

1-1-2010

Who Am I? Well, It Depends: How Frame-of-Reference Imposes Context In Non-Contextualized Personality Inventories

Elizabeth Anne McCune
Portland State University

Let us know how access to this document benefits you.

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Recommended Citation

McCune, Elizabeth Anne, "Who Am I? Well, It Depends: How Frame-of-Reference Imposes Context In Non-Contextualized Personality Inventories" (2010). *Dissertations and Theses*. Paper 725.

10.15760/etd.725

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.

Who Am I? Well, It Depends:
How Frame-of-Reference Imposes Context
In Non-Contextualized Personality Inventories

by

Elizabeth Anne McCune

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Applied Psychology

Dissertation Committee Members:

Donald Truxillo, Chair
Talya Bauer
Todd Bodner
Leslie Hammer
Wayne Wakeland

Portland State University
©2010

ABSTRACT

The frame-of-reference (FOR) effect refers to the finding that validities for personality measures can be improved by asking respondents to consider how they behave in a particular context (e.g., “at work”). Recently, Lievens, De Corte, and Schollaert (2008) demonstrated that a FOR serves to reduce within-person inconsistencies in responding, which then improves the reliability and validity of personality measures. Despite this important step forward in FOR research, Lievens et al. note that there is still very little known with regard to how respondents complete non-contextualized personality inventories (i.e., inventories where no FOR is provided).

The present studies sought to fill this significant gap in the literature by addressing the question: Do people think of themselves in particular situations or contexts when responding to non-contextualized personality inventories and, if so, what are these contexts? In addition, does the use of context vary by the personality dimension being studied? Two studies were conducted in order to fully address these Research Questions. The first of these studies was a qualitative study which examined the number and types of contexts spontaneously generated by test-takers for non-contextualized personality items. Twenty-eight interviews were conducted with college students who held a variety of life roles (e.g., student, employee, parent, spouse). Interview data demonstrated that participants considered themselves in general, at school, at work, with friends, with family, at home, and in other more specific situations (e.g., driving a car) when responding to non-contextualized inventories.

Data for Study 2 were collected from 463 college students using a self-report methodology that asked participants to indicate which FORs they were using in responding to the same non-contextualized inventory used in Study 1. Results indicated significant differences in FOR endorsement across factors, such that participants endorsed the highest number of FORs for agreeableness items and the lowest number of FORs for openness to experience items. In addition, there were significant differences in the use of FORs within factors such that, for example, the “With Family” FOR was used most frequently for agreeableness but the “At School” FOR was used most frequently for openness to experience. Finally, results of Study 2 indicated that while the using more FORs in responding may increase error variances, it does not have a substantial impact on the factor structure of the Big 5.

The present studies contribute to the literature by being the first to examine the role that situations play in responding to a non-contextualized inventory, and they do so using both qualitative and quantitative methods. In addition, the present studies represent a person-centric approach to the study of I/O psychology in that they focus on the individual experience as the basis for research.

DEDICATION

For my parents, for never wanting me to be anything except who I am, and for Donald,
for showing me who I could become.

ACKNOWLEDGEMENTS

I first need to thank my chair, advisor, mentor, and friend, Dr. Donald Truxillo. These few lines are nowhere near what is needed to thank him for everything he has taught me and all of the support he has provided over the last five years. He is truly one-of-a-kind and I cannot imagine how I got lucky enough to have him as my advisor. I would also like to thank the rest of my committee: Dr. Talya Bauer, for setting high expectations and giving me the tools necessary to meet them, Dr. Todd Bodner, for fostering my love for all things quantitative, Dr. Leslie Hammer, for her input on the theoretical development of this work, and Dr. Wayne Wakeland, for his enthusiasm and positive attitude toward this project.

I also need to thank a number of people without whom the project could not have been completed. Brian Breniser and Victoria Prescott helped conduct Study 1 interviews, Josh Carlsen served as a second rater for Study 1 interviews, and David Meier and Tony Sinclair assisted in recruitment for Study 2. I would also like to thank all of the instructors who allowed me to recruit participants from their courses, and Julie Bilski for helping me find the space to conduct the Study 1 interviews.

I would also like to acknowledge the support I have received from Kronos, Inc. and Microsoft, Corp. The motivation for this study came in part from my work in selection that took place at Kronos, and I would specifically like to acknowledge (in alphabetical order) Kristin Charles, Autumn Krauss, Annette Kuhn, Ryan Robinson, and Aarti Shyamsunder for all of the support they provided me with during my time with Kronos. I would also like to acknowledge Rupert Bader, Shawn Del Duco, and

Alexis Fink from Microsoft because without their flexibility, support, and understanding I would not have been able to complete this dissertation in my desired timeframe.

Finally, I owe a debt of gratitude to my friends, family, and loved ones who have supported me in various and wonderful ways. My dearest friends, Jennifer Lampi, Gina Magharious, Tivra Marks, and Jennifer VanDuker have served as patient sounding boards and an endless source of laughter and joy. I would especially like to thank Staci Wendt for everything that she has done to support me in this process (including serving as a courier of my dissertation documents) and just for being an incredible person who I am grateful to call my friend. My family is due immense thanks for their patience with me, and for always providing me with a respite from the stress of graduate school. My great aunt and uncle, Ken and Margaret Smith, have provided me with a home away from home on countless occasions and I am so grateful for them. My sisters, Jennifer and Karen, my five amazing nephews (Anthony, Corbin, Deven, Ethan, and Nicholas) and one very special niece (Brenae) remind me of the joy that is found in the little things in life. I thank my parents for their persistence in trying to understand exactly what it is that I want to be when I grow up, and for their unconditional love and unwavering confidence in my ability to succeed. Lastly, I will try (and fail) to express how thankful I have been to have Jeffrey Stocum in my life for the last four years. Jeff has made my life more rich, rewarding, and fulfilling, yet somehow simultaneously easier. I thank him for all that he is, and can't wait for our post-graduate adventures to begin.

