5.12	1.40
8	I carry out my obligations to the best of my ability.	0.25	1.88	-1.68	2.8	2.11	-1.41	2.30
9	I often feel responsible for making sure that all group project assignments are completed.	2.00	2.12	-1.16	2.8	2.07	-1.23	2.90
10	My friends know they can count on me in times of need.	3.00	1.85	-1.92	1.6	1.63	-1.99	1.40
11	When I make mistakes I often blame others. (R)	0.25	1.03	-1.70	2.2	1.08	-1.55	1.00
12	I hold myself accountable for most of my mistakes.	0.25	0.97	-1.63	2.6	1.18	-1.43	1.90
13	When working on a group project, my group members often rely on me to get things done.	2.00	1.51	-0.75	2.7	1.49	-0.76	2.50

Table 3, continued

	Item	Specificity Rating	Work FOR			School FOR		
			2PL Estimate a	2PL Estimate b	Relevance Rating	2PL Estimate a	2PL Estimate b	Relevance Rating
14	I would gladly spend some of my leisure time trying to improve my community.	2.75	0.67	0.01	0.9	0.67	-0.06	1.00
15	If I am running late, I try to call ahead to notify those who are waiting for me.	1.25	1.47	-1.69	2.8	0.96	-2.24	1.45
16	I go out of my way to keep my promises.	0.75	1.23	-1.40	2.5	1.38	-1.48	0.85
17	When others are counting on me, I am extra careful to show up on time.	1.50	1.32	-1.89	2.7	1.12	-2.20	1.85
18	When others depend on me, I try to put in an extra effort.	1.25	1.46	-2.27	2.7	1.35	-2.16	2.25
19	There is no use in doing things for people; they never really appreciate your efforts. (R)	0.00	1.41	-1.92	1.6	1.05	-2.22	0.40
20	I see no reason for belonging to clubs or community organizations. (R)	2.50	1.13	-2.03	1.3	1.44	-1.74	1.15

*Notes.* (R) denotes a reverse-scored item.

Table 4. Work- and School-Rewritten Industriousness Items

Item Number	Original Item	Work Rewrite	School Rewrite
1	I lower my standards as tasks get more difficult.	.	.
2	I often need a push to get started.	It is often hard for me to get started on my work unless a deadline is approaching or I am being monitored.	It is often hard for me to get started on my coursework unless someone is pushing me to do it or it is due soon.
3	I do not work as hard as the majority of people around me.	I do not work as hard as most people do at their jobs.	I do not work as hard as the majority of college students.
4	Setting goals and achieving them is not very important to me.	Setting goals in order to deliver a good work product is not very important to me.	Setting goals in order to be a successful student is not very important to me.
5	Getting average grades is enough for me.	.	.
6	I exert medium effort on most tasks I do.	.	.
7	I do what is required, but rarely anything more.	I only do the minimum work that my boss or supervisor requires of me.	I only do the minimum work required by my instructors.
8	I can never guarantee completion of even the simplest projects.	.	.
9	I have high standards and work toward them.	.	.
10	I handle even difficult tasks smoothly.	I handle even difficult job requirements smoothly.	I competently handle even the most challenging class projects.
11	I go above and beyond what is required.	I go above and beyond the minimum requirements of my job.	My performance exceeds the University's minimum requirements for a degree.
12	I strive to improve myself.	I strive to become an even better employee.	I strive to become an even better student.
13	I am a very competent person.	I am a very competent employee.	I am a very competent student.

Table 4, continued

Item Number	Original Item	Work Rewrite	School Rewrite
14	I make every effort to do more than what is expected of me.	If I finish a required work task early, I will try to make it even better or look for something else to do instead of taking a break.	I work hard to do more than my teachers expect of me.
15	I demand the highest quality in everything I do.	I demand the highest quality from myself when performing job tasks.	I demand the highest quality from myself on class assignments.
16	I am effective at my work.	I complete all of my required work tasks in a thorough and timely manner.	I complete all of my coursework in a thorough and timely manner.
17	I set extremely high standards for myself.	The standards I set for myself are as high or higher than those expected of me by my employer.	The standards I set for myself are as high or higher than those expected of me by my teachers and/or family.
18	I try to be the best at anything I do.	I strive to do the best job I can on all work tasks.	I try to be the best student that I can be.
19	I often contribute little to group projects.	.	.
20	I invest little effort in my work.	.	.

Table 5. Work- and School-Rewritten Order Items

Item Number	Original Item	Work Rewrite	School Rewrite
1	I frequently forget to put things back in their proper place.	I frequently forget to return materials where they belong after I use them.	I frequently forget to put my class notes and handouts into the correct folder or file.
2	I prefer keeping my options open and rarely plan in advance.	I do what seems important to me at the moment instead of planning out what I need to accomplish.	I prefer keeping my options open and rarely plan in advance how I will complete assignments with upcoming deadlines.
3	I seldom make detailed “to do” lists.	Even when it would be helpful, I seldom make a “to do” list of the work I need to complete.	I seldom make detailed “to do” lists to help me complete my schoolwork.
4	I do not like work spaces that are too clean and tidy.	.	.
5	For me, being organized is unimportant.	For me, it is not important to keep things organized while I work.	For me, it is not important to keep my schoolwork organized.
6	Being neat is not exactly my strength.	I am not a very neat and tidy employee.	I do not maintain a very neat and tidy environment.
7	I frequently do not put things in their proper place.	I frequently do not put work materials back where they belong.	When I am done with my schoolwork, I frequently leave a mess of books, notes, and other materials lying around.
8	As long as I have a little bit of clear space on my desk, I am happy to do my work.	I don’t mind working in a messy environment as long as there is a little bit of clear space to work.	.
9	I have a daily routine and stick to it.	I have established a daily routine to help me perform well on the job.	I have a daily routine for completing schoolwork that I stick to.
10	I need a neat environment in order to work well.	.	.
11	I dislike doing things without proper planning.	I do not like to work on a job task before making a good plan.	I dislike studying or completing an assignment without having a good plan.

Table 5, continued

Item Number	Original Item	Work Rewrite	School Rewrite
12	I prefer to do things in a logical order.	I prefer to plan a logical order for completing work assignments before I begin working.	I prefer to plan a logical order for completing my coursework before I begin working.
13	Organization is a key component of most things I do.	.	.
14	I hardly ever lose or misplace things.	I almost always know where I can find the materials that I put away for later use.	I hardly ever lose or misplace class notes and assignments.
15	I become annoyed when things around me are disorganized.	I become annoyed when my workspace becomes disorganized.	I become annoyed when my notes and handouts are disorganized.
16	I keep detailed notes of important meetings and lectures.	.	.
17	I hate when people are sloppy.	I hate when I see coworkers doing sloppy work.	I hate when I see my peers turn in sloppy work.
18	Every item in my room and on my desk has its own designated place.	I make sure that everything in my workspace has its own designated place.	I make sure that everything in my room or office has its own designated place.
19	Most of the time my belongings are in complete disarray.	Most of the time my workspace is in complete disarray.	Most of the time my room or office is in complete disarray.
20	I feel comfortable even in very disorganized settings.	I feel comfortable on the job even if the workplace is very disorganized.	I feel comfortable doing my schoolwork even in very disorganized settings.

Table 6. Work- and School-Rewritten Responsibility Items

Item Number	Original Item	Work Rewrite	School Rewrite
1	I only think of myself and rarely of others.	I don't take my coworkers into consideration when I make decisions.	I don't take my group members' needs into consideration when working in a group.
2	I am not the most responsible person among my friends.	I am not usually the person who can be counted on to get a job done.	I am not the most responsible person when it comes to completing my coursework.
3	Sometimes it is too much of a bother to do exactly what is promised.	Sometimes it is too much of a bother to follow my supervisor's directions exactly.	Sometimes it is too much of a bother to follow the instructions for an assignment exactly.
4	If I am running late to an appointment, I may decide not to go at all.	If I am going to be late for work, I might decide not to go at all.	If I am going to be late to class, I might decide not to go at all.
5	Although I think charitable causes are important, I rarely become involved in charitable activities.	Although I understand that coworkers sometimes need assistance with a task, I see no reason for helping them more than is required.	Although I think that contributing to class discussions makes classtime more valuable to everyone, I rarely do so.
6	I am usually not the most responsible group member, but I will not shirk on my duties either.	I can be depended on to do the work that I am assigned, but nothing more.	.
7	I have a reputation for being late for almost every meeting or event.	.	I have a reputation for being late to classes and meetings.
8	I carry out my obligations to the best of my ability.	I can be counted on to do my assigned work to the best of my ability.	I complete all course requirements to the best of my ability.
9	I often feel responsible for making sure that all group project assignments are completed.	.	.
10	My friends know they can count on me in times of need.	My coworkers know they can count on me to help out when they need my assistance.	My classmates know that they can count on me when they need my help to complete a group assignment.



Table 6, continued

Item Number	Original Item	Work Rewrite	School Rewrite
11	When I make mistakes I often blame others.	When I make a mistake, I will often try to blame the problem on someone else's actions.	When I do poorly on an exam or assignment I often try to blame it on someone or something outside of my control.
12	I hold myself accountable for most of my mistakes.	.	I hold myself accountable if I fail to deliver what I promised to my instructor.
13	When working on a group project, my group members often rely on me to get things done.	.	.
14	I would gladly spend some of my leisure time trying to improve my community.	I would gladly take on extra hours at work to help my employer with a special project or event.	I would gladly spend some of my leisure time helping an instructor to improve a course.
15	If I am running late, I try to call ahead to notify those who are waiting for me.	.	.
16	I go out of my way to keep my promises.	When necessary, I will put in extra time or effort to complete the work I said I would do.	I will go out of my way to do what I promised to my group.
17	When others are counting on me, I am extra careful to show up on time.	.	When others are depending on me to help complete an assignment, I will try to put in an extra effort.
18	When others depend on me, I try to put in an extra effort.	When my contributions will help to ensure the success of everyone on a work team, I will try to put in an extra effort.	When my classmates are counting on me, I am extra careful to show up on time.
19	There is no use in doing things for people; they never really appreciate your efforts.	There is no use in helping out coworkers, because no one will praise me for my efforts.	There is no use for helping out other group members, because it will not benefit me.
20	I see no reason for belonging to clubs or community organizations.	Although I think there are ways that an employee can add value to a company through unique contributions, I rarely do so.	Although I understand that my group members sometimes need help in order to finish an assignment, I see no reason for being the one to do so.

Table 7.

Frequency Table of Nonadjusted  $\chi^2/\text{df}$  Ratios for All Item Singles, Doubles, and Triples

Scale		Range							Mean	SD
		< 1	1 < 2	2 < 3	3 < 4	4 < 5	5 < 7	> 7		
Industriousness	Singles	20	0	0	0	0	0	0	0.0	0.0
	Doubles	153	28	6	2	0	0	1	0.6	0.9
	Triples	733	318	63	11	10	4	1	0.9	0.8
Order	Singles	20	0	0	0	0	0	0	0.0	0.0
	Doubles	156	21	8	1	1	2	1	0.6	1.3
	Triples	730	285	68	34	4	10	9	1.0	1.1
Responsibility	Singles	20	0	0	0	0	0	0	0.0	0.0
	Doubles	158	19	7	2	0	1	3	0.7	1.9
	Triples	761	261	54	13	15	17	19	1.1	1.5

Table 8.

Frequency Table of Nonadjusted  $\chi^2/\text{df}$  Ratios for All Item Singles, Doubles, and Triples: School FOR

Scale		Range							Mean	SD
		< 1	1 < 2	2 < 3	3 < 4	4 < 5	5 < 7	> 7		
Industriousness	Singles	20	0	0	0	0	0	0	0.0	0.0
	Doubles	149	30	8	0	2	1	0	0.6	0.8
	Triples	755	281	82	19	3	0	0	0.9	0.7
Order	Singles	20	0	0	0	0	0	0	0.0	0.0
	Doubles	156	17	7	2	3	3	2	0.7	1.4
	Triples	685	261	94	45	33	18	4	1.2	1.2
Responsibility	Singles	20	0	0	0	0	0	0	0.0	0.0
	Doubles	155	22	4	2	3	3	1	0.7	1.6
	Triples	760	216	78	57	7	7	15	1.1	1.2

Table 9.  
Work FOR Original Item Pairs with Nonadjusted  $\chi^2/\text{df}$  ratios  $\geq 3$

Scale	$\chi^2/\text{df}$ Ratio	Item Pair	
Industriousness	3.1	Getting average grades is enough for me.	I exert medium effort on most tasks I do.
	9.6	I am a very competent person.	I am effective at my work.
Order	3.2	I often contribute little to group projects.	I invest little effort in my work.
	13.3	I frequently forget to put things back in their proper place.	I frequently do not put things in their proper place.
	5.2	I frequently forget to put things back in their proper place.	I hardly ever lose or misplace things.
	4.6	I prefer keeping my options open and rarely plan in advance.	I seldom make detailed "to do" lists.
	5.9	I dislike doing things without proper planning.	I prefer to do things in a logical order.
	3.4	I become annoyed when things around me are disorganized.	I hate when people are sloppy.
Responsibility	5.1	I am not the most responsible person among my friends.	I am not the most responsible group member, but I will not shirk on my duties either.
	20.4	Although I think charitable causes are important, I rarely become involved in charitable activities.	I would gladly spend some of my leisure time trying to improve my community.
	3.3	I am not the most responsible group member, but I will not shirk on my duties either.	When others depend on me, I try to put in an extra effort.
	3.8	I have a reputation for being late for almost every meeting or event.	When others are counting on me, I am extra careful to show up on time.
	8.6	When I make mistakes I often blame others.	I hold myself accountable for most of my mistakes.
	10.6	When others are counting on me, I am extra careful to show up on time.	When others depend on me, I try to put in an extra effort.

Table 10.  
School FOR Original Item Pairs with Nonadjusted  $\chi^2/\text{df}$  ratios  $\geq 3$

Scale	$\chi^2/\text{df}$ Ratio	Item Pair	
Industriousness	4.5	I handle even difficult tasks smoothly.	I am a very competent person.
	4.3	I go above and beyond what is required.	I make every effort to do more than what is expected of me.
Order	5.4	I am a very competent person.	I am effective at my work.
	10.3	I frequently forget to put things back in their proper place.	I frequently do not put things in their proper place.
	5.6	I frequently forget to put things back in their proper place.	I hardly ever lose or misplace things.
	8.0	I prefer keeping my options open and rarely plan in advance.	I seldom make detailed "to do" lists.
	5.7	I prefer keeping my options open and rarely plan in advance.	I dislike doing things without proper planning.
	5.2	As long as I have a little bit of clear space on my desk, I am happy to do my work.	I feel comfortable even in very disorganized settings.
	4.5	I have a daily routine and stick to it.	I dislike doing things without proper planning.
	3.4	I have a daily routine and stick to it.	I prefer to do things in a logical order.
	4.4	I need a neat environment in order to work well.	I become annoyed when things around me are disorganized.
	3.0	I dislike doing things without proper planning.	I prefer to do things in a logical order.
Responsibility	4.7	I become annoyed when things around me are disorganized.	I hate when people are sloppy.
	4.4	I only think of myself and rarely of others.	There is no use in doing things for people; they never really appreciate your efforts.
	6.8	I am not the most responsible person among my friends.	I am usually not the most responsible group member, but I will not shirk on my duties either.
	15.4	Although I think charitable causes are important, I rarely become involved in charitable activities.	I would gladly spend some of my leisure time trying to improve my community.
	3.9	Although I think charitable causes are important, I rarely become involved in charitable activities.	When others are counting on me, I am extra careful to show up on time.