TABLE OF CONTENTS

Abstract.....*i-ii*

Dedication.....*iii*

Acknowledgements.....*iv-v*

List of Tables.....*viii-x*

List of Figures.....*xi*

Chapter 1: Introduction.....1

Chapter 2: The Five Factor Model of Personality in Personnel Selection.....5

Chapter 3: Frame of Reference Effects in Personality Measures.....27

Chapter 4: Development of Research Questions for Study 149

Chapter 5: Study 1.....53

 Section 5.1: Method..... 56

 Section 5.2: Results.....64

 Section 5.3: Discussion.....76

Chapter 6: Development of Research Questions for Study 286

 Section 6.1: Method..... 90

 Section 6.2: Results.....95

 Section 6.3: Discussion.....114

Chapter 7: General Discussion.....129

References.....183

Appendix A: Study 1 Screening Survey.....194

Appendix B: Interviewer Cheat Sheet.....196

Appendix C: Researcher Script and Follow-Up Questions for Study 1197

Appendix D: IPIP Personality Inventory.....199

Appendix E: Study 2 Survey.....201

List of Tables

Table 1. Research Questions and Methodology for the Proposed Study.....145

Table 2. Labels and Definitions for the FFM Domains and Facets
(Costa & McCrae, 1992).....146-147

Table 3. Examples of Key Quotes for Each Study 1 FOR Category.....148-149

Table 4. Percent of FORs Appearing within Each Factor for Study 1150

Table 5. Item Order for Each Survey Version.....151-152

Table 6. Means, Standard Deviations, and Confidence Intervals for
Each FOR within Each Item.....153-155

Table 7. RQ3: Factor Level Mean Differences in Total FOR Endorsement.....156

Table 8. RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise
Mean Differences for Total FOR Endorsement across Agreeableness Items.....157

Table 9. RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise
Mean Differences for Total FOR Endorsement across Conscientiousness Items.....158

Table 10. RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise
Mean Differences for Total FOR Endorsement across Extraversion Items159

Table 11. RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise
Mean Differences for Total FOR Endorsement across Neuroticism Items.....160

Table 12. RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise
Mean Differences for Total FOR Endorsement across
Openness to Experience Items.....161

Table 13. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for FORs within Agreeableness.....162

Table 14. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for FORs within Conscientiousness.....163

Table 15. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for FORs within Extraversion.....164

Table 16. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for FORs within Neuroticism.....165

Table 17. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for FORs within Openness to Experience.....166

Table 18. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for Factors within the In General Response Option.....167

Table 19. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for Factors within the Family FOR.....168

Table 20. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for Factors within the Friends FOR.....169

Table 21. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for Factors within the Home FOR.....170

Table 22. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for Factors within the School FOR.....171

Table 23. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for Factors within the Work FOR.....172

Table 24. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise

Mean Differences for Factors within the Other Response Option.....173

Table 25. Descriptive Statistics for Endorsement of Response Options.....174

Table 26. Correlations between Response Option Endorsement Frequency
and Individual Differences.....175

Table 27. Results of CFA Models.....176-177

Table 28. Means, SDs, and t-test Results for Differences in Mean FOR
Endorsement Based on Actual Survey Response Context.....178-179

List of Figures

Figure 1. McCrae and Costa’s (1996) Meta-Theoretical Framework of
Personality.....180

Figure 2. FOR Endorsement by Factor Interaction.....181

Figure 3. Meyer (2010) Taxonomy of Work Situations.....182

Chapter 1

Introduction

The Five Factor Model (FFM; Costa & McCrae, 1992) of personality is a useful framework for measuring normal personality and has been highly utilized in the field of industrial/organizational (I/O) psychology. Although the FFM has been used in a variety of contexts in I/O psychology, the utility and validity of using personality measures in employment selection is an often debated topic that has garnered much attention over the past two decades (e.g., Barrick & Mount, 1991; Gatewood, Feild, & Barrick, 2007; Morgeson et al., 2007). While practitioners tout the importance of considering one's personality when determining whether or not an individual will be a "good fit" for a position and anecdotal evidence speaks strongly regarding the pivotal role that personality plays in job and career success, validity coefficients for the FFM are generally modest (e.g., Barrick, Mount, & Judge, 2001).

As suggested by Mount, Barrick, and Strauss (1994), one possible reason for the modest validity coefficients of personality measures may be differences in the frames of reference (FORs) used by respondents. For example, some respondents might be considering their behavior at work, while others are thinking of themselves at school, at home, or in a multitude of contexts. Schmit, Ryan, Stierwalt, and Powell (1995) were the first to formally test the FOR effect, and their results indicated that validity coefficients could be improved by asking respondents to think of how they behave "at work" when responding to a personality measure. Additional studies conducted in this same vein (e.g., Bing, Whanger, Davison, & VanHook, 2004;

Hunthausen, Truxillo, Bauer, & Hammer, 2003; Small & Diefendorff, 2006) have found similar results.

Most recently, Lievens, De Corte, and Schollaert (2008) challenged the traditional explanation of the FOR effect. Prior to the Lievens et al. study, researchers had assumed that providing a FOR reduced between-person variability in responding, and in doing so improved reliability which then improved validity. However, Lievens et al. showed that providing a FOR has only a modest impact on between-person variability; rather, the improvement in reliability (and validity) is due to the reduction in within-person variability in responding.

Although the Lievens et al. (2008) study made an important contribution by disproving the traditional explanation of the FOR effect and presenting a valid alternative explanation, the authors note that there is still much that researchers do not fully understand about how FORs are used by respondents when completing personality measures. Clearly, as demonstrated by the FOR literature, the context plays an important role in determining how respondents will complete a personality measure, but what happens when a context is not specified, as is the case with the vast majority of personality measures?

The present studies are aimed at filling significant gaps in the literature on personality and FOR. The first contribution will be a greater understanding of the contexts, or FORs, used by respondents when completing a non-contextualized personality inventory. To my knowledge, this basic research question has not yet been addressed, and doing so should provide valuable information with regard to the extent

to which situational behavior may be considered even when responding to so-called non-contextualized or “generic” personality inventories. The second contribution will involve a deeper investigation of the FORs used in non-contextualized personality inventories and the potentially meaningful ways these FORs may differ across items, and across and within individuals. Certain items or factors may elicit more or fewer FORs, and the pattern of these FORs may be more or less consistent across items and factors. Further, individuals should be more or less likely to endorse relatively more or fewer FORs, which may impact the factor structure of the FFM. The final contribution of the proposed studies is that they utilize multiple methodologies in order to answer the research questions. Personality research in I/O psychology has been criticized for a reliance on self-report questionnaires (e.g., Hough & Furnham, 2003). Thus, the proposed studies rely on qualitative methods in addition to an innovative self-report methodology that allows participants to elaborate on their responses. Table 1 summarizes the research questions that will be addressed by each of the studies.

Overview of the Dissertation

The dissertation proceeds as follows. First, an overarching theory of personality is presented in order to situate the FFM in a theoretical context, since it is often criticized as being atheoretical. Second, the FOR literature is reviewed. Initial research questions are then developed by synthesizing the literature from the first two sections. At this point the first study, including the rationale, method, and results, are presented. In brief, Study 1 is a qualitative study aimed at understanding the range of contexts in which participants consider themselves when responding to a non-

contextualized personality inventory. This study provides a greater understanding of the contexts implied by a non-contextualized personality inventory and serves as input for Study 2. The discussion for Study 1 addresses the insights gained and potential limitations.