Table 10, continued

Scale	$\chi^2/\text{df}$ Ratio	Item Pair	
	3.9	I am not the most responsible group member, but I will not shirk on my duties either.	When others depend on me, I try to put in an extra effort.
	4.5	I have a reputation for being late for almost every meeting or event.	When others are counting on me, I am extra careful to show up on time.
	4.5	I often feel responsible for making sure that all group project assignments are completed.	When working on a group project, my group members often rely on me to get things done.
	6.7	When I make mistakes I often blame others.	I hold myself accountable for most of my mistakes.
	6.9	When others are counting on me, I am extra careful to show up on time.	When others depend on me, I try to put in an extra effort.

Table 11.

Correlations of Perceived Relevance and Specificity with 2PL Discrimination Estimates

<b>Frame of Reference</b>	<b>Scale</b>	<b>Perceived Relevance</b>	<b>Specificity</b>
Work	Industriousness	0.45*	-0.50*
	Order	0.10	0.06
	Responsibility	0.54*	-0.14
	Overall	0.33*	-0.18
School	Industriousness	0.09	-0.25
	Order	-0.15	-0.09
	Responsibility	0.51*	-0.13
	Overall	0.21	-0.17

*Note.* Asterisk (\*) denotes a correlation that is significant at  $p < .05$  (non-directional).

Table 12.  
Correlation Matrix: Performance Criteria and All Personality Scales

	Mean	SD	N	IRB	OCBI	GPA	Orig. Ind (W)	Orig. Ord (W)	Orig. Res (W)	Work Rewrit. Ind	Work Rewrit. Ord	Work Rewrit. Res
IRB	4.64	0.41	362	0.81								
OCBI	3.96	0.66	362	0.47	0.83							
GPA	0.85	0.16	271	0.08	-0.04	.						
Original Ind (Work)	4.02	0.56	475	0.52	0.42	0.12	<b>0.92</b>					
Original Ord (Work)	3.67	0.63	475	0.33	0.21	0.07	0.53	<b>0.91</b>				
Original Res (Work)	4.07	0.48	475	0.50	0.48	0.13	0.78	0.47	<b>0.86</b>			
Work Rewritten Ind	3.93	0.60	475	0.51	0.46	0.09	0.86	0.56	0.76	<b>0.91</b>		
Work Rewritten Ord	3.76	0.57	475	0.38	0.30	0.08	0.59	0.86	0.56	0.62	<b>0.86</b>	
Work Rewritten Res	4.14	0.50	475	0.47	0.52	0.06	0.72	0.39	0.80	0.75	0.50	<b>0.86</b>
Work-Relevant Ind*	3.99	0.56	475	0.52	0.45	0.10	0.92	0.56	0.79	0.98	0.63	0.78
Work-Relevant Ord	3.76	0.58	475	0.37	0.29	0.08	0.61	0.91	0.56	0.64	0.98	0.50
Work-Relevant Res	4.13	0.48	475	0.51	0.49	0.07	0.77	0.44	0.89	0.80	0.55	0.96
Original Ind (School)	3.92	0.59	394	0.41	0.27	0.22	0.72	0.49	0.58	0.62	0.53	0.47
Original Ord (School)	3.60	0.71	394	0.28	0.14	0.07	0.45	0.87	0.38	0.47	0.75	0.26
Original Res (School)	4.02	0.48	394	0.45	0.40	0.19	0.66	0.43	0.79	0.62	0.52	0.63
School Rewritten Ind	3.81	0.64	394	0.39	0.22	0.22	0.68	0.47	0.56	0.60	0.50	0.43
School Rewritten Ord	3.64	0.66	394	0.25	0.13	0.07	0.45	0.81	0.36	0.46	0.74	0.26
School Rewritten Res	3.94	0.48	394	0.38	0.36	0.22	0.65	0.42	0.70	0.61	0.49	0.63
School Relevant Ind	3.91	0.57	394	0.40	0.22	0.24	0.71	0.48	0.59	0.61	0.53	0.46
School Relevant Ord	3.63	0.66	394	0.26	0.14	0.07	0.46	0.84	0.37	0.48	0.76	0.28
School Relevant Res	3.92	0.47	394	0.41	0.37	0.23	0.68	0.46	0.74	0.64	0.54	0.63



Table 12, continued

	Work Relev. Ind*	Work Relev. Ord	Work Relev. Res	Orig. Ind (S)	Orig. Ord (S)	Orig. Res (S)	School Rewrit. Ind	School Rewrit. Ord	School Rewrit. Res	School Relev. Ind	School Relev. Ord	School Relev. Res
<b>Work Relevant Ind*</b>	<b>0.91</b>											
<b>Work Relevant Ord</b>	0.65	<b>0.89</b>										
<b>Work Relevant Res</b>	0.82	0.55	<b>0.88</b>									
<b>Original Ind (School)</b>	0.66	0.55	0.52	<b>0.92</b>								
<b>Original Ord (School)</b>	0.48	0.80	0.31	0.53	<b>0.92</b>							
<b>Original Res (School)</b>	0.65	0.53	0.69	0.72	0.44	<b>0.86</b>						
<b>School Rewritten Ind</b>	0.63	0.52	0.50	0.84	0.52	0.62	<b>0.88</b>					
<b>School Rewritten Ord</b>	0.47	0.78	0.31	0.53	0.90	0.40	0.56	<b>0.86</b>				
<b>School Rewritten Res</b>	0.65	0.50	0.67	0.68	0.42	0.75	0.69	0.44	<b>0.81</b>			
<b>School Relevant Ind</b>	0.65	0.55	0.52	0.91	0.53	0.67	0.97	0.56	0.72	<b>0.90</b>		
<b>School Relevant Ord</b>	0.49	0.80	0.33	0.54	0.94	0.43	0.56	0.98	0.45	0.57	<b>0.90</b>	
<b>School Relevant Res</b>	0.68	0.55	0.70	0.75	0.48	0.83	0.73	0.48	0.97	0.76	0.50	<b>0.84</b>

*Notes.* Values in bold on the diagonal are coefficient alphas. IRB: in-role behavior; OCBI: organizational citizenship behavior – individual. In order to place all self-reported GPAs on the same scale, GPAs were transformed to a percentage scale (e.g., both a 3.75 GPA on a 4-point scale and a 4.69 GPA on a 5-point scale would be transformed to .94). N = 475 for personality scales; N = 362 for correlations between all personality scales and both job performance measures; N = 271 for correlation between all personality scales and GPA. Number of items in each Work-Rewritten scale: Industriousness, 13; Order, 16; Responsibility, 14. Number of items in each School-Rewritten scale: Industriousness, 13; Order, 15; Responsibility, 16. There are 20 items in each of the remaining personality scales.

\*Due to an error in administering the Work-Relevant Industriousness scale, this scale contains only 19 items.

Table 13.  
Convergent/Discriminant Validity Analysis, Work-Rewritten Industriousness

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Orig.	1	I lower my standards as tasks get more difficult.	<b>0.40</b>	0.17
Ind.	2	I often need a push to get started.	<b>0.32</b>	0.25
	3	I do not work as hard as the majority of people around me.	<b>0.47</b>	0.19
	4	Setting goals and achieving them is not very important to me.	<b>0.51</b>	0.00
	5	Getting average grades is enough for me.	<b>0.30</b>	0.03
	6	I exert medium effort on most tasks I do.	<b>0.52</b>	0.11
	7	I do what is required, but rarely anything more.	<b>0.51</b>	0.20
	8	I can never guarantee completion of even the simplest projects.	<b>0.27</b>	0.17
	9	I have high standards and work toward them.	<b>0.69</b>	0.04
	10	I handle even difficult tasks smoothly.	<b>0.64</b>	-0.01
	11	I go above and beyond what is required.	<b>0.78</b>	-0.06
	12	I strive to improve myself.	<b>0.82</b>	-0.22
	13	I am a very competent person.	<b>0.64</b>	-0.05
	14	I make every effort to do more than what is expected of me.	<b>0.75</b>	-0.07
	15	I demand the highest quality in everything I do.	<b>0.81</b>	-0.07
	16	I am effective at my work.	<b>0.62</b>	0.03
	17	I set extremely high standards for myself.	<b>0.76</b>	-0.08
	18	I try to be the best at anything I do.	<b>0.68</b>	-0.06
	19	I often contribute little to group projects.	<b>0.33</b>	0.11
	20	I invest little effort in my work.	<b>0.46</b>	0.13
Rewrit.	2	It is often hard for me to get started on my work unless a deadline is approaching or I am being monitored.	<b>0.31</b>	0.21
Ind.	3	I do not work as hard as most people do at their jobs.	<b>0.40</b>	0.26
	4	Setting goals in order to deliver a good work product is not very important to me.	<b>0.34</b>	0.24
	7	I only do the minimum work that my boss or supervisor requires of me.	<b>0.48</b>	0.18
	10	I handle even difficult job requirements smoothly.	<b>0.56</b>	0.02
	11	I go above and beyond the minimum requirements of my job.	<b>0.70</b>	-0.05
	12	I strive to become an even better employee.	<b>0.77</b>	-0.13
	13	I am a very competent employee.	<b>0.64</b>	-0.04
	14	If I finish a required work task early, I will try to make it even better or look for something else to do instead of taking a break.	<b>0.54</b>	0.05
	15	I demand the highest quality from myself when performing job tasks.	<b>0.80</b>	-0.04
	16	I complete all of my required work tasks in a thorough and timely manner.	<b>0.65</b>	0.01
	17	The standards I set for myself are as high or higher than those expected of me by my employer.	<b>0.75</b>	-0.02
	18	I strive to do the best job I can on all work tasks.	<b>0.78</b>	-0.07

Table 13, continued

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Rewrit. Ord.	1	I frequently forget to return materials where they belong after I use them.	0.07	<b>0.58</b>
	2	I do what seems important to me at the moment instead of planning out what I need to accomplish.	0.08	<b>0.45</b>
	3	Even when it would be helpful, I seldom make a "to do" list of the work I need to complete.	0.13	<b>0.31</b>
	5	For me, it is not important to keep things organized while I work.	0.07	<b>0.56</b>
	6	I am not a very neat and tidy employee.	0.04	<b>0.63</b>
	7	I frequently do not put work materials back where they belong.	0.08	<b>0.53</b>
	8	I don't mind working in a messy environment as long as there is a little bit of clear space to work.	-0.09	<b>0.67</b>
	9	I have established a daily routine to help me perform well on the job.	0.27	<b>0.35</b>
	11	I do not like to work on a job task before making a good plan.	0.18	<b>0.20</b>
	12	I prefer to plan a logical order for completing work assignments before I begin working.	<b>0.30</b>	0.20
	14	I almost always know where I can find the materials that I put away for later use.	<b>0.32</b>	0.28
	15	I become annoyed when my workspace becomes disorganized.	0.06	<b>0.55</b>
	17	I hate when I see coworkers doing sloppy work.	<b>0.35</b>	0.13
	18	I make sure that everything in my workspace has its own designated place.	0.20	<b>0.47</b>
	19	Most of the time my workspace is in complete disarray.	0.03	<b>0.65</b>
	20	I feel comfortable on the job even if the workplace is very disorganized.	-0.21	<b>0.66</b>

*Note.* The larger factor loading for each item is indicated in bold. The structure matrix loadings reveal the exact same pattern of results. Item numbers for the Rewritten scale items correspond to the Original scale counterparts.

Table 14.  
Convergent/Discriminant Validity Analysis, Work-Rewritten Order

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Orig.	1	I frequently forget to put things back in their proper place.	<b>0.47</b>	0.16
Ord.	2	I prefer keeping my options open and rarely plan in advance.	<b>0.39</b>	0.17
	3	I seldom make detailed "to do" lists.	<b>0.42</b>	0.07
	4	I do not like work spaces that are too clean and tidy.	<b>0.50</b>	0.19
	5	For me, being organized is unimportant.	<b>0.68</b>	0.14
	6	Being neat is not exactly my strength.	<b>0.73</b>	-0.02
	7	I frequently do not put things in their proper place.	<b>0.58</b>	0.17
	8	As long as I have a little bit of clear space on my desk, I am happy to do my work.	<b>0.64</b>	-0.15
	9	I have a daily routine and stick to it.	<b>0.38</b>	0.09
	10	I need a neat environment in order to work well.	<b>0.80</b>	-0.18
	11	I dislike doing things without proper planning.	<b>0.53</b>	-0.09
	12	I prefer to do things in a logical order.	<b>0.33</b>	0.21
	13	Organization is a key component of most things I do.	<b>0.72</b>	0.07
	14	I hardly ever lose or misplace things.	<b>0.48</b>	-0.02
	15	I become annoyed when things around me are disorganized.	<b>0.67</b>	-0.05
	16	I keep detailed notes of important meetings and lectures.	<b>0.27</b>	0.17
	17	I hate when people are sloppy.	<b>0.51</b>	0.03
	18	Every item in my room and on my desk has its own designated place.	<b>0.70</b>	-0.07
	19	Most of the time my belongings are in complete disarray.	<b>0.70</b>	0.01
	20	I feel comfortable even in very disorganized settings.	<b>0.74</b>	-0.28
Rewrit. Order	1	I frequently forget to return materials where they belong after I use them.	<b>0.43</b>	0.26
	2	I do what seems important to me at the moment instead of planning out what I need to accomplish.	<b>0.32</b>	0.19
	3	Even when it would be helpful, I seldom make a "to do" list of the work I need to complete.	<b>0.30</b>	0.17
	5	For me, it is not important to keep things organized while I work.	<b>0.50</b>	0.20
	6	I am not a very neat and tidy employee.	<b>0.49</b>	0.25
	7	I frequently do not put work materials back where they belong.	<b>0.40</b>	0.27
	8	I don't mind working in a messy environment as long as there is a little bit of clear space to work.	<b>0.72</b>	-0.15
	9	I have established a daily routine to help me perform well on the job.	<b>0.39</b>	0.23
	11	I do not like to work on a job task before making a good plan.	<b>0.41</b>	-0.04
	12	I prefer to plan a logical order for completing work assignments before I begin working.	<b>0.38</b>	0.16
	14	I almost always know where I can find the materials that I put away for later use.	0.30	<b>0.37</b>
	15	I become annoyed when my workspace becomes disorganized.	<b>0.66</b>	-0.02
	17	I hate when I see coworkers doing sloppy work.	0.24	<b>0.28</b>
	18	I make sure that everything in my workspace has its own designated place.	<b>0.63</b>	0.10
	19	Most of the time my workspace is in complete disarray.	<b>0.55</b>	0.16
	20	I feel comfortable on the job even if the workplace is very disorganized.	<b>0.73</b>	-0.30

Table 14, continued

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Rewrit. Res.	1	I don't take my coworkers into consideration when I make decisions.	-0.02	<b>0.51</b>
	2	I am not usually the person who can be counted on to get a job done.	0.06	<b>0.44</b>
	3	Sometimes it is too much of a bother to follow my supervisor's directions exactly.	0.13	<b>0.46</b>
	4	If I am going to be late for work, I might decide not to go at all.	0.06	<b>0.49</b>
	5	Although I understand that coworkers sometimes need assistance with a task, I see no reason for helping them more than is required.	-0.02	<b>0.57</b>
	6	I can be depended on to do the work that I am assigned, but nothing more.	0.07	<b>0.57</b>
	8	I can be counted on to do my assigned work to the best of my ability.	0.10	<b>0.59</b>
	10	My coworkers know they can count on me to help out when they need my assistance.	0.00	<b>0.67</b>
	11	When I make a mistake, I will often try to blame the problem on someone else's actions.	-0.03	<b>0.39</b>
	14	I would gladly take on extra hours at work to help my employer with a special project or event.	-0.09	<b>0.54</b>
	16	When necessary, I will put in extra time or effort to complete the work I said I would do.	0.03	<b>0.61</b>
	18	When my contributions will help to ensure the success of everyone on a work team, I will try to put in an extra effort.	-0.02	<b>0.58</b>
	19	There is no use in helping out coworkers, because no one will praise me for my efforts.	-0.06	<b>0.59</b>
	20	Although I think there are ways that an employee can add value to a company through unique contributions, I rarely do so.	0.08	<b>0.47</b>

*Note.* The larger factor loading for each item is indicated in bold. The structure matrix loadings reveal the exact same pattern of results. Item numbers for the Rewritten scale items correspond to the Original scale counterparts.

Table 15.  
Convergent/Discriminant Validity Analysis, Work-Rewritten Responsibility

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Orig.	1	I only think of myself and rarely of others.	<b>0.52</b>	-0.03
Res.	2	I am not the most responsible person among my friends.	0.13	<b>0.31</b>
	3	Sometimes it is too much of a bother to do exactly what is promised.	<b>0.34</b>	0.28
	4	If I am running late to an appointment, I may decide not to go at all.	<b>0.34</b>	0.16
	5	Although I think charitable causes are important, I rarely become involved in charitable activities.	0.14	<b>0.16</b>
	6	I am usually not the most responsible group member, but I will not shirk on my duties either.	0.08	<b>0.45</b>
	7	I have a reputation for being late for almost every meeting or event.	0.21	<b>0.22</b>
	8	I carry out my obligations to the best of my ability.	<b>0.73</b>	0.01
	9	I often feel responsible for making sure that all group project assignments are completed.	<b>0.57</b>	0.12
	10	My friends know they can count on me in times of need.	<b>0.70</b>	-0.10
	11	When I make mistakes I often blame others.	<b>0.36</b>	0.08
	12	I hold myself accountable for most of my mistakes.	<b>0.50</b>	-0.07
	13	When working on a group project, my group members often rely on me to get things done.	<b>0.52</b>	0.14
	14	I would gladly spend some of my leisure time trying to improve my community.	<b>0.35</b>	0.01
	15	If I am running late, I try to call ahead to notify those who are waiting for me.	<b>0.59</b>	-0.08
	16	I go out of my way to keep my promises.	<b>0.67</b>	-0.03
	17	When others are counting on me, I am extra careful to show up on time.	<b>0.65</b>	-0.05
	18	When others depend on me, I try to put in an extra effort.	<b>0.75</b>	-0.14
	19	There is no use in doing things for people; they never really appreciate your efforts.	<b>0.48</b>	0.04
	20	I see no reason for belonging to clubs or community organizations.	<b>0.31</b>	0.16
Rewrit.	1	I don't take my coworkers into consideration when I make decisions.	<b>0.41</b>	0.08
Res.	2	I am not usually the person who can be counted on to get a job done.	<b>0.38</b>	0.12
	3	Sometimes it is too much of a bother to follow my supervisor's directions exactly.	<b>0.31</b>	0.24
	4	If I am going to be late for work, I might decide not to go at all.	<b>0.43</b>	0.12
	5	Although I understand that coworkers sometimes need assistance with a task, I see no reason for helping them more than is required.	<b>0.48</b>	0.05
	6	I can be depended on to do the work that I am assigned, but nothing more.	<b>0.40</b>	0.23
	8	I can be counted on to do my assigned work to the best of my ability.	<b>0.68</b>	0.03

Table 15, continued

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Rewrit. Res. (cont.)	10	My coworkers know they can count on me to help out when they need my assistance.	<b>0.73</b>	-0.04
	11	When I make a mistake, I will often try to blame the problem on someone else's actions.	<b>0.34</b>	0.04
	14	I would gladly take on extra hours at work to help my employer with a special project or event.	<b>0.53</b>	-0.08
	16	When necessary, I will put in extra time or effort to complete the work I said I would do.	<b>0.68</b>	-0.06
	18	When my contributions will help to ensure the success of everyone on a work team, I will try to put in an extra effort.	<b>0.72</b>	-0.14
	19	There is no use in helping out coworkers, because no one will praise me for my efforts.	<b>0.48</b>	0.05
	20	Although I think there are ways that an employee can add value to a company through unique contributions, I rarely do so.	<b>0.30</b>	0.21
Rewrit. Ord	1	I frequently forget to return materials where they belong after I use them.	0.06	<b>0.57</b>
	2	I do what seems important to me at the moment instead of planning out what I need to accomplish.	0.03	<b>0.47</b>
	3	Even when it would be helpful, I seldom make a "to do" list of the work I need to complete.	0.09	<b>0.34</b>
	5	For me, it is not important to keep things organized while I work.	0.01	<b>0.62</b>
	6	I am not a very neat and tidy employee.	0.06	<b>0.62</b>
	7	I frequently do not put work materials back where they belong.	0.08	<b>0.54</b>
	8	I don't mind working in a messy environment as long as there is a little bit of clear space to work.	-0.15	<b>0.68</b>
	9	I have established a daily routine to help me perform well on the job.	0.20	<b>0.40</b>
	11	I do not like to work on a job task before making a good plan.	0.02	<b>0.33</b>
	12	I prefer to plan a logical order for completing work assignments before I begin working.	0.20	<b>0.29</b>
	14	I almost always know where I can find the materials that I put away for later use.	<b>0.33</b>	0.29
	15	I become annoyed when my workspace becomes disorganized.	-0.01	<b>0.60</b>
	17	I hate when I see coworkers doing sloppy work.	<b>0.31</b>	0.20
	18	I make sure that everything in my workspace has its own designated place.	0.11	<b>0.57</b>
	19	Most of the time my workspace is in complete disarray.	0.04	<b>0.63</b>
	20	I feel comfortable on the job even if the workplace is very disorganized.	-0.28	<b>0.68</b>

*Note.* The larger factor loading for each item is indicated in bold. The structure matrix loadings reveal the exact same pattern of results. Item numbers for the Rewritten scale items correspond to the Original scale counterparts.

Table 16.  
Convergent/Discriminant Validity Analysis, School-Rewritten Industriousness

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Orig.	1	I lower my standards as tasks get more difficult.	<b>0.38</b>	0.18
Ind.	2	I often need a push to get started.	0.23	<b>0.36</b>
	3	I do not work as hard as the majority of people around me.	<b>0.45</b>	0.32
	4	Setting goals and achieving them is not very important to me.	<b>0.32</b>	0.29
	5	Getting average grades is enough for me.	<b>0.33</b>	0.11
	6	I exert medium effort on most tasks I do.	<b>0.44</b>	0.29
	7	I do what is required, but rarely anything more.	<b>0.50</b>	0.21
	8	I can never guarantee completion of even the simplest projects.	<b>0.31</b>	0.14
	9	I have high standards and work toward them.	<b>0.74</b>	0.00
	10	I handle even difficult tasks smoothly.	<b>0.61</b>	-0.12
	11	I go above and beyond what is required.	<b>0.73</b>	0.09
	12	I strive to improve myself.	<b>0.70</b>	-0.13
	13	I am a very competent person.	<b>0.64</b>	-0.23
	14	I make every effort to do more than what is expected of me.	<b>0.75</b>	0.03
	15	I demand the highest quality in everything I do.	<b>0.76</b>	0.06
	16	I am effective at my work.	<b>0.63</b>	0.00
	17	I set extremely high standards for myself.	<b>0.80</b>	-0.08
	18	I try to be the best at anything I do.	<b>0.68</b>	0.02
	19	I often contribute little to group projects.	<b>0.32</b>	0.04
	20	I invest little effort in my work.	<b>0.40</b>	0.28
Rewrit.	2	It is often hard for me to get started on my coursework unless someone is pushing me to do it or it is due soon.	0.22	<b>0.45</b>
Ind.	3	I do not work as hard as the majority of college students.	<b>0.36</b>	0.33
	4	Setting goals in order to be a successful student is not very important to me.	0.23	<b>0.29</b>
	7	I only do the minimum work required by my instructors.	<b>0.37</b>	0.35
	10	I competently handle even the most challenging class projects.	<b>0.60</b>	-0.08
	11	My performance exceeds the University's minimum requirements for a degree.	<b>0.37</b>	-0.06
	12	I strive to become an even better student.	<b>0.61</b>	0.00
	13	I am a very competent student.	<b>0.65</b>	-0.12
	14	I work hard to do more than my teachers expect of me.	<b>0.65</b>	0.13
	15	I demand the highest quality from myself on class assignments.	<b>0.72</b>	0.17
	16	I complete all of my coursework in a thorough and timely manner.	<b>0.50</b>	0.16
	17	The standards I set for myself are as high or higher than those expected of me by my teachers and/or family.	<b>0.69</b>	0.05
	18	I try to be the best student that I can be.	<b>0.62</b>	0.12



Table 16, continued

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Rewrit. Ord.	1	I frequently forget to put my class notes and handouts into the correct folder or file.	-0.11	<b>0.60</b>
	2	I prefer keeping my options open and rarely plan in advance how I will complete assignments with upcoming deadlines.	0.16	<b>0.50</b>
	3	I seldom make detailed "to do" lists to help me complete my schoolwork.	-0.01	<b>0.40</b>
	5	For me, it is not important to keep my schoolwork organized.	0.04	<b>0.40</b>
	6	I do not maintain a very neat and tidy environment.	-0.04	<b>0.71</b>
	7	When I am done with my schoolwork, I frequently leave a mess of books, notes, and other materials lying around.	-0.11	<b>0.73</b>
	9	I have a daily routine for completing schoolwork that I stick to.	0.16	<b>0.42</b>
	11	I dislike studying or completing an assignment without having a good plan.	0.22	<b>0.27</b>
	12	I prefer to plan a logical order for completing my coursework before I begin working.	0.27	<b>0.33</b>
	14	I hardly ever lose or misplace class notes and assignments.	0.10	<b>0.44</b>
	15	I become annoyed when my notes and handouts are disorganized.	0.12	<b>0.42</b>
	17	I hate when I see my peers turn in sloppy work.	0.17	<b>0.20</b>
	18	I make sure that everything in my room or office has its own designated place.	0.01	<b>0.60</b>
	19	Most of the time my room or office is in complete disarray.	-0.13	<b>0.75</b>
	20	I feel comfortable doing my schoolwork even in very disorganized settings.	-0.15	<b>0.67</b>

*Note.* The larger factor loading for each item is indicated in bold. The structure matrix loadings reveal the exact same pattern of results. Item numbers for the Rewritten scale items correspond to the Original scale counterparts.

Table 17.  
Convergent/Discriminant Validity Analysis, School-Rewritten Order

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Orig.	1	I frequently forget to put things back in their proper place.	<b>0.70</b>	-0.12
Ord.	2	I prefer keeping my options open and rarely plan in advance.	<b>0.53</b>	0.01
	3	I seldom make detailed "to do" lists.	<b>0.46</b>	0.05
	4	I do not like work spaces that are too clean and tidy.	<b>0.54</b>	-0.04
	5	For me, being organized is unimportant.	<b>0.79</b>	-0.06
	6	Being neat is not exactly my strength.	<b>0.81</b>	-0.13
	7	I frequently do not put things in their proper place.	<b>0.78</b>	-0.15
	8	As long as I have a little bit of clear space on my desk, I am happy to do my work.	<b>0.63</b>	-0.09
	9	I have a daily routine and stick to it.	<b>0.49</b>	0.18
	10	I need a neat environment in order to work well.	<b>0.70</b>	0.02
	11	I dislike doing things without proper planning.	<b>0.46</b>	0.17
	12	I prefer to do things in a logical order.	<b>0.34</b>	0.28
	13	Organization is a key component of most things I do.	<b>0.77</b>	0.10
	14	I hardly ever lose or misplace things.	<b>0.56</b>	-0.01
	15	I become annoyed when things around me are disorganized.	<b>0.60</b>	0.06
	16	I keep detailed notes of important meetings and lectures.	<b>0.37</b>	0.31
	17	I hate when people are sloppy.	<b>0.46</b>	0.04
	18	Every item in my room and on my desk has its own designated place.	<b>0.70</b>	-0.09
	19	Most of the time my belongings are in complete disarray.	<b>0.79</b>	-0.12
	20	I feel comfortable even in very disorganized settings.	<b>0.67</b>	-0.16
Rewrit. Order	1	I frequently forget to put my class notes and handouts into the correct folder or file.	<b>0.52</b>	-0.05
	2	I prefer keeping my options open and rarely plan in advance how I will complete assignments with upcoming deadlines.	<b>0.47</b>	0.13
	3	I seldom make detailed "to do" lists to help me complete my schoolwork.	<b>0.35</b>	0.07
	5	For me, it is not important to keep my schoolwork organized.	<b>0.36</b>	0.12
	6	I do not maintain a very neat and tidy environment.	<b>0.73</b>	-0.06
	7	When I am done with my schoolwork, I frequently leave a mess of books, notes, and other materials lying around.	<b>0.70</b>	-0.11
	9	I have a daily routine for completing schoolwork that I stick to.	<b>0.43</b>	0.12
	11	I dislike studying or completing an assignment without having a good plan.	<b>0.37</b>	0.13
	12	I prefer to plan a logical order for completing my coursework before I begin working.	<b>0.41</b>	0.20
	14	I hardly ever lose or misplace class notes and assignments.	<b>0.47</b>	0.18
	15	I become annoyed when my notes and handouts are disorganized.	<b>0.54</b>	0.10
	17	I hate when I see my peers turn in sloppy work.	<b>0.30</b>	0.02
	18	I make sure that everything in my room or office has its own designated place.	<b>0.70</b>	-0.01
	19	Most of the time my room or office is in complete disarray.	<b>0.75</b>	-0.12
	20	I feel comfortable doing my schoolwork even in very disorganized settings.	<b>0.73</b>	-0.19

Table 17, continued

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Rewrit. Res.	1	I don't take my group members' needs into consideration when working in a group.	0.00	<b>0.42</b>
	2	I am not the most responsible person when it comes to completing my coursework.	0.17	<b>0.28</b>
	3	Sometimes it is too much of a bother to follow the instructions for an assignment exactly.	<b>0.23</b>	0.22
	4	If I am going to be late to class, I might decide not to go at all.	<b>0.28</b>	0.15
	5	Although I understand that my group members sometimes need help in order to finish an assignment, I see no reason for being the one to do so.	-0.05	<b>0.48</b>
	7	I have a reputation for being late to classes and meetings.	<b>0.23</b>	0.19
	8	I complete all course requirements to the best of my ability.	0.22	<b>0.47</b>
	10	My classmates know that they can count on me when they need my help to complete a group assignment.	0.02	<b>0.66</b>
	11	When I do poorly on an exam or assignment I often try to blame it on someone or something outside of my control.	-0.01	<b>0.23</b>
	12	I hold myself accountable if I fail to deliver what I promised to my instructor.	0.00	<b>0.48</b>
	14	I would gladly spend some of my leisure time helping an instructor to improve a course.	0.08	<b>0.35</b>
	16	I will go out of my way to do what I promised to my group.	-0.06	<b>0.73</b>
	17	When my classmates are counting on me, I am extra careful to show up on time.	-0.02	<b>0.70</b>
	18	When others are depending on me to help complete an assignment, I will try to put in an extra effort.	-0.06	<b>0.79</b>
	19	There is no use for helping out other group members, because it will not benefit me.	-0.09	<b>0.58</b>
	20	Although I think that contributing to class discussions makes classtime more valuable to everyone, I rarely do so.	<b>0.09</b>	0.07

*Note.* The larger factor loading for each item is indicated in bold. The structure matrix loadings reveal the exact same pattern of results. Item numbers for the Rewritten scale items correspond to the Original scale counterparts.