Following the discussion of the first study are additional research questions that are examined in Study 2. These questions stem from both the literature review and the results of the first study. The Study 2 method section then goes on to describe the quantitative approach taken to understanding the context implied by non-contextualized personality inventories. Study 2 shows how items and factors differ in the degree to which they elicit multiple FORs, which FORs certain items/factors tend to elicit, and that individual differences in the tendency to utilize multiple FORs has little to no impact on the factor structure of a non-contextualized personality inventory. The Study 2 discussion further explores these findings and addresses potential limitations. The last chapter is a general discussion that weaves together the results of both studies, addresses implications for research and practice, and provides possible directions for future research.

Chapter 2

The Five Factor Model of Personality in Personnel Selection

The purpose of this chapter is threefold. First, it will describe an overarching framework of personality that is useful in understanding how personality leads to behavior and how behavior can be expected to change across situations. It will review the use of the FFM in personnel selection, and lastly there will be a discussion of how the validity of personality measures might be improved by a greater consideration of context.

A Meta-Theoretical Framework for Personality

The question of what defines a person's fundamental nature, or personality, is at its essence a deeply philosophical one that the field of psychology has only recently begun to attempt to answer. The concept of character from Aristotle's *Nicomachean Ethics* dates back to 350 B.C., and more than 2,000 years elapsed before trait theorists in psychology entered the discussion in the 1930s. In the relatively short period of time since then, personality theories have proliferated in a highly unsystematic fashion, resulting in many interesting ideas that have little or no connection to one another. In an attempt to remedy this problem, McCrae and Costa (1996) present a meta-theoretical framework for personality under which, they believe, basically all personality theories can be encompassed, and which provides a structure that will allow new personality theories to develop and flourish. This chapter begins with an explanation of that meta-theoretical framework, followed by a discussion of the FFM of personality which is easily situated in the framework.

McCrae and Costa (1996) argue that personality theories serve three functions for personality theorists and researchers. The first function is to “serve as a vehicle for addressing basic philosophical questions about human nature” (McCrae & Costa, p. 55). For example, the psychoanalytic view of personality implied certain assumptions about human nature (e.g., that it is basically hedonistic) that other theorists disagreed with strongly. The theories of personality that emerged as a response to the psychoanalytic view (e.g., the humanistic perspective) were contesting the accuracy of those philosophical assumptions. The debates were not data-driven, rather they reflected differences in beliefs about human nature.

The second function of personality theories according to McCrae and Costa (1996) is to house insights regarding psychological mechanisms and individual characteristics. A strong theory of personality should contain information about individual differences and the psychological mechanisms by which these differences lead to behavior. Finally, personality theories themselves should help to define the field of personality psychology by identifying its scope and limits. That is, personality theories should explicitly identify variables of interest and the behaviors or phenomena they are attempting to explain.

In addition to the three basic functions a personality theory should serve, it should also encompass the six elements depicted in Figure 1 (McCrae & Costa, 1996). Dynamic processes appear throughout the model and they identify the nature of the interactions between the five elements that appears in the boxes. These dynamic processes might often be psychological mechanisms, but it is the responsibility of the

theorist to determine whether that is the case and, if so, what the precise mechanisms are.

Regarding the elements that are framed in boxes, the first of these is basic tendencies which are the “universal raw material of personality” (p. 66). That is, basic tendencies refer to dispositions and other characteristics that are often unobservable, as well some basic physical characteristics. Examples given by McCrae and Costa include genetics, physical appearance, intelligence, physiological drives, focal vulnerabilities (e.g., proneness to alcoholism, tendencies toward bipolar disorder), and personality. Basic tendencies may be inherited or formed at an early life stage, and they are thought to be the essential part of the person and definitive of one’s potential. McCrae and Costa consider personality traits to be the most important of the basic tendencies, and they define them not as frequencies or consistencies in behavior, but as “abstract dispositions” (p.69).

The next element in the meta-theoretical framework is characteristic adaptations, which are the product of the individual and their environment. That is, an individual’s basic tendencies provide the raw material with which their environments interact, and the result is the acquisition of habits, attitudes, and skills. McCrae and Costa (1996) refer to characteristic adaptations as the “concrete manifestation of basic tendencies” (p. 69). McCrae and Costa view the distinction between basic tendencies and characteristic adaptations as critical to a meta-theoretical framework of personality. In their view, the distinction serves to explain how theoretically universal dimensions of personality can actually vary across cultures, and it also explains the

changes in personality that occur over the lifespan (Costa & McCrae, 1980). The basic idea, as explained by McCrae and Costa, is that basic *tendencies* are exactly that, tendencies. They reflect inclinations or predispositions toward patterns of behavior, but they are quite abstract in nature. Characteristic adaptations, on the other hand, reflect the role that the environment plays in the expression of these basic tendencies. Certain environmental influences may promote or diminish the extent to which basic tendencies are expressed. Additionally, different environmental influences will lead to, potentially, different expressions of the same basic tendency. This distinction between basic tendencies and characteristic adaptations, and the ability of the latter to explain how behavior can vary across situations, is a concept that is integral to the proposed study.

The fourth element in McCrae and Costa's (1996) meta-theoretical framework is the self-concept. As indicated in Figure 1, the self-concept can actually be considered a specific characteristic adaptation that emerges based on the interaction between an individual's basic tendencies and their environment. Indeed, as demonstrated in the figure, basic tendencies directly influence the self-concept. However, so many personality theories are focused around the concept of "self" that McCrae and Costa call it out as a separate, but highly related, element. According to McCrae and Costa, the self-concept "consists of knowledge, views, and evaluations of the self, ranging from miscellaneous facts of personal history to the identity that gives a sense of purpose and coherence to life" (p. 70). The self-concept is viewed differently by different types of personality theorists, but perhaps the most relevant

perspective on the self-concept is that of the trait theorists who believe it to be a reasonably accurate description of an individual, while accepting the fact that there are occasions in which it may be distorted. The assumption that one's self-concept is an accurate representation of a person's inner nature or basic tendencies is a critical assumption in self-report methodology.

The fifth element in McCrae and Costa's (1996) framework is the objective biography which, for all intents and purposes, is basically one's behavior. Depending on the theory, objective biographies may take the form of observable behaviors, retelling of dreams, emotional reactions, and more. For most theories, the objective biography is the outcome that the researcher is trying to predict. It is important to note that although the objective biography is typically the outcome in a particular study, the figure demonstrates that behaviors actually feed back into the self-concept, helping to shape and change it.