Table 18.  
Convergent/Discriminant Validity Analysis, School-Rewritten Responsibility

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Orig.	1	I only think of myself and rarely of others.	<b>0.44</b>	-0.08
Res.	2	I am not the most responsible person among my friends.	0.17	<b>0.34</b>
	3	Sometimes it is too much of a bother to do exactly what is promised.	0.30	<b>0.33</b>
	4	If I am running late to an appointment, I may decide not to go at all.	0.26	<b>0.30</b>
	5	Although I think charitable causes are important, I rarely become involved in charitable activities.	<b>0.19</b>	0.14
	6	I am usually not the most responsible group member, but I will not shirk on my duties either.	0.14	<b>0.37</b>
	7	I have a reputation for being late for almost every meeting or event.	0.19	<b>0.19</b>
	8	I carry out my obligations to the best of my ability.	<b>0.53</b>	0.17
	9	I often feel responsible for making sure that all group project assignments are completed.	<b>0.50</b>	0.26
	10	My friends know they can count on me in times of need.	<b>0.63</b>	-0.06
	11	When I make mistakes I often blame others.	<b>0.28</b>	-0.05
	12	I hold myself accountable for most of my mistakes.	<b>0.54</b>	-0.09
	13	When working on a group project, my group members often rely on me to get things done.	<b>0.50</b>	0.18
	14	I would gladly spend some of my leisure time trying to improve my community.	<b>0.42</b>	-0.04
	15	If I am running late, I try to call ahead to notify those who are waiting for me.	<b>0.44</b>	0.08
	16	I go out of my way to keep my promises.	<b>0.63</b>	0.04
	17	When others are counting on me, I am extra careful to show up on time.	<b>0.64</b>	0.03
	18	When others depend on me, I try to put in an extra effort.	<b>0.72</b>	-0.05
	19	There is no use in doing things for people; they never really appreciate your efforts.	<b>0.41</b>	-0.10
	20	I see no reason for belonging to clubs or community organizations.	<b>0.32</b>	0.15
Rewrit.	1	I don't take my group members' needs into consideration when working in a group.	<b>0.41</b>	0.04
Res.	2	I am not the most responsible person when it comes to completing my coursework.	0.17	<b>0.33</b>
	3	Sometimes it is too much of a bother to follow the instructions for an assignment exactly.	0.15	<b>0.36</b>
	4	If I am going to be late to class, I might decide not to go at all.	0.08	<b>0.40</b>
	5	Although I think that contributing to class discussions makes classtime more valuable to everyone, I rarely do so.	0.03	<b>0.15</b>
	7	I have a reputation for being late to classes and meetings.	0.15	<b>0.30</b>
	8	I complete all course requirements to the best of my ability.	<b>0.39</b>	0.34
	10	My classmates know that they can count on me when they need my help to complete a group assignment.	<b>0.64</b>	0.09
	11	When I do poorly on an exam or assignment I often try to blame it on someone or something outside of my control.	<b>0.23</b>	0.02

Table 18, continued

Scale	Item #		Pattern Matrix	
			Factor 1	Factor 2
Rewrit. Res. (cont.)	12	I hold myself accountable if I fail to deliver what I promised to my instructor.	<b>0.51</b>	-0.03
	14	I would gladly spend some of my leisure time helping an instructor to improve a course.	<b>0.33</b>	0.10
	16	I will go out of my way to do what I promised to my group.	<b>0.71</b>	-0.01
	17	When my classmates are counting on me, I am extra careful to show up on time.	<b>0.65</b>	0.02
	18	When others are depending on me to help complete an assignment, I will try to put in an extra effort.	<b>0.73</b>	-0.01
	19	There is no use for helping out other group members, because it will not benefit me.	<b>0.54</b>	-0.03
	20	Although I understand that my group members sometimes need help in order to finish an assignment, I see no reason for being the one to do so.	<b>0.47</b>	0.02
Rewrit. Order	1	I frequently forget to put my class notes and handouts into the correct folder or file.	-0.14	<b>0.58</b>
	2	I prefer keeping my options open and rarely plan in advance how I will complete assignments with upcoming deadlines.	0.01	<b>0.58</b>
	3	I seldom make detailed "to do" lists to help me complete my schoolwork.	0.00	<b>0.41</b>
	5	For me, it is not important to keep my schoolwork organized.	0.04	<b>0.42</b>
	6	I do not maintain a very neat and tidy environment.	-0.11	<b>0.75</b>
	7	When I am done with my schoolwork, I frequently leave a mess of books, notes, and other materials lying around.	-0.18	<b>0.73</b>
	9	I have a daily routine for completing schoolwork that I stick to.	0.00	<b>0.49</b>
	11	I dislike studying or completing an assignment without having a good plan.	0.03	<b>0.38</b>
	12	I prefer to plan a logical order for completing my coursework before I begin working.	0.09	<b>0.45</b>
	14	I hardly ever lose or misplace class notes and assignments.	0.07	<b>0.52</b>
	15	I become annoyed when my notes and handouts are disorganized.	0.04	<b>0.49</b>
	17	I hate when I see my peers turn in sloppy work.	0.01	<b>0.28</b>
	18	I make sure that everything in my room or office has its own designated place.	-0.07	<b>0.67</b>
	19	Most of the time my room or office is in complete disarray.	-0.14	<b>0.74</b>
	20	I feel comfortable doing my schoolwork even in very disorganized settings.	-0.20	<b>0.65</b>

*Note.* The larger factor loading for each item is indicated in bold. The structure matrix loadings reveal the exact same pattern of results. Item numbers for the Rewritten scale items correspond to the Original scale counterparts.

Table 19.  
Rewritten Item Estimates, Work FOR Industriousness

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
2	I often need a push to get started.	More	More	It is often hard for me to get started on my work unless a deadline is approaching or I am being monitored.	0.94	-0.32	-0.21
3	I do not work as hard as the majority of people around me.	More	More	I do not work as hard as most people do at their jobs.	1.69	-1.54	0.16
4	Setting goals and achieving them is not very important to me.	More	Less	Setting goals in order to deliver a good work product is not very important to me.	1.10	-1.42	-0.42
7	I do what is required, but rarely anything more.	More	More	I only do the minimum work that my boss or supervisor requires of me.	1.47	-1.35	-0.24
10	I handle even difficult tasks smoothly.	More	Less	I handle even difficult job requirements smoothly.	1.27	-0.75	-0.04
11	I go above and beyond what is required.	More	More	I go above and beyond the minimum requirements of my job.	1.59	-0.62	-1.10
12	I strive to improve myself.	More	Less	I strive to become an even better employee.	1.76	-1.24	-0.35
13	I am a very competent person.	More	Less	I am a very competent employee.	1.28	-1.76	0.03
14	I make every effort to do more than what is expected of me.	More	Less	If I finish a required work task early, I will try to make it even better or look for something else to do instead of taking a break.	1.30	-0.49	-0.86

Table 19, continued

Item #	Original Scale Item	Relevance (More/ Less)	Specificity (More/ Less)	Rewritten Scale Item	Rewritten Item Estimate		a- Parameter Difference
					a	b	
15	I demand the highest quality in everything I do.	More	Less	I demand the highest quality from myself when performing job tasks.	2.52	-1.04	-0.45
16	I am effective at my work.	More	More	I complete all of my required work tasks in a thorough and timely manner.	1.43	-1.66	-0.23
17	I set extremely high standards for myself.	More	Less	The standards I set for myself are as high or higher than those expected of me by my employer.	1.81	-0.80	-0.72
18	I try to be the best at anything I do.	More	Less	I strive to do the best job I can on all work tasks.	1.91	-1.30	0.03

Table 20.  
Rewritten Item Estimates, Work FOR Order

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
1	I frequently forget to put things back in their proper place.	Less	Less	I frequently forget to return materials where they belong after I use them.	1.29	-1.56	0.13
2	I prefer keeping my options open and rarely plan in advance.	Less	Less	I do what seems important to me at the moment instead of planning out what I need to accomplish.	0.82	-0.21	-0.19
3	I seldom make detailed "to do" lists.	More	Less	Even when it would be helpful, I seldom make a "to do" list of the work I need to complete.	0.95	-1.03	-0.12
5	For me, being organized is unimportant.	More	Less	For me, it is not important to keep things organized while I work.	1.59	-1.14	-1.33
6	Being neat is not exactly my strength.	Less	Less	I am not a very neat and tidy employee.	1.75	-1.12	-1.00
7	I frequently do not put things in their proper place.	Less	Less	I frequently do not put work materials back where they belong.	1.15	-1.98	-0.80
8	As long as I have a little bit of clear space on my desk, I am happy to do my work.	More	More	I don't mind working in a messy environment as long as there is a little bit of clear space to work.	1.61	-0.25	0.30
9	I have a daily routine and stick to it.	More	Less	I have established a daily routine to help me perform well on the job.	1.11	-0.49	0.31
11	I dislike doing things without proper planning.	More	Less	I do not like to work on a job task before making a good plan.	0.86	0.61	-0.35



Table 20, continued

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	a	b	a-Parameter Difference
12	I prefer to do things in a logical order.	More	Less	I prefer to plan a logical order for completing work assignments before I begin working.	1.00	-0.82	-0.05
14	I hardly ever lose or misplace things.	More	Less	I almost always know where I can find the materials that I put away for later use.	1.03	-1.56	0.15
15	I become annoyed when things around me are disorganized.	More	Less	I become annoyed when my workspace becomes disorganized.	1.71	-0.47	-0.11
17	I hate when people are sloppy.	Less	Less	I hate when I see coworkers doing sloppy work.	0.59	-1.59	-0.51
18	Every item in my room and on my desk has its own designated place.	More	More	I make sure that everything in my workspace has its own designated place.	2.13	-0.40	0.33
19	Most of the time my belongings are in complete disarray.	Less	More	Most of the time my workspace is in complete disarray.	1.81	-1.36	-0.02
20	I feel comfortable even in very disorganized settings.	More	Less	I feel comfortable on the job even if the workplace is very disorganized.	1.43	-0.16	-0.30

Table 21.  
Rewritten Item Estimates, Work FOR Responsibility

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
1	I only think of myself and rarely of others.	Less	Less	I don't take my coworkers into consideration when I make decisions.	1.09	-2.31	-0.16
2	I am not the most responsible person among my friends.	Less	More	I am not usually the person who can be counted on to get a job done.	1.16	-2.30	0.02
3	Sometimes it is too much of a bother to do exactly what is promised.	More	Less	Sometimes it is too much of a bother to follow my supervisor's directions exactly	0.93	-1.36	-0.43
4	If I am running late to an appointment, I may decide not to go at all.	More	More	If I am going to be late for work, I might decide not to go at all.	1.46	-2.09	0.38
5	Although I think charitable causes are important, I rarely become involved in charitable activities.	Less	More	Although I understand that coworkers sometimes need assistance with a task, I see no reason for helping them more than is required.	1.49	-1.59	0.91
6	I am usually not the most responsible group member, but I will not shirk on my duties either.	More	More	I can be depended on to do the work that I am assigned, but nothing more.	1.05	-0.35	-0.01
8	I carry out my obligations to the best of my ability.	More	Less	I can be counted on to do my assigned work to the best of my ability.	1.63	-1.86	-0.26
10	My friends know they can count on me in times of need.	Less	More	My coworkers know they can count on me to help out when they need my assistance.	1.86	-1.75	0.01

Table 21, continued

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
11	When I make mistakes I often blame others.	More	Less	When I make a mistake, I will often try to blame the problem on someone else's actions.	0.81	-2.31	-0.22
14	I would gladly spend some of my leisure time trying to improve my community.	Less	More	I would gladly take on extra hours at work to help my employer with a special project or event.	0.69	-0.69	0.02
16	I go out of my way to keep my promises.	More	Less	When necessary, I will put in extra time or effort to complete the work I said I would do.	1.28	-1.98	0.05
18	When others depend on me, I try to put in an extra effort.	More	Less	When my contributions will help to ensure the success of everyone on a work team, I will try to put in an extra effort.	1.37	-1.85	-0.10
19	There is no use in doing things for people; they never really appreciate your efforts.	Less	Less	There is no use in helping out coworkers, because no one will praise me for my efforts.	1.25	-2.23	-0.15
20	I see no reason for belonging to clubs or community organizations.	Less	More	Although I think there are ways that an employee can add value to a company through unique contributions, I rarely do so.	0.93	-1.28	-0.21

Table 22.  
Rewritten Item Estimates, School FOR Industriousness

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
2	I often need a push to get started.	More	More	It is often hard for me to get started on my coursework unless someone is pushing me to do it or it is due soon.	1.02	0.11	0.06
3	I do not work as hard as the majority of people around me.	Less	More	I do not work as hard as the majority of college students.	1.46	-0.67	-0.55
4	Setting goals and achieving them is not very important to me.	Less	Less	Setting goals in order to be a successful student is not very important to me.	0.80	-2.04	-0.68
7	I do what is required, but rarely anything more.	More	More	I only do the minimum work required by my instructors.	1.36	-0.55	-0.75
10	I handle even difficult tasks smoothly.	Less	Less	I competently handle even the most challenging class projects.	1.16	-0.54	0.08
11	I go above and beyond what is required.	More	More	My performance exceeds the University's minimum requirements for a degree.	0.48	-1.04	-2.40
12	I strive to improve myself.	More	Less	I strive to become an even better student.	1.68	-1.76	-0.01
13	I am a very competent person.	Less	Less	I am a very competent student.	1.25	-1.61	0.15
14	I make every effort to do more than what is expected of me.	Less	Less	I work hard to do more than my teachers expect of me.	2.03	-0.11	-0.39
15	I demand the highest quality in everything I do.	Less	Less	I demand the highest quality from myself on class assignments.	2.68	-0.61	-0.16

Table 22, continued

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
16	I am effective at my work.	More	More	I complete all of my coursework in a thorough and timely manner.	1.26	-0.98	-0.02
17	I set extremely high standards for myself.	More	Less	The standards I set for myself are as high or higher than those expected of me by my teachers and/or family.	1.82	-0.64	-0.59
18	I try to be the best at anything I do.	Less	Less	I try to be the best student that I can be.	1.95	-1.23	-0.08

Table 23.  
Rewritten Item Estimates, School FOR Order

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
1	I frequently forget to put things back in their proper place.	Less	Less	I frequently forget to put my class notes and handouts into the correct folder or file.	1.13	-1.57	-0.61
2	I prefer keeping my options open and rarely plan in advance.	Less	Less	I prefer keeping my options open and rarely plan in advance how I will complete assignments with upcoming deadlines.	1.19	-0.74	0.07
3	I seldom make detailed "to do" lists.	Less	Less	I seldom make detailed "to do" lists to help me complete my schoolwork.	0.83	-1.25	-0.29
5	For me, being organized is unimportant.	Less	Less	For me, it is not important to keep my schoolwork organized.	0.94	-1.81	-2.66
6	Being neat is not exactly my strength.	Less	Less	I do not maintain a very neat and tidy environment.	2.39	-0.52	-0.59
7	I frequently do not put things in their proper place.	Less	Less	When I am done with my schoolwork, I frequently leave a mess of books, notes, and other materials lying around.	1.60	-0.28	-0.93
9	I have a daily routine and stick to it.	Less	Less	I have a daily routine for completing schoolwork that I stick to.	0.97	0.57	-0.13
11	I dislike doing things without proper planning.	Less	Less	I dislike studying or completing an assignment without having a good plan.	0.96	-0.04	-0.36

Table 23, continued

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
12	I prefer to do things in a logical order.	More	Less	I prefer to plan a logical order for completing my coursework before I begin working.	1.23	-1.00	0.01
14	I hardly ever lose or misplace things.	Less	Less	I hardly ever lose or misplace class notes and assignments.	1.25	-1.20	0.12
15	I become annoyed when things around me are disorganized.	Less	Less	I become annoyed when my notes and handouts are disorganized.	1.48	-0.92	-0.34
17	I hate when people are sloppy.	Less	Less	I hate when I see my peers turn in sloppy work.	0.36	1.74	-0.72
18	Every item in my room and on my desk has its own designated place.	Less	More	I make sure that everything in my room or office has its own designated place.	1.99	-0.28	0.18
19	Most of the time my belongings are in complete disarray.	Less	More	Most of the time my room or office is in complete disarray.	1.70	-0.74	-0.18
20	I feel comfortable even in very disorganized settings.	Less	Less	I feel comfortable doing my schoolwork even in very disorganized settings.	1.83	-0.20	0.10

Table 24.  
Rewritten Item Estimates, School FOR Responsibility

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
1	I only think of myself and rarely of others.	Less	Less	I don't take my group members' needs into consideration when working in a group	0.96	-2.58	-0.50
2	I am not the most responsible person among my friends.	Less	More	I am not the most responsible person when it comes to completing my coursework.	1.16	-1.60	-0.03
3	Sometimes it is too much of a bother to do exactly what is promised.	Less	Less	Sometimes it is too much of a bother to follow the instructions for an assignment exactly.	0.86	-1.39	-0.43
4	If I am running late to an appointment, I may decide not to go at all.	Less	More	If I am going to be late to class, I might decide not to go at all.	0.60	-0.72	-0.22
5	Although I think charitable causes are important, I rarely become involved in charitable activities.	Less	More	Although I think that contributing to class discussions makes classtime more valuable to everyone, I rarely do so.	0.35	1.17	-0.38
7	I have a reputation for being late for almost every meeting or event.	Less	Less	I have a reputation for being late to classes and meetings.	0.50	-2.94	0.19
8	I carry out my obligations to the best of my ability.	More	Less	I complete all course requirements to the best of my ability.	1.24	-1.54	-0.87
10	My friends know they can count on me in times of need.	Less	More	My classmates know that they can count on me when they need my help to complete a group assignment.	1.77	-1.35	0.14



Table 24, continued

Item #	Original Scale Item	Relevance (More/Less)	Specificity (More/Less)	Rewritten Scale Item	Rewritten Item Estimate		a-Parameter Difference
					a	b	
11	When I make mistakes I often blame others.	Less	Less	When I do poorly on an exam or assignment I often try to blame it on someone or something outside of my control.	0.56	-2.43	-0.52
12	I hold myself accountable for most of my mistakes.	More	Less	I hold myself accountable if I fail to deliver what I promised to my instructor.	1.42	-1.54	0.24
14	I would gladly spend some of my leisure time trying to improve my community.	Less	More	I would gladly spend some of my leisure time helping an instructor to improve a course.	0.31	2.40	-0.36
16	I go out of my way to keep my promises.	Less	Less	I will go out of my way to do what I promised to my group.	1.40	-1.27	0.01
17	When others are counting on me, I am extra careful to show up on time.	More	More	When others are depending on me to help complete an assignment, I will try to put in an extra effort.	0.99	-2.39	-0.12
18	When others depend on me, I try to put in an extra effort.	More	Less	When my classmates are counting on me, I am extra careful to show up on time.	1.77	-1.71	0.42
19	There is no use in doing things for people; they never really appreciate your efforts.	Less	Less	There is no use for helping out other group members, because it will not benefit me.	1.25	-2.21	0.20
20	I see no reason for belonging to clubs or community organizations.	Less	More	Although I understand that my group members sometimes need help in order to finish an assignment, I see no reason for being the one to do so.	1.11	-1.68	-0.32

Table 25.  
Original and Work-Relevant Scale Reactions Results

	Original Scale			Work-Relevant Scale			Mean Difference	Effect Size (Cohen's d)	Scale Reliability	
	Mean	SD	N	Mean	SD	N			Development Sample	Current Sample
Predictive Validity	2.64	0.65	23	2.87	0.70	21	0.24	0.35	.71	.78
Face Validity	3.95	0.77	23	4.31	0.61	21	0.37*	0.53	.78	.85
Fairness	3.11	0.61	22	3.30	0.54	20	0.19	0.32	.69	.62
Overall Reactions	3.23	0.43	23	3.49	0.42	21	0.26*	0.62	.82	.84
Personal Impact	3.96	0.63	22	4.22	0.48	20	0.26	0.47	.74	.78

*Note.* Asterisks (\*) denote a statistically significant one-tailed t-test,  $p < .05$ . The Overall Reactions score is the mean of predictive validity, face validity, and fairness. Scale reliabilities are alpha values from the sample used for scale development ( $N = 206$ ) and the sample used in the current research.