The sixth and final element in McCrae and Costa's (1996) model is external influences, which they define as the psychological environment. Since this element is not given much discussion by McCrae and Costa, applying Bronfenbrenner's conceptualization of the environment as described in his bio-ecological model (e.g., Bronfenbrenner, 1979; Bronfenbrenner and Morris, 1998) is useful. Bronfenbrenner argues that the environment is a nested system comprised of three basic levels, the microsystem, the mesosystem, and the macrosystem, with the microsystem incorporating the most narrow components of the environment and the macrosystem incorporating the most broad, overarching components. Imposing this structure onto

the McCrae and Costa framework helps to generate examples of the environmental influences that could play important roles in a theory of personality.

At the macrosystem level, the environment is comprised of cultural and subcultural norms and values, as well as the *zeitgeist* in which individuals find themselves. An individual's race and even socioeconomic status can be an influence at the macrosystem level. Although elements in the macrosystem are quite distal and thus the least likely to be predictive of individual behavior, they are critical to a comprehensive understanding of the environment in which an individual functions.

At the mesosystem level are influences such as important relationships in a person's life (e.g., mother-child, siblings, coworkers and supervisors), or more generally, the various roles held by individuals. Kahn, Wolfe, Quinn, Snoek, and Rosenthal (1964) defined roles as expectations of appropriate behavior held by others. For example, a person holding the role of "wife" may be expected to behave differently than a person holding the role of "husband." Stereotypically, the wife role involves cleaning, cooking and performing other household chores while the husband role is that of the bread-winner. Behavior within a role is typically guided by social norms that dictate appropriate and acceptable behavior for that role. Importantly, a single person holds multiple roles (e.g., mother, wife, employee). As discussed by Katz and Kahn (1978), the nature of these roles can come into conflict, essentially meaning that the expectations of one role conflict with the expectations of another role. Individuals are often having to balance the demands of multiple roles in order to

minimize role conflict and reduce tension. Thus, understanding the roles held by an individual is key to understanding the environment in which they function.

Finally, the microsystem is comprised of influences that exist in a specific situation. These types of environmental influences include situational constraints, social cues, and reinforcement and punishment. These influences are the most proximal and exhibit the most direct influence on behavior in any particular situation.

The role of external influences in McCrae and Costa's (1996) framework helps to emphasize the distinction between basic tendencies and characteristic adaptations described earlier. As shown in Figure 1, characteristic adaptations are a highly interactive component in the model. Importantly, they influence objective biographies which then affect external influences, and external influences then reciprocally affect characteristic adaptations. Also, as indicated in the figure by the symbol in the top left corner of the characteristic adaptations box, there are dynamic processes occurring within the context of characteristic adaptations themselves, which speaks to their somewhat malleable nature. Basic tendencies, on the other hand, stand alone at the beginning of the model. Although they influence characteristic adaptations (and the self-concept), no other element in the model has an influence on basic tendencies. These are the enduring features that form the core of personality. Indeed, the definition of personality provided by McCrae and Costa is centered around personality traits, which they argue are the main component of one's basic tendencies: "Personality consists of a system defined by personality traits and the dynamic processes by which they affect the individual's psychological functioning" (p. 76).

McCrae and Costa (1996) present an overarching framework of personality within which they encourage researchers to generate both “grand” and midlevel theories of personality. McCrae and Costa suggest the FFM of personality is an example of one such grand theory that attempts to explain, broadly, the relationships between the elements in the larger framework. They also note that more specific mid-level theories of personality are needed to explain *how* certain elements are related to one another. Grand theories, such as the FFM, propose relationships between various elements in the meta-theoretical framework, but mid-level theories identify the psychological mechanisms that explain how two or more elements are related. The remainder of this chapter will discuss the relevance of the FFM as a grand theory, and then discuss the theory of conditional dispositions (Wright & Mischel, 1987) and the cognitive-affective system theory of personality (Mischel & Shoda, 1995) as mid-level theories well suited to addressing the research questions of interest in the present study.

The Five Factor Model of Personality

The FFM is rooted in the lexical hypothesis, first proposed by Galton (1884). The suggestion made by the lexical hypothesis is that personality traits are such a fundamental component of human interactions that any word needed to describe a person exists currently in language. The implication of this assumption is that if we want to understand personality traits, we need only examine language. Galton began a list of adjectives used to describe people, and this list was expanded by Allport and Odbert (1936). With the emergence of factor analysis in the early part of the twentieth

century, the efforts of many researchers (e.g., Cattell, 1943; Fiske, 1949; Tupes & Christal, 1961) were focused on refining this list and identifying underlying factors.

Decades later, general agreement has been reached on labels and definitions for a five factor model, and these are presented in Table 2. Hough and Ones (2001) conducted a comprehensive review of the FFM factors and reached several important conclusions. First, their review indicated that FFM is robust across different cultures, languages, types of assessment (e.g., questionnaires, interviews), rating sources (e.g., self, spouse, coworkers), and genders, and it is also replicable across different factor extraction and rotation methods. Hough and Ones concluded that of the five factors, Extraversion and Neuroticism were the most robust as they were replicated in nearly all studies, and that Conscientiousness is the next most robust. Less support was found for the replicability of Agreeableness, and the least amount of evidence regarding replicable factor structure was found for Openness to Experience. FFM researchers also agree that each of these broad factors, or domains, consists of more narrow facets. Personality facets can be defined as lower level traits that are relatively more narrow than broad domains, and that together generally reflect the entire scope of the broad domain (Costa & McCrae, 1995). Although there is still disagreement surrounding the nature of the facets (see Goldberg, 1997), the facet structure offered by Costa and McCrae for the NEO-PI-R personality measure is presented along with the domain labels and definitions in Table 2 because it appears to be the most heavily researched and replicated.

As mentioned earlier, the FFM can easily be situated as a grand theory of personality within McCrae and Costa's (1996) meta-theoretical framework of personality. Although McCrae and Costa present several postulates, only those with the greatest relevance to the proposed study will be discussed. With regard to basic tendencies, McCrae and Costa suggest that the FFM has the properties of individuality, development, and structure. Individuality refers to the idea that all adults can be characterized on the basis of the five factors, and that these factors influence thoughts, feelings, and behaviors. Development refers to the notion that the FFM traits continue to develop as a person ages and reach stability around age 30. Structure refers to the hierarchical nature of the factors, that is, that factors characterize broad dispositions while facets characterize dispositions more narrowly.

Regarding characteristic adaptations, McCrae and Costa (1996) suggest that the FFM has the properties of adaptation and plasticity. The concept of adaptation refers to the idea that people develop and evolve in a manner that is highly consistent with their personality traits. For example, extraverts surround themselves with people by joining clubs and various other organizations. The concept of plasticity states that characteristic adaptations (but *not* personality traits themselves) will not necessarily be stable over time because they are susceptible to changes in maturation and the environment. Also, characteristic adaptations can change as a result of making an intentional effort to change (e.g., by enrolling in psychotherapy).