## FIGURES

Figure 1. Hypothesized Interaction: Item Discrimination

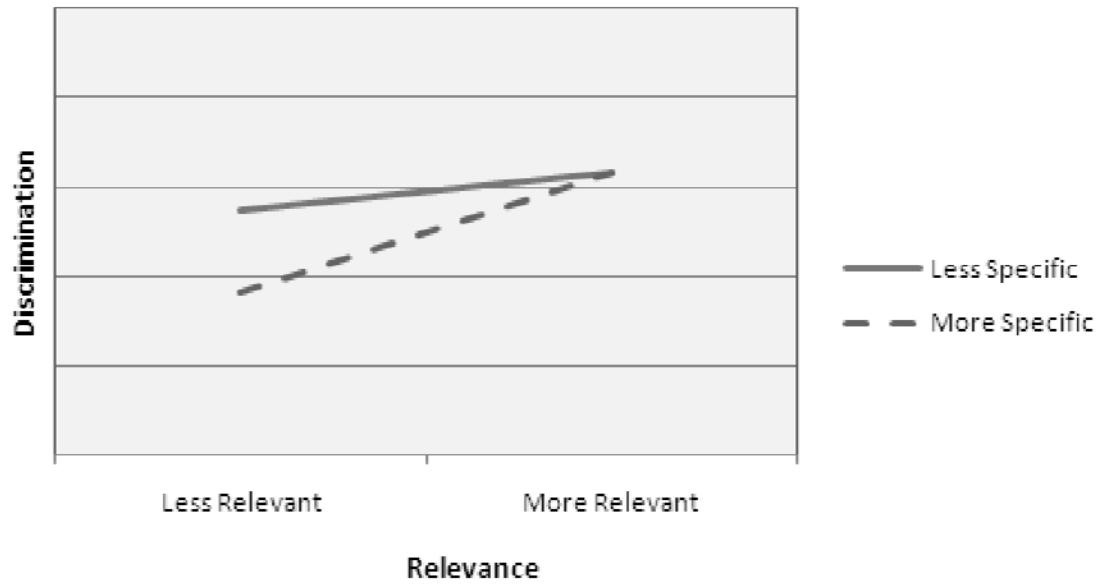


Figure 2: Parallel Analysis: Work FOR Industriousness

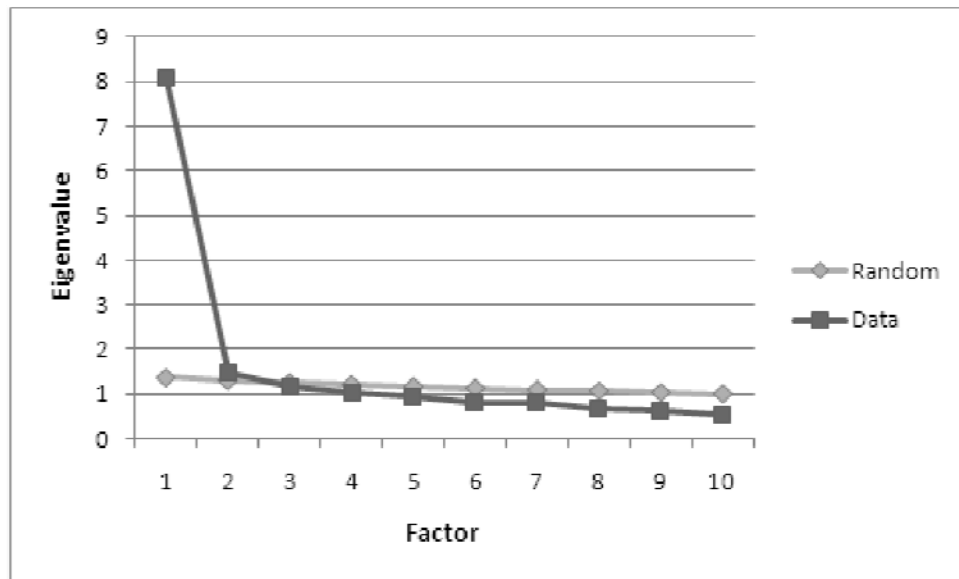


Figure 3. Parallel Analysis: Work FOR Order

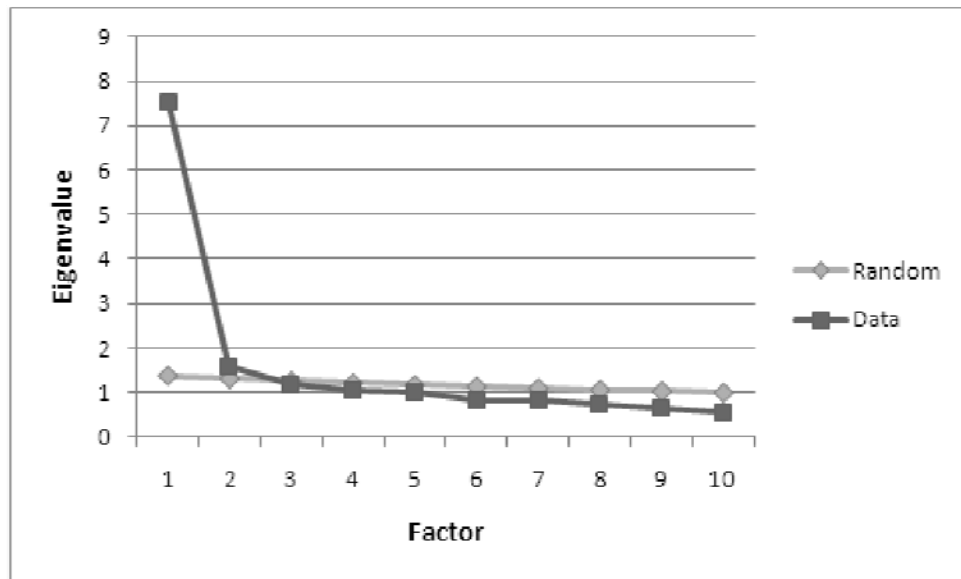


Figure 4. Parallel Analysis: Work FOR Responsibility

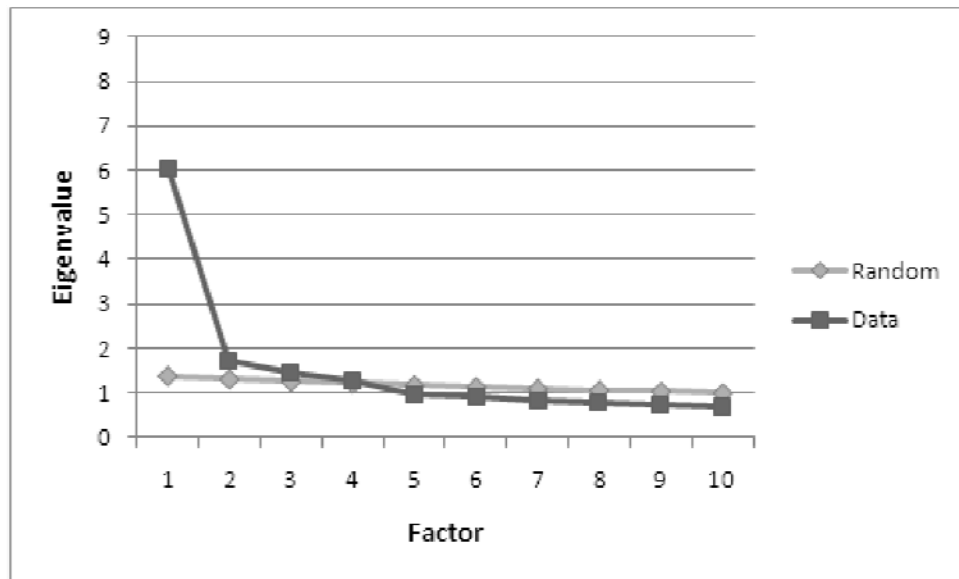


Figure 5. Parallel Analysis: School FOR Industriousness

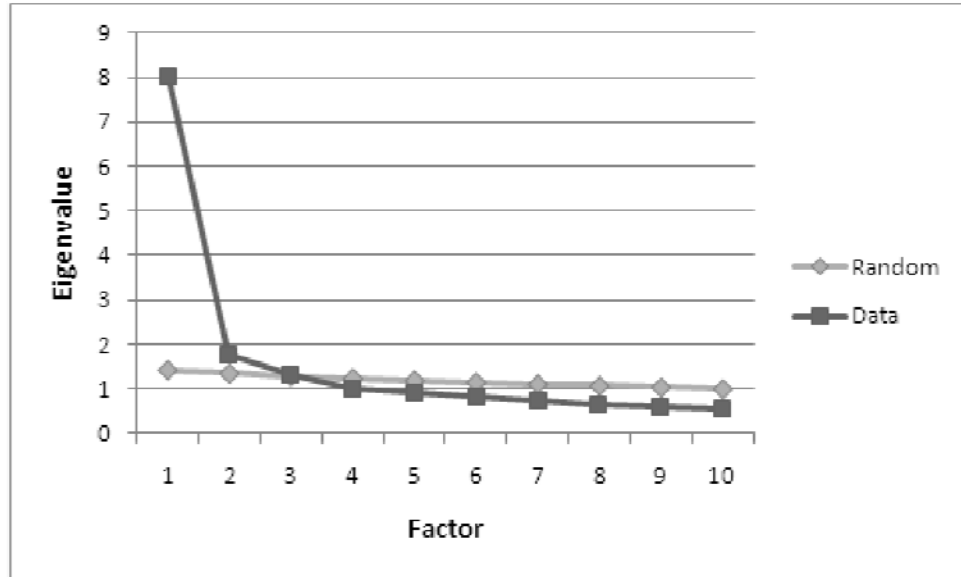


Figure 6. Parallel Analysis: School FOR Order

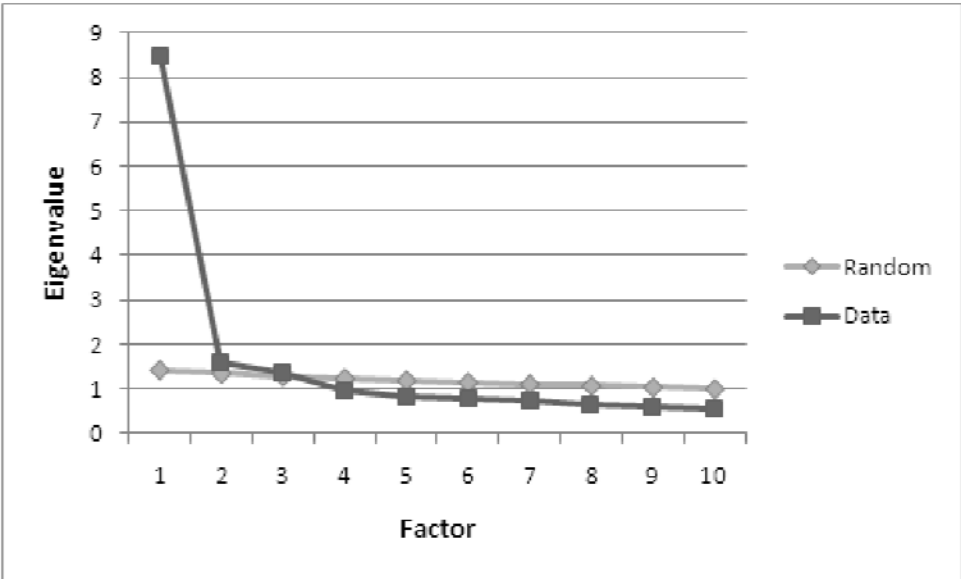




Figure 7. Parallel Analysis: School FOR Responsibility

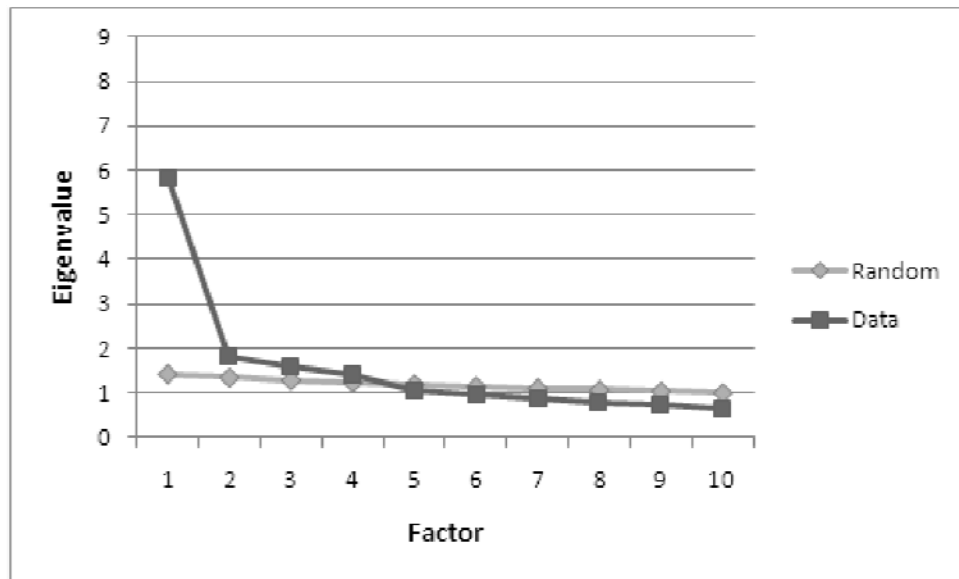


Figure 8. Plot of Relevance Ratings and 2PL Discrimination Estimates, Work FOR

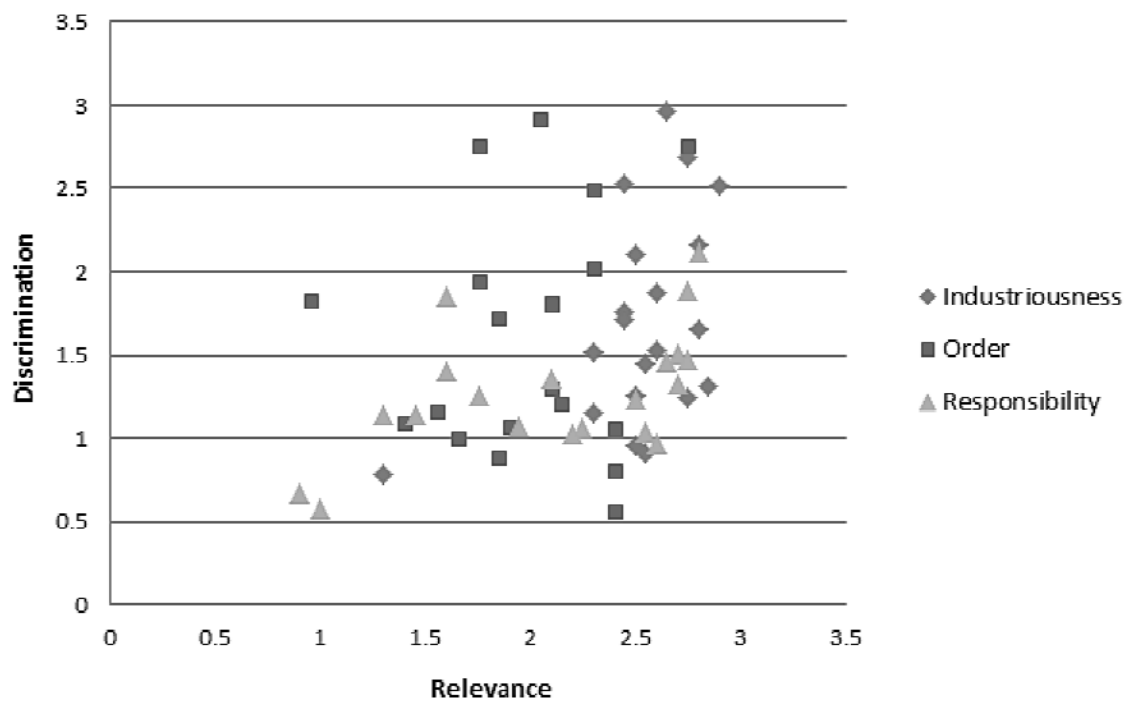


Figure 9. Plot of Relevance Ratings and 2PL Discrimination Estimates, School FOR

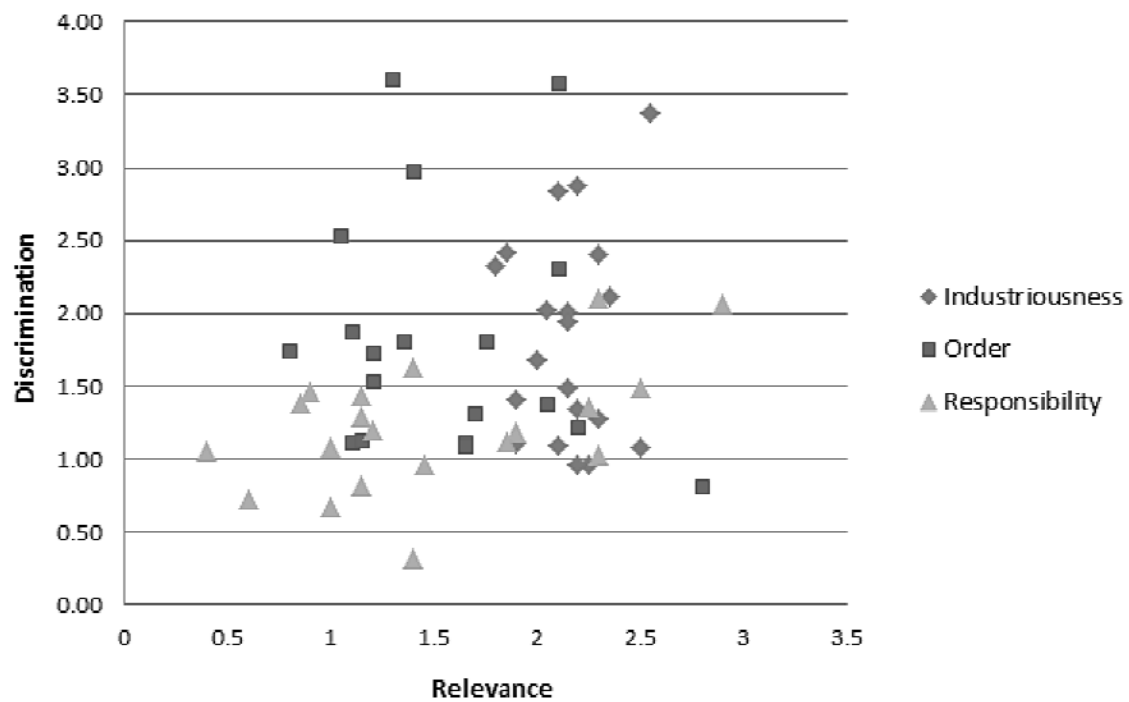


Figure 10. Interaction of Item Relevance and Specificity on Discrimination Estimates,  
Work FOR: Order Facet

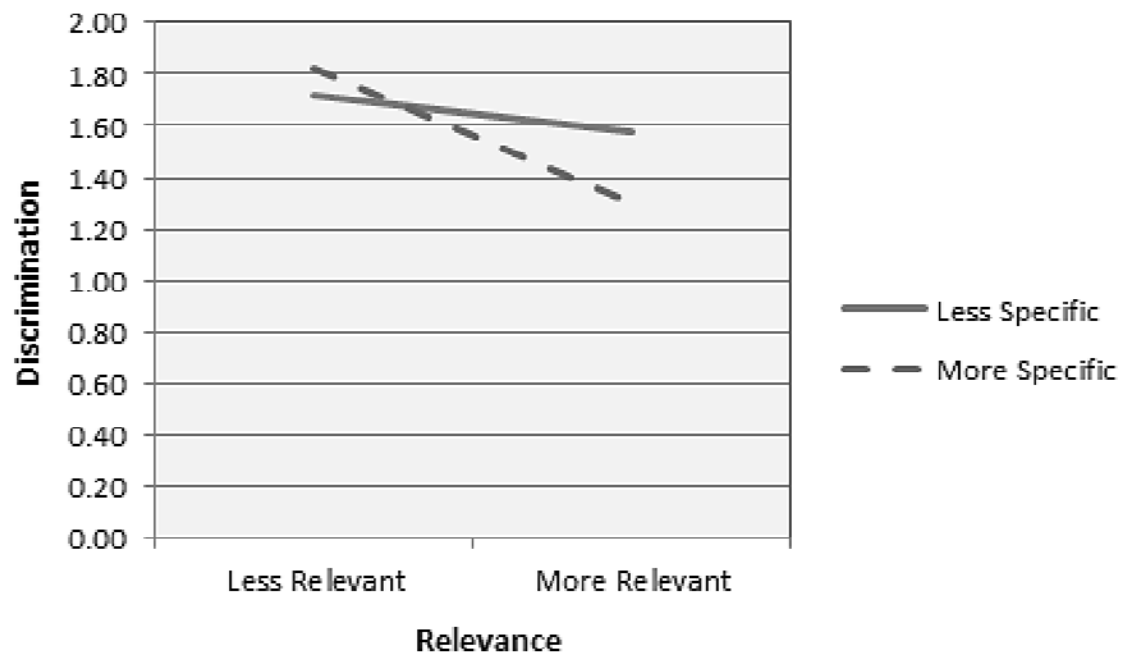


Figure 11. Interaction of Item Relevance and Specificity on 2PL Discrimination Estimates, Work FOR: Responsibility Facet

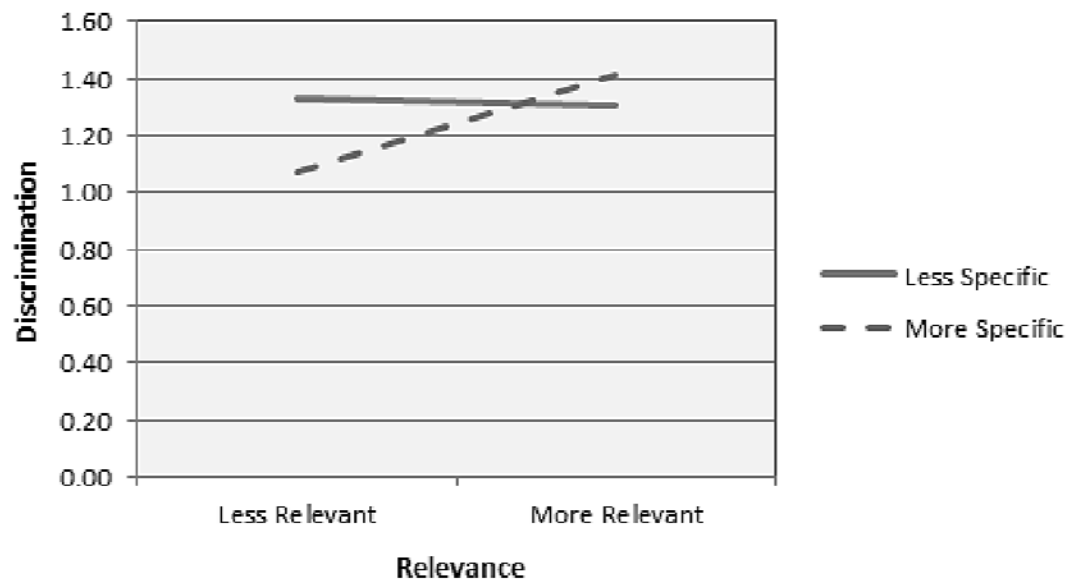


Figure 12. Interaction of Item Relevance and Specificity on 2PL Discrimination Estimates, Work FOR: Overall

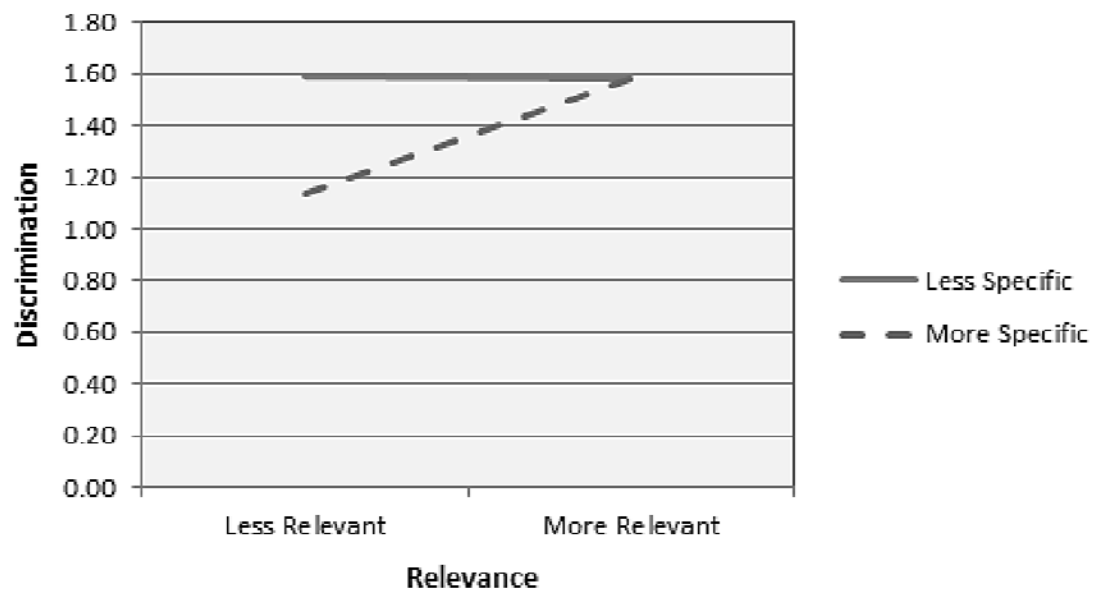


Figure 13. Interaction of Item Relevance and Specificity on 2PL Discrimination  
Estimates, School FOR: Industriousness Facet