With respect to objective biographies, McCrae and Costa (1996) describe the principle of multiple determination. Multiple determination refers to the idea that any

particular behavior is the result of each and every characteristic adaptation that is induced by the situation. As pointed out by Ahadi and Diener (1989) there is rarely a direct association between a single trait and a particular behavior. Rather, behavior can be instigated as a response to multiple trait-related drives. The example given by McCrae and Costa is the behavior of reading a book. Reading a book can satisfy an introversion-related need for privacy and also satisfy an openness-related need for intellectual stimulation. One important implication of this postulate is that it is basically impossible to infer the presence or absence of a trait on the basis of observing a single behavior.

McCrae and Costa (1996) suggest that a self-schema is an integral component of the self-concept. The idea of the self-schema is that individuals maintain a view of themselves that is influenced both by cognitive and affective appraisals of themselves and the world, and that this view resides in consciousness. In other words, individuals hold a view of themselves that they are capable of describing to others. Finally, McCrae and Costa (1996) propose that the principle of interaction, which suggests that the environment helps to shape characteristic adaptations themselves, and the environment also interacts with characteristic adaptations to regulate the flow of behavior.

Now that the FFM has been presented in its theoretical context, the next step is to discuss its relevance to the prediction of performance. Although the FFM is certainly relevant to other aspects of the work environment, the focus will be restricted to the use of FFM in employee selection.

Validity of the Five Factor Model in the Prediction of Performance

The FFM of personality is a very popular framework for examining individual differences in the field of industrial/organizational psychology. Prior to the emergence of the FFM, individual differences were often overlooked in organizational research. Not only were there a plethora of personality constructs, making it difficult to draw comparisons across studies, but the existing models and measures were based primarily upon theories of abnormal personality (e.g., the Minnesota Multiphasic Personality Inventory; Cascio & Aguinis, 2006). In order for a personality theory to be useful in organizational research, it would need to describe patterns of behavior common to the majority of individuals, rather than the minority. Thus, the FFM provided researchers with a much-needed and very useful taxonomy of normal personality. Barrick and Mount's (1991) meta-analysis affirmed the utility of the FFM in organizational research. In their study, the authors linked each of the five personality dimensions to important workplace outcomes including job proficiency, training proficiency, and personnel data (i.e., salary level, turnover, status change, and tenure). Averaging the mean effect sizes across these criteria yielded the following corrected validity estimates for each factor: Extraversion ($\rho = .13$), Neuroticism ($\rho = .08$), Agreeableness ($\rho = .07$), Conscientiousness ($\rho = .22$), and Openness to Experience ($\rho = .04$).

Since Barrick and Mount's (1991) meta-analysis, many studies in I/O psychology have examined the relevance of personality in the prediction of various outcomes, and there have been additional meta-analyses as well (e.g., Hough, Ones, &

Viswesvaran, 1998; Mount, Barrick, & Stewart, 1998; Pulakos & Schmitt, 1995).

Perhaps the most informative of all these studies was a meta-analysis conducted by Barrick, Mount, and Judge (2001) wherein the authors meta-analyzed the results from 15 independent meta-analyses in order to reach some general conclusions about the relevance of the five factors to various aspects of performance across different jobs.

Criteria in the Barrick et al. (2001) meta-analysis were divided into two general categories, work performance and specific occupations, following the general format adopted by Barrick and Mount (1991). The work performance category consisted of the following components: supervisor ratings, objective performance, training performance, and teamwork. The specific occupations category included sales performance, managerial performance, professionals, police, and skilled or semi-skilled workers. Results were most favorable for Conscientiousness, with estimated true correlations ranging from .19 to .26, indicating that Conscientiousness is a robust predictor of performance across different types of criteria and jobs. Results for Neuroticism (coded as Emotional Stability) indicated a significant estimated true correlation with overall work performance, although most estimated true correlations for specific criteria were near .10. A notable exception is the estimated true correlation for teamwork which was .22. Although it did not yield a significant estimated true score for overall performance, Extraversion also yielded favorable results for more specific criteria, with almost all estimated true correlations exceeding .10, and an estimated true correlation for training proficiency of .33. Agreeableness and Openness to Experience did not predict overall work performance, and most estimated true

correlations for specific criteria were weak (i.e., less than .15). Notable exceptions are the estimated true correlation between Agreeableness and teamwork (.34) and Openness to Experience and training proficiency (.33). Thus, the results indicate that Conscientiousness and Neuroticism are robust predictors of performance, with Conscientiousness being even more robust. And while Extraversion, Agreeableness, and Openness to Experience may not predict overall work performance, they definitely have relevance to more specific work criteria (e.g., teamwork, training proficiency).

Less research has been conducted regarding the validity of facet measures of the FFM, likely because the literature is still trying to come to consensus regarding the structure of the facets. Most research on facets has been focused on Conscientiousness, and a meta-analysis by Dudley, Orvis, Lebiecki, and Cortina (2006) summarizes these studies. Dudley et al. conceptualized Conscientiousness as consisting of four facets (achievement, order, dependability, and cautiousness). Using meta-analytic regression, the researchers were able to demonstrate that the facets contributed incremental variance to the prediction of several outcomes beyond that of global Conscientiousness. These outcomes include overall job performance, task performance, job dedication, interpersonal facilitation, and counterproductive work behaviors. Thus, with the exception of Conscientiousness, few studies have examined the validity of facet measures of the FFM.

Despite the supporting evidence of the utility of the FFM demonstrated by Barrick et al. (2001) and other meta-analyses, several researchers have criticized the

FFM on theoretical grounds. For example, Briggs (1989) argued that the five dimensions lack precise specification. Other researchers argue that five dimensions are not adequate in explaining the whole of personality. For example, Ashton et al. (2004) argued that a six factor structure is more appropriate. A central critique of the FFM seems to be that the dimensions may be too broad to accurately predict outcomes (Hough, 1992). Although there is little argument that global measures are more useful than facet measures in the prediction of overall performance, facet measures may maximize the predictive validity of specific performance criteria (Murphy & Lee, 1994).

In addition to criticisms of the FFM model specifically, some I/O researchers and practitioners have been critical of the use of all personality measures in an employee selection context (e.g., Morgeson et al., 2007). In a recent, highly controversial article, Morgeson and colleagues argue that personality tests generally exhibit extremely low validity as selection tools, and they suggest that the field as a whole reconsider their use. The authors are not necessarily arguing that the evaluation of personality ought to be removed from the selection process, but that we need to generate more valid ways of measuring personality (e.g., by using non self-report methods). Reactions to this article were strong, impassioned even, with responses by Tett and Christiansen (2007) and Ones, Dilchert, Viswesvaran, and Judge (2007) stating that the claims made in the original study were unfounded, and citing meta-analytic evidence demonstrating the utility of personality measures in a selection context.

The Morgeson et al. (2007) article and the responses to it are important for several reasons. First, the articles demonstrate that although most researchers and practitioners acknowledge that personality plays a role in performance, they all certainly do not agree on how or whether personality assessments should be used in selection. Second, it demonstrates that even meta-analytical evidence is not necessarily a gold standard by which we should be evaluating the utility of various selection tools. Rather, meta-analyses are one piece of information amongst many others that can be used to help make informed decisions. Finally, and perhaps most importantly, it shows that there is much work that can be done to improve the utility and validity of personality measures in employee selection. That is the one point on which both the authors of the original article and the authors of the responses agreed. Ones et al. (2007) suggested that collecting both observer and self-report ratings may be another way to improve the validity of personality measures. In the original Morgeson et al. article, Robert Dipboye argues that self-report methodology need not be abandoned entirely, rather, researchers just need to be more creative in how the information is collected. For example, he suggests that participants be allowed to elaborate on their responses in some way. The present studies offer an innovative approach to addressing this call for alternative self-report methodologies.

In one sense, the question of how to improve the validity of a measure is a psychometric one. The reliability of a scale places an upper limit on the validity coefficient. Specifically, the correlation between two measures cannot exceed the product of the square root of their reliabilities (McDonald, 1999). For example, the

maximum correlation that could be obtained between a predictor with reliability equal to .80 and a criterion with reliability equal to .85 is $r = .82$, or $\sqrt{.80}(\sqrt{.85})$. Only a perfectly reliable predictor and criterion could correlate at 1.00. With this knowledge, a reasonable approach to improving the validity of a measure would be to attempt to improve its reliability. Indeed, there is evidence that personality measures yield only modest reliability estimates. Internal consistency reliability estimates for the five factors measured by the NEO-PI-R, McCrae and Costa's (1992) measure of FFM, are described as ranging from .77 to .86, although many primary studies have found reliabilities as low as .60 (see Hough & Furnham, 2003). One common explanation for these low to modest validities centers around the idea that people may demonstrate inconsistencies in their behavior across situations. Given that personality is generally understood to reflect enduring tendencies and patterns of behavior, this explanation has led to a fair amount of concern and caused some to seriously question the use of personality measures (e.g., Mischel, 1968). Presented next are a discussion of these concerns and a theoretical framework that describes how cross situational variability in behavior is essential to the understanding of personality.

Cross Situational Variability and Stability of Personality

Mischel's (1968) strong critique of personality measures led to an almost immediate and relatively drastic decline in personality research in the decade or so after it was published (Hogan, 1991). In the critique, Mischel enumerated what he viewed as the many problems that existed in personality research at the time. Hogan provides a succinct summary of Mischel's critiques. The first of Mischel's main

complaints is that individual differences in behavior should be consistent across situations and across time, but he notes that the literature does not provide evidence of this consistency. The second of Mischel's critiques is that validities for personality measures are rarely larger than .30. Therefore, since personality accounts for such a small proportion of the variance in behavior, the situation must play a much more important role, and researchers should focus most of their attention on the situation.

Hogan's (1991) response to Mischel's first claim (i.e., that behavior should be consistent across situations and across time) is that it is too vague to be useful, given that Mischel does not clearly define what he means by "consistency." Hogan argues that in order to demonstrate true consistency in behavior across situations, people would actually have to change their behavior to fit the circumstances. He also argues that behavior does not need to be perfectly consistent across situations, rather, the goal is to demonstrate "functionally equivalent" behavior that will support making inferences about the existence of stable personality traits. A more fundamental point, however, is made by McCrae and Costa (1996), who argue that "sophisticated personality psychologists have never claimed that traits determine behavior independently of situational context" (p. 57). Indeed, this view is reflected in McCrae and Costa's meta-theoretical framework of personality and made apparent by the distinction between basic tendencies and characteristic adaptations (discussed earlier). Individuals do have personality traits that develop throughout young adulthood and stabilize around age 30, but the manner in which these traits are expressed vary as a

result of characteristic adaptations, which are subject to external influences in the environment.

With regard to Mischel's second critique, Hogan (1991) argues that many studies report correlations that exceed .30 (e.g., Hogan, DeSoto, & Solano, 1975). However, the meta-analytic evidence discussed earlier confirms that validities for personality measures *are* generally small to moderate in magnitude. Yet the conclusion that Mischel draws from this finding, that attention should be refocused primarily on the situation, is somewhat unwarranted. McCrae and Costa (1996) argue that although traits do not determine behavior independently of situations, they do occupy an important position in the explanation of behavior. The assertion that the situation is the primary driver of observable behavior is a reflection of the behaviorist paradigm that dominated theory and research in psychology at the time of Mischel's writing. A more moderate view of the impact of the situation was posited by Costa, McCrae, and Zonderman (1987) who argued that, just as personality cannot be expected to exclusively determine behavior, there are upper limits to the roles played by the environment in the determination of behavior.

CAPS Theory

Interestingly, since publishing his critique of personality research, Mischel has made some important contributions to the field's understanding of the person by situation interaction, and the role that each has to play in the determination of behavior. Wright and Mischel (1987) published a study that described the notion of conditional dispositions. Using a sample of children at a summer camp, Wright and

Mischel demonstrated that in situations (e.g., fishing) where the demand for a competency such as coping skills was low, children considered to be aggressive demonstrated few aggressive behaviors. However, in situations where competency demands were high (e.g., being denied something, having conflict with a peer) aggressive children demonstrated higher levels of aggressive behavior. Mischel and colleagues continued to build upon these initial findings based on the summer camp children, and their work eventually led to the development of a cognitive-affective system theory of personality (CAPS; Mischel & Shoda, 1995).

The CAPS theory was proposed in an attempt to reconcile the apparently contradictory findings in personality research that indicated that although personality was supposed to be stable, there also exists evidence of cross-situational variability in behavior. I will explain the basic tenets of the theory, and in doing so attempt to situate these elements into McCrae and Costa's (1996) meta-theoretical framework to promote greater understanding and cohesion between the two.

Mischel and Shoda (1995) point out that personality researchers have generally considered cross situational variance in behavior as noise or measurement error that detracts from being able to detect "true" behavioral tendencies across situations. However, Mischel and Shoda argue that differences in behavior across situations are not simply noise, but are characteristic patterns of behavior that are, in fact, reflections of an underlying personality structure. Furthermore, they argue that, with the appropriate amount of methodological rigor, these fluctuations in behavior across situations can be predicted.