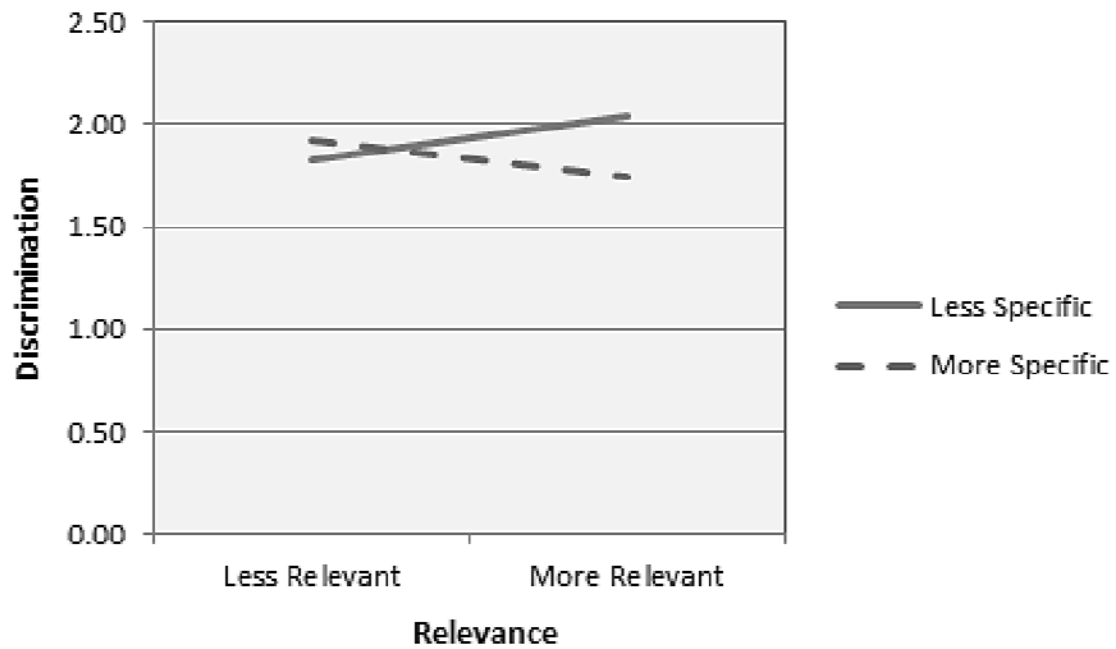


Figure 14. Interaction of Item Relevance and Specificity on 2PL Discrimination  
Estimates, School FOR: Order Facet

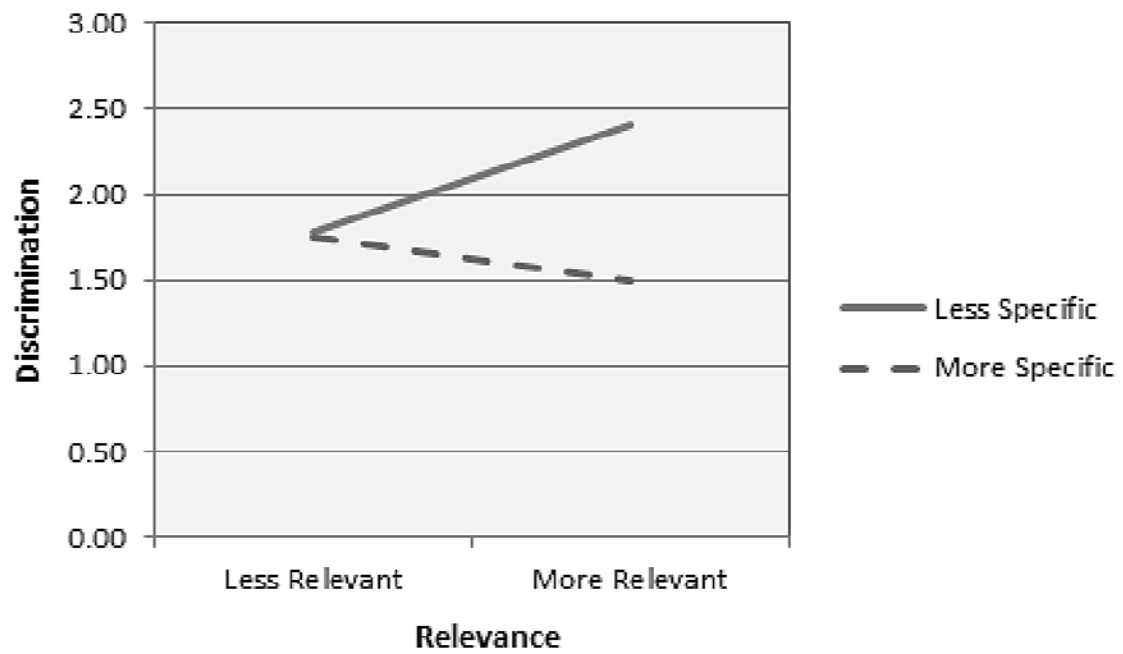




Figure 15. Interaction of Item Relevance and Specificity on 2PL Discrimination  
Estimates, School FOR: Responsibility Facet

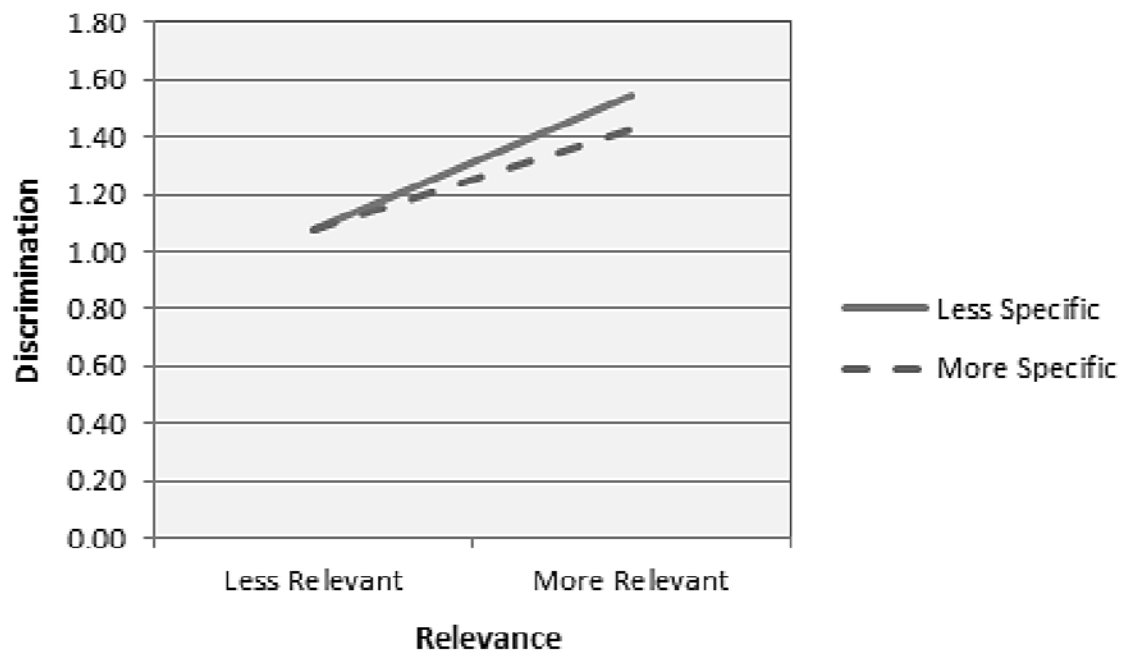


Figure 16. Interaction of Item Relevance and Specificity on 2PL Discrimination  
Estimates, School FOR: Overall

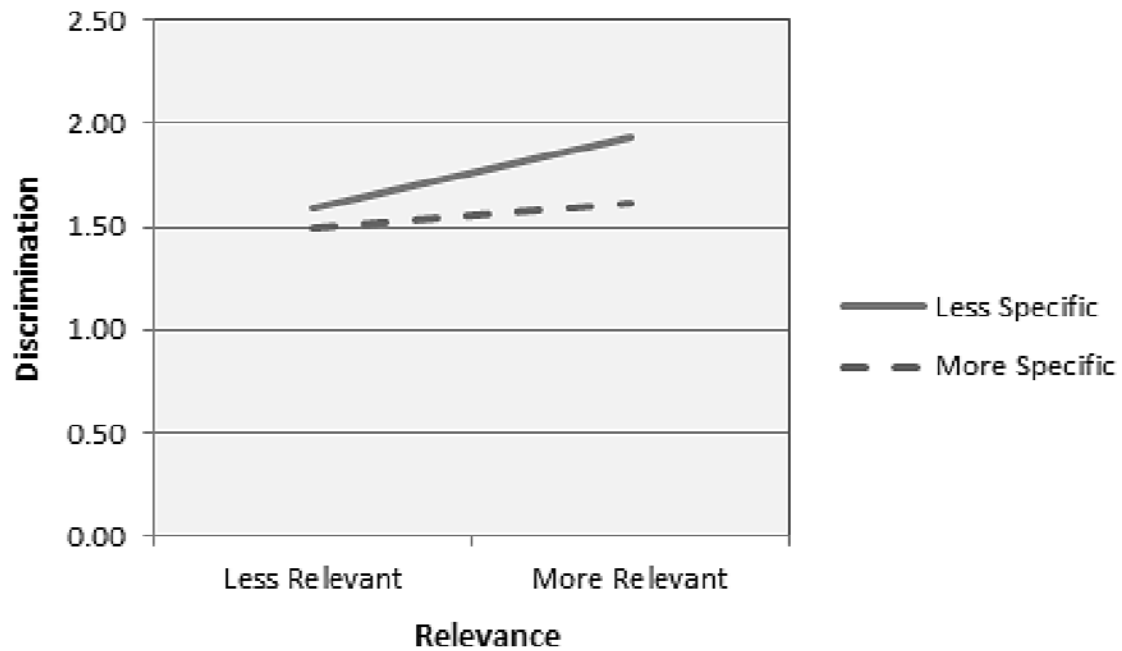
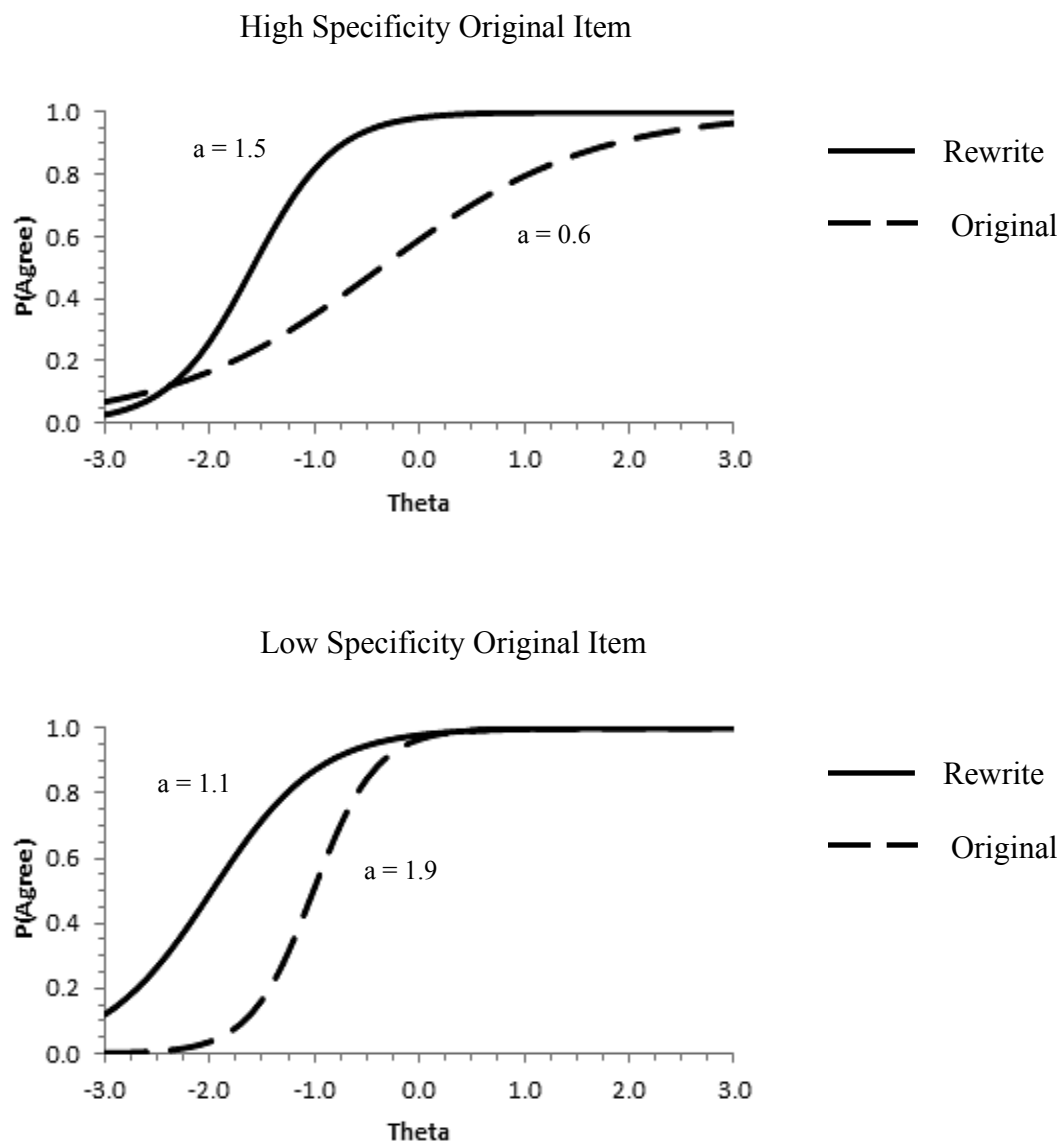


Figure 17. Impact of Original item specificity on Rewritten item change in discrimination



## REFERENCES

- Allport, G. W. & Odbert, H. S. (1936). Trait names: A psycholexical study. *Psychological Monographs*, 47, 211.
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., De Vries, R. E., Di Blas, L., Boies, K., & De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, 86, 356-366.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9(1-2), 9-30.
- Bartram, D., Brown, A., Fleck, S., Inceoglu, S. & Ward, K.(2006). OPQ Technical Manual. Thames Ditton, UK: SHL Group.
- Birenbaum, M., & Montag, I. (1986). On the location of the sensation seeking construct in the personality domain. *Multiple Behavioral Research*, 21, 357-373.
- Carter, N.T., Lake, C.J., & Zickar, M.J. (2010). Toward understanding the psychology of unfolding. *Industrial and Organizational Psychology*, 3, 511-514.
- Cattell, R.B. (1978). Scientific use of factor analysis in behavioral and life sciences. Plenum Press, 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., Jennings, D., Clause, C.S., & Delbridge, K. (1998). Applicant perceptions of test fairness: Integrating justice and self-serving bias perspectives. *International Journal of Selection and Assessment*, 6, 232-239.
- Chernyshenko, O. S., (2002). Applications of ideal point approaches to scale construction and scoring in personality measurement: The development of a six-faceted measure of conscientiousness. Unpublished Dissertation, ProQuest Information & Learning, US.
- Chernyshenko, O.S., Stark, S., Drasgow, F., & Roberts, B.W. (2007). Constructing personality scales under the assumptions of an ideal point response process: Toward increasing the flexibility of personality measures. *Psychological Assessment*, 19, 88 – 106.
- Costa, P.T., Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEOFFI) professional manual. Odessa, FL: Psychological Assessment Resources, Inc.
- Drasgow, F., Chernyshenko, O.S., & Stark, S. (2010). 75 years after Likert: Thurstone was right! *Industrial and Organizational Psychology*, 3, 465-476.
- Drasgow, F., Levine, M. V., Tsien, S., Williams, B., & Mead, A. D. (1995). Fitting polytomous item response theory models to multiple-choice tests. *Applied Psychological Measurement*, 19, 146-165.
- Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology*, 91(1), 40-57.