The key elements in Mischel and Shoda's (1995) theory are referred to as *cognitive-affective mediating units*. These would be considered basic tendencies in McCrae and Costa's framework. Types of cognitive-affective mediating units include: encodings (constructs for the self, people, events, and situations), expectancies and beliefs (e.g., about the world, outcomes of behavior, and self-efficacy), affects (feelings, emotions, affective responses), goals and values, and competencies and self-regulatory plans (behaviors that one is capable of performing and plans for organizing action). Obviously, individuals can and do differ with respect to each of these units, and Mischel and Shoda argue that these units are what form the core of the person.

The next basic tenet of CAPS theory is that the cognitive-affective units interact with one another in ways that are unique to each individual. Thus, the organizational structure of the units differs across persons. These patterns of interactions among the units, or their organizational structure, can be considered characteristic adaptations in McCrae and Costa's (1996) language for at least two reasons. First, Mischel and Shoda (1995) argue that interactions among certain units will be critical to the formation of a self-concept. As discussed by McCrae and Costa, the self-concept is a particular kind of characteristic adaptation. Second, Mischel and Shoda note that patterns of units will become activated in response to situations. As indicated in McCrae and Costa's framework, external influences have a direct impact on characteristic adaptations.

The final central tenet of Mischel and Shoda's (1995) theory is that situations, depicted as external influences in McCrae and Costa's (1996) framework, have

predictable influences on the organizational structure of cognitive-affective mediating units (i.e., characteristic adaptations). That is, although various situations may result in differences in behavior across situations, these differences can be predicted and are stable in their own way. Mischel and Shoda argue that identifying the psychologically relevant features of situations will aid in efforts to predict behavior across situations.

A key point in Mischel and Shoda's (1995) theory is that by accurately specifying the situation, researchers will observe a greater degree of cross-situational stability in personality. The implication for attempting to reliably measure personality, then, is to provide an individual with a specific context in which they can think of themselves. By identifying the situation for the respondent and trying to reduce the number of contexts in which they are imagining themselves while responding, the reliability of the measure should improve, which would in turn (ideally) improve validity. This approach has been adopted by several researchers and practitioners in I/O psychology, and the results of these efforts are discussed in the next section.

Chapter 3

Frame of Reference Effects in Personality Measures

This chapter will first discuss the origin of the FOR and describe the first study conducted to test it (Schmit et al., 1995) in some detail. After this discussion will be a review of the studies that have examined the ability of a FOR to improve criterion-related validity, followed by a review of the findings regarding the psychometric properties of personality measures that use a FOR.

Origin of the FOR Effect

Within the personality testing research, a FOR refers to the environments or contexts in which individuals think of themselves when responding to a personality inventory (e.g., at work, at school, at home; Mount, Barrick, & Strauss, 1994). The notion that one's FOR could affect the validity of a personality test was first discussed by Mount, Barrick, and Strauss (1994). Mount et al. conducted a study with the intention of comparing the validity of observer-rated (e.g., supervisor, coworker, and customer) personality measures to the validity of self-rated personality measures. Results indicated that observer ratings contributed incremental variance, beyond that accounted for by self-ratings, to the prediction of supervisor-rated performance for measures of conscientiousness, extraversion, openness to experience, and agreeableness. Following Hogan (1991), Mount et al. argue that from an observer's perspective, personality is a reflection of how a person is perceived by others in a given context. That is, an observer's understanding of a target's personality is limited to the public self the target displays in a particular context. Personality from the

target's perspective, however, is a reflection of a multitude of "selves" both public and private, across a variety of situations, and may consist of thoughts, beliefs, and feelings that are not necessarily expressed or even observable. Thus, Mount et al. argued that observer ratings of personality may be more valid predictors of performance because they are restricted to work-related observations of behavior, resulting in a stronger predictor-criterion match than self ratings of personality. The researchers suggested that a simple way to test this hypothesis would be to ask participants to think of themselves at work when responding to a personality test, and test to see whether the validity of this personality test was higher than that of standard personality tests. This hypothesis was first tested by Schmit, Ryan, Stierwalt, and Powell (1995).

Schmit, Ryan, Stierwalt, and Powell (1995) conducted two studies aimed at understanding how frames of reference affect responses to personality tests. The researchers focused on how two factors might serve to create a FOR for participants. The first factor was the testing situation, which was manipulated by presenting participants either with instructions for a general personality test (general instructions) or with instructions indicating that they should respond to the questions as if they were applying for a customer service job (applicant instructions). The second factor was item contextualization, which was manipulated by adding the words "at work" to the personality items. The example given by the authors involved changing the statement, "I try to be courteous to everyone I meet" to "I try to be courteous to everyone I meet at work." This factor had just two levels, work-specific and non-contextual frames of

reference. Because this was a student sample, “work” referred to the job of being a student. It is worth noting that although the researchers used a Big Five measure of personality, the Openness to Experience factor was not included because the addition of “at work” to many of the items on this scale did not make sense.

In the first study, Schmit et al. (1995) designed two substudies, a between-subjects factorial design and a mixed factorial design. The between-subjects factorial design (substudy 1) randomly assigned participants into one of four conditions: general instructions – non-contextual items, general instructions – work-specific items, applicant instructions – non-contextual items, applicant instructions – work-specific items. The mixed factorial design (substudy 2) randomly assigned participants to either the general or applicant instructions condition (between-subjects factor), but all participants responded to both the work-specific and non-contextualized personality items (within-subjects factor). Results for substudy 1 indicated significant main effects, but no significant interactions, for both testing situation and item contextualization for measures of Conscientiousness, Agreeableness, and Neuroticism, such that means were higher for Conscientiousness and Agreeableness in the applicant instructions and work-specific items conditions, and means were lower for Neuroticism in these conditions. Results from substudy 2 were largely consistent with substudy 1, with the addition of significant testing situation by item contextualization interactions for Conscientiousness and Neuroticism. As expected, the interactions indicated that factor scores were the most favorable (i.e., highest for Conscientiousness and lowest for Neuroticism) in the applicant instructions – work-

specific items condition and least favorable (i.e., lowest for Conscientiousness and highest for Neuroticism) in the general instructions – non-contextualized items condition.

Schmit et al. (1995) also conducted multiple group confirmatory factor analysis using the data from the two substudies by first comparing the work-specific and non-contextualized items in the general instructions condition, and then comparing the two item types in the applicant instructions condition. Both comparisons showed strong evidence for psychometric equivalence, yielding invariant structure, invariant latent-factor correlations, and invariant factor loadings across the two item types. Only error variances were found to differ across the two measures in both conditions, with the error variances being larger for the non-contextualized items. This first study by Schmit et al. is important because it demonstrates that significant mean differences can result depending on one's FOR, and it also provides at least preliminary evidence that the factor structure of the personality measure did not appear to differ substantially based on item contextualization.