- Duff, F.L. (1965). Item subtlety in personality inventory scales. *Journal of Consulting Psychology*, 29(6), 565-570.
- English, A. D. (2001). *When personality traits need a frame of reference: Enhancing the predictive validity of non-cognitive measures*. Unpublished Dissertation, Florida Tech.
- Ghiselli, E.E., & Barthol, R.P. (1953). The validity of personality inventories in the selection of employees. *Journal of Applied Psychology*, 37(1), 18-20.
- Gill, C. M., & Hodgkinson, G. P. (2007). Development and validation of the Five-Factor Model Questionnaire (FFMQ): An adjectival-based personality inventory for use in occupational settings. *Personnel Psychology*, 60, 731-766.
- Gilliland, S. W. (1994). Effects of procedural and distributive justice on reactions to a selection system. *Journal of Applied Psychology*, 79, 691-701.
- Glode, J. (2002). *Expanding the justice dilemma: A comparison of stakeholder reactions to personnel selection tools*. Dissertation Abstracts International: Section B: the Sciences and Engineering, 63, 3053.
- Gordon, E.D., & Holden, R.R. (1998). Personality test item validity: insights from “self” and “other” research and theory. *Personality and Individual Differences*, 25, 103-117.
- Griffin, B., & Hesketh, B. (2004). Why openness to experience is not a good predictor of job performance. *International Journal of Selection and Assessment*, 12(3), 243-251.
- Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology*, 8, 135—164.

- Harland, L.K., Rauzi, T., & Biasotto, M.M. (1995). Perceived fairness of personality tests and the impact of explanations for their use. *Employee Responsibilities and Rights Journal*, 8(3), 183-192.
- Heggestad, E.D., & Gordon, H.L. (2008). An argument for context-specific personality assessments. *Industrial and Organizational Psychology*, 1(3), 320–322.
- Heller, D., Watson, D., Komar, J., Min, J.A., & Perunovic, W.Q. (2007). Contextualized personality: Traditional and new assessment procedures. *Journal of Personality*, 75(6), 1229-1229.
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88, 100-112.
- Huang, C. D., Church, A. T., & Katigbak, M.S. (1997). Identifying cultural differences in items and traits: Differential item functioning in NEO personality inventory. *Journal of Cross-Cultural Psychology*, 28, 192-218.
- Hubley, A. M., Wu, A. D., & Zumbo, B. D. (2009). Interpreting IRT parameters: Putting psychological meat on the psychometric bone. Presented at the 117th Annual Meeting of the American Psychological Association (APA), Toronto, ON, Canada.
- Ingerick, M., McFarland, L.A., Vasilopoulos, N.L., & Cucina, J.M. (2004, April). Modeling the impact of applicant reactions on noncognitive test validity: An interactionist perspective. Paper presented at the 19th Annual Conference of the Society for Industrial and Organizational Psychology, Chicago.

- Johnson, J.A. (2004). The impact of item characteristics on item and scale validity. *Multivariate Behavioral Research*, 3(2), 273-302.
- Kaiser, R.B., & Craig, S.B. (2004). Building a better mouse trap: Item characteristics associated with rating discrepancies in 360-degree feedback, *Consulting Psychology Journal: Practice and Research*, 57(4), 235–245.
- Krishnamoorthy, K. & Xia, Y. (2007). Inferences on correlation coefficients: one-sample, independent and correlated cases. *Journal of Statistical Planning and Inference*, 137, 2362–2379.
- Kuncel, R.B., & Kuncel, N.R. (1995). Response-process models: Toward an integration of cognitive-processing models, psychometric models, latent-trait theory, and self-schemas. In P.E. Shrout and S.T. Fiske (Eds.), *Personality, Research Methods, and Theory: A Festschrift Honoring Donald W. Fiske*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lee, K., & Ashton, M.C. (2008). The HEXACO personality factors in the indigenous personality lexicons of English and 11 other languages. *Journal of Personality*, 76(5), 1001-1054.
- Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, 93(2), 268–279.
- Maraun, M. D. (1997). Appearance and reality: Is the Big Five the structure of trait descriptors? *Personality and Individual Differences*, 22(5), 629-647.



- McCrae, R.R. and Costa, P.T. (1989). Reinterpreting the Myers-Briggs Type Indicators from the perspective of the five-factor model of personality. *Journal of Personality*, 57, 17–40.
- Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, 80(4), 252-283.
- 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.
- Neill, J.J., Dunn, O.J., 1975. Equality of dependent correlation coefficients. *Biometrics*, 31(2), 531–543.
- Organ, D.W. (1994). Personality and organizational citizenship. *Journal of Management*, 20(2), 465-478.
- Organ, D.W. & Ryan, A. M. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behaviors. *Personnel Psychology*, 48, 776-801.
- Pace, V. L., & Brannick, M. T. (2010b). Improving prediction of work performance through frame-of-reference consistency: Empirical evidence using openness to experience. *International Journal of Selection and Assessment*, 18, 230-235.
- Paulhus, D.L., Bruce, M.N., & Trapnell, P.D. (1995). Effects of self-presentation strategies on personality profiles and their structure. *Personality and Social Psychology Bulletin*, 21, 100-108.

- 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.
- Podsakoff, P. M., MacKenzie, S. B., Paine, J. B., & Bachrach, D. G. (2000). Organizational citizenship behaviors: A critical review of the theoretical and empirical literature and suggestions for future research. *Journal of Management, 26*, 513-563.
- Reise, S.P., & Waller, N.G. (1990). Fitting the two-parameter model to personality data. *Applied Psychological Measurement, 14*(1), 45-58.
- Roberts, B. W. (2006). Personality development and organizational behavior. In B. M. Staw (Ed.). *Research on Organizational Behavior*. Elsevier Science/JAI Press.
- Roberts, B. W. (2009). Back to the future: *Personality and Assessment* and personality development. *Journal of Research in Personality, 43*, 137-145.
- Roberts, B. W., Bogg, T., Walton, K. E., Chernyshenko, O. S., & Stark, S. E. (2004). A lexical investigation of the lower-order structure of conscientiousness. *Journal of Research in Personality, 38*(2), 164-178.
- 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*(1), 103-139.
- Robertson, I.T., & Kinder, A. (1993). Personality and job competences: The criterion-related validity of some personality variables. *Journal of Occupational and Organizational Psychology, 66*, 225-244.

- Robie, C., Born, M. P., & Schmit, M. J. (2001). Personal and situational determinants of personality responses: A partial reanalysis and reinterpretation of the Schmit et al. (1995) data. *Journal of Business and Psychology*, 16(1), 101-117.
- Robie, C., Schmit, M. J., Ryan, A. M., & Zickar, M. J. (2000). Effects of item context specificity on the measurement equivalence of a personality inventory. *Organizational Research Methods*, 3(4), 348-365.
- Rolland, J.P. (2002). The cross-cultural generalizability of the Five-Factor Model of personality. In R. R. McCrae & J. Allik (Eds.), *The Five-Factor Model of Personality Across Cultures*. New York: Kluwer Academic/Plenum.
- 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.
- Roskam, E. E. (1985). Current issues in item response theory. In E. E. Roskam (Ed.), *Measurement and personality assessment* (pp. 3-19). Amsterdam, The Netherlands: Elsevier Science.
- Ryan, A. M., & Sackett, P.R. (1987). Pre-employment honesty testing: Fakability, reactions of test takers, and company image. *Journal of Business and Psychology*, 1, 248-256.
- Rynes, S. L., & Connerly, M.R. (1993). Applicant reactions to alternative selection procedures. *Journal of Business and Psychology*, 7, 261-277.
- Sackett, P.R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology*, 59, 419-450.

- Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology*, 82, 36—43.
- Saucier, G., & Goldberg, L.R. (1996). The language of personality: Lexical perspectives on the five-factor model. In J.S. Wiggins (Ed.), *The Five-Factor Model of Personality: Theoretical Perspectives*. New York: The Guilford Press.
- Schmit, M.J., Kihm, J.A., & Robie, C. (2000). Development of a global measure of personality. *Personnel Psychology*, 53(1), 153-193.
- Schmit, M. J., & Ryan, A.M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology*, 78, 966-974.
- 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(5), 607-620.
- Schmitt, N., Gooding, R.Z., Noe, R.A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Seiler, S.N. (2009). The measurement of reactions to non-cognitive assessments. Unpublished research.
- Seiler, S.N., & Kuncel, N.R. (2005, April). Modeling the Individual Difference Determinants of Faking: Integration and Extension. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles.

- Seiler, S.N. & Kuncel, N.R. (2007, April). Predicting Faking Using the Faking Dispositions and Reactions Questionnaire (FDRQ). Paper presented at the 22<sup>nd</sup> Annual Conference of the Society for Industrial and Organizational Psychology, New York .
- Shoda, Y, & Mischel, W. (2000). Reconciling Contextualism with the Core Assumptions of Personality Psychology, *European Journal of Personality*, 14, 407-428.
- Silver, N.C., Hittner, J.B., & May, K. (2006). A Fortran 77 program for comparing dependent correlations. *Applied Psychological Measurement*, 30, 152-153.
- Smith, D. B., Hanges, P. J., & Dickson, M. W. (2001). Personnel selection and the five-factor model: Reexamining the effects of applicant's frame of reference. *Journal of Applied Psychology*, 86(2), 304-315.
- 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.
- Stark, S., Chernyshenko, O. S., Chan, K.Y., Lee, W. C., & Drasgow, F. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of Applied Psychology*, 86(5), 943-953.
- Tellegen, A., & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The Sage Handbook of Personality Theory and Assessment: Vol. II. Personality measurement and testing*. London: Sage.
- Tett, R.P., Jackson, D.N., & Rothstein M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703–742.

- Waller, N.G., & Zavala, J.D. (1993). Evaluating the Big Five. *Psychological Inquiry*, 4(2), 131-134.
- Werner, P.D., & Pervin, L.A. (1986). The content of personality inventory items. *Journal of Personality and Social Psychology*, 51(3), 622-628).
- Williams, E.J. (1959). Significance of difference between two non-independent correlation coefficients. *Biometrics*, 15(1), 135–136.
- Williams, L.J., & Anderson, S.E. (1991). Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management*, 17(3), 601-617.
- Wood, D. (2007). Using the PRISM to compare the explanatory value of general and role-contextualized trait ratings. *Journal of Personality*, 75(6), 1103-1126.
- Wood, D., & Roberts, B.W. (2006). Cross-sectional and longitudinal tests of the personality and role identity structural model (PRISM). *Journal of Personality*, 74(3), 780-810.
- Wright, J. C. and Mischel, W. (1987). A conditional approach to dispositional constructs: The local predictability of social behavior. *Journal of Personality and Social Psychology*, 55, 454-469.
- Zickar, M. J., & Robie, C. (1999). Modeling faking good on personality items: An item-level analysis. *Journal of Applied Psychology*, 84(4), 551-563.
- Zickar, M. J., & Ury, K.L. (2002). Developing an interpretation of item parameters for personality items: Content correlates of parameter estimates. *Educational and Psychological Measurement*, 62(1), 19-31.

Zumbo, B.D., Pope, G.A., Watson, J.E., & Hubley, A.M. (1997). An empirical test of Roskam's conjecture about the interpretation of an ICC parameter in personality inventories. *Educational and Psychological Measurement*, 57, 963-969.

## APPENDIX A

### Instructions for Perceived Work Relevance Rating Task (Study 1)

#### *Part 1: Work-Relevance Ratings*

For this task, you will rate each item in terms of your perception of its direct relevance to the workplace. Do not rate the item in terms of your judgment of how the item might be related to work. Instead, provide a rating of how much the item directly reflects a work behavior or characteristic. That is, how much does the item sound like it could describe a person at work?

Example: I would judge that the item, “I am the life of the party” does not reflect a work characteristic, even though I can infer that it is related to how I interact with people at work. So, I would probably rate it 0 or 1. I would judge the item, “I find it easy to interact with people I have just met” to be much more work-relevant, so I would probably rate it a 2 or 3.

Use the following 0-3 point scale for your ratings:

- 0: Does not describe a work behavior/characteristic at all**
- 1: Describes a work behavior/characteristic a little**
- 2: Describes a work behavior/characteristic fairly well**
- 3: Describes a work behavior/characteristic extremely well**



## APPENDIX A, continued

### *Part 2: Suggest Rewrites*

For any or all of the items you rated less than a 3, suggest a rewrite to the item so that it measures the same type of behavior or characteristic but uses language that is more work-relevant.

#### **Example 1**

##### Original Item

When I have too many chores to do, I often ignore them all.

##### Rewritten Item

When I have too many organizing and administrative tasks to complete, I often ignore them all.

#### **Example 2**

##### Original Item

When I make a commitment, I can always be counted on to follow through.

##### Rewritten Item

My coworkers can count on me to do the work I said I would do.

## APPENDIX B

### Frame of Reference Instructions

#### *Work FOR Instructions*

“On the following pages, there are statements and words describing people's characteristics and behaviors. Please use the rating scale below to describe how accurately each statement or word describes you when you are at work. You can think about yourself in terms of your work characteristics in a job you currently have, one you held in the past, or a combination of jobs you have held. Describe yourself as you honestly see yourself. Please read each item carefully, and then fill in the bubble that corresponds to the number on the scale. Remember, please respond to the statements in terms of your characteristics at work.”

#### *General FOR Instructions*

“On the following pages, there are statements and words describing people's characteristics and behaviors. Please use the rating scale below to describe how accurately each statement or word describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself. Please read each item carefully, and then fill in the bubble that corresponds to the number on the scale.”

## APPENDIX C

### Williams and Anderson (1991) Job Performance Scales Used in the Current Research

Scale	Item
In-Role Behavior	1. I adequately complete my assigned duties.
	2. I fulfill the responsibilities specified in my job description.
	3. I perform the tasks that are expected of me.
	4. I meet the formal performance requirements of the job.
	5. I engage in activities that will positively affect my performance evaluation.
	6. I neglect aspects of the job I am obligated to perform.
	7. I fail to perform essential duties.
Organizational Citizenship Behavior – Individual	1. I help others who have been absent.
	2. I help others who have heavy work loads.
	3. I assist my supervisor with his/her work (when not asked).
	4. I take time to listen to co-workers' problems and worries.
	5. I go out of the way to help new employees.
	6. I take a personal interest in other employees.
	7. I pass along information to co-workers.

## APPENDIX D

### Personality Test Reactions Items (Seiler, 2009)

#### **Predictive Validity**

- 1 I am confident that the assessment would predict how well an applicant will perform on the job.
- 2 An employer can't tell much about an applicant's ability to do the job from his/her responses to the measure.
- 3 Applicants who perform well on this measure will perform well on the job.
- 4 I am confident that an employer can use the test to predict how well an applicant will perform a job.
- 5 This test does not really measure a person's true abilities for a job.
- 6 Using this test is a good way to find out if a person is really suited for a job.

#### **Face Validity**

- 7 I understand what the items have to do with a job.
- 8 The content of the assessment is related to what is required in a job.
- 9 I cannot see any relationship between the test and what is required on the job.
- 10 The actual content of the personality instrument is clearly related to job tasks.
- 11 There is no real connection between the examination and job tasks.

#### **Fairness**

- 12 I feel that using the test to select applicants for a job is unfair.
- 13 This test would give all people an equal opportunity to compete for the job.
- 14 It is perfectly appropriate for an employer to administer the instrument to applicants.
- 15 Employers have the right to obtain information from applicants by using this method.
- 16 The content appears prejudiced against certain people.
- 17 This type of test is a fair way of assessing an applicant's potential.
- 18 Employers have the right to obtain information from applicants through this type of assessment.
- 19 Applicants would get a fair outcome as a result of their responses to the items.

#### **Personal Impact**

- 20 I would resent being asked to take this test for a job.
- 21 This test would be an invasion of my privacy.
- 22 I would enjoy being asked to take this test for a job.
- 23 I would refuse to take the test, even if it meant losing a chance at the job.
- 24 This type of assessment asks too much personal information.

## APPENDIX D, continued

### **Personal Impact, continued**

- 25 This type of test does not respect the rights and dignity of an applicant.
- 26 I would be comfortable taking this test in a selection setting.
- 27 This procedure treats applicants respectfully.
- 28 Some of the items are offensive.