The second study conducted by Schmit et al. (1995) was a criterion-related validity study aimed at determining whether a FOR would improve the validity of a personality measure. The researchers randomly assigned college students to each of the four conditions described in Study 1. Participants were administered a facet-level measure of Conscientiousness, and their college GPA was used as the criterion. Results for global Conscientiousness indicated that the highest validity coefficient was found for the applicant instructions – work-specific items condition, followed by the

general instructions – work-specific items condition, general instructions – non-contextualized items condition, and finally the applicant instructions – non-contextualized items condition. The results for the facets of Conscientiousness were largely consistent with those for the global measure. Also, using ANOVAs and confirmatory factor analysis, the authors were able to replicate the results from Study 1.

The Schmit et al. (1995) studies are important for several reasons. First, they were the first to examine the FOR effect in personality tests. Second, the results demonstrated that mean personality scale scores differ significantly based on the FOR provided, and these results were replicated across two samples. Third, the study provided evidence that criterion-related validities are enhanced by providing subjects with a FOR. Finally, the researchers were able to demonstrate that the psychometric properties and factor structure of a personality scale did not differ based on whether a FOR was provided. These last two contributions, enhancing criterion-related validity by using FORs and evaluating the psychometric properties of FOR scales, have been expanded on by other FOR researchers. The next two sections will provide a review of these studies.

FOR and Criterion-Related Validity

Following in the tradition of Schmit et al. (1995), Bing, Whanger, Davison, and VanHook (2004) attempted to address some of the limitations of the original study. Bing et al. collected data on a facet measure of Conscientiousness using a student sample and replicated the 2 X 2 between-subjects design from Schmit et al.

Bing et al.'s results indicated a different rank ordering of the four conditions than found by Schmit et al., with the highest criterion validity coefficient obtained in the general instructions – contextualized items condition, followed by applicant instructions – contextualized items condition, the general instructions – contextualized items condition, and finally the applicant instructions – non-contextualized item condition. The primary difference between these findings across the two studies is that the applicant instructions produced higher validities in the Schmit et al. study, whereas the general instructions produced higher validities in the Bing et al. study (though the differences in validities between the applicant and general instructions for both the contextualized and non-contextualized items were not significant). Consistent with the findings of Schmit et al., the validities of the contextualized items were higher for both applicant and general instructions than were the non-contextualized items.

Bing et al. (2004) were interested in addressing two other remaining concerns based on Schmit et al. (1995). First, the researchers included a measure of cognitive ability in order to test the possibility that the enhanced validities Schmit et al. found for the contextualized personality items may have been spurious. In other words, Bing et al. suggested that a correlation between cognitive ability and the contextualized measure of Conscientiousness would be an alternative explanation for the enhanced validities in the work specific items condition. Results indicated that contextualized items did not have a higher correlation with cognitive ability under either instruction condition, thus eliminating at least one explanation of spurious findings. Secondly, the researchers wanted to test the incremental validity of the contextualized personality

items over the non-contextualized personality items. Regression results indicated that contextualized Conscientiousness demonstrated incremental validity beyond cognitive ability and non-contextualized personality scores for both the general and applicant instruction conditions.

Bing et al. (2004) is an important contribution to the FOR literature because it replicated, to a large extent, the results found by Schmit et al. (1995). Bing et al. also included a measure of cognitive ability in order to address the possibility that Schmit et al.'s findings may have been spurious, and in doing so removed a viable alternative explanation for the FOR effect in the Schmit et al. study. Finally, and perhaps most importantly, Bing et al. demonstrated that a contextualized personality test has incremental validity above both cognitive ability and a standard, non-contextualized personality inventory.

Although the studies conducted by Schmit et al. (1995) and Bing et al. (2004) made several significant contributions, the primary limitation of these studies was that they were conducted using student samples. Hunthausen, Truxillo, Bauer, and Hammer (2003) addressed this limitation by testing the FOR effect in a sample of entry-level customer service managers for a major airline. In addition to the type of sample, another difference between the two studies is that Hunthausen et al. chose to contextualize the personality measures by placing the "at work" FOR in the instructions, rather than within the items themselves. This allowed them to overcome the awkwardness that Schmit et al. faced when trying to attach "at work" to the Openness to Experience items, and they were able to include this scale in the study.

Hunthausen et al. also included cognitive ability and test-taking motivation as control variables. Since cognitive ability (*g*) has been consistently shown to be a strong predictor of job performance (Hunter & Hunter, 1984; Pearlman, Schmidt, & Hunter, 1980; Schmidt & Hunter, 1998), other performance predictors are often evaluated in light of their ability to contribute incremental variance to the prediction of performance beyond *g*. Test-taking motivation was collected because it is assumed that incumbent samples are generally less motivated than applicant samples. Although test taking-motivation was reasonably high in the total sample, the means of test-taking motivation differed slightly across the non-contextualized and “at work” FOR conditions, resulting in it being used as a control variable.

Hunthausen et al.’s (2003) results provide further support for the use of a specific FOR in personality tests. Comparisons of validity coefficients in the non-contextualized versus “at work” FORs indicated significantly higher validities for Conscientiousness, Extraversion, and Openness to Experience, but not for Neuroticism or Agreeableness. In addition, FOR condition was tested as a moderator of the relationship between each of the personality factors and job performance. When considered as a group, the interaction terms contributed incremental variance to the prediction of performance beyond the control variables (i.e., test-taking motivation, job tenure, gender), cognitive ability, and main effects. When considering the interaction terms independently, only the betas for the FOR by Extraversion and FOR by Openness to Experience were significant, indicating that criterion-related validities for these factors were higher in the “at work” FOR condition than in the non-

contextualized condition. Additional hierarchical regressions for each condition indicated that in the non-contextualized condition, Extraversion, Openness to Experience, and Conscientiousness did not show incremental validity beyond cognitive ability and test taking motivation, but the ΔR^2 for the group of predictors and the individual betas were significant in the “at work” FOR condition. Thus, the three factors contributed incremental validity beyond cognitive ability and test taking motivation in the “at work” FOR condition, but not in the non-contextualized condition.

The Hunthausen et al. (2003) study was a very informative extension of the Schmit et al. (1995) study. First, the results indicated that the FOR effect does exist in an incumbent sample outside of the laboratory. Secondly, the study showed that the FOR effect produces incremental validity beyond cognitive ability. Finally, the results provide an interesting comparison to those of Schmit et al., particularly with regard to Extraversion. Schmit et al. did not find significant mean differences in Extraversion scores across item type (work-specific versus non-contextualized) in either sample from Study 1. The researchers suggest that measures of Extraversion may not enjoy the enhanced criterion-related validities associated with a FOR to the same extent as measures of Conscientiousness, for example, because an Extraversion scale may generalize better across situations. In other words, they suggest that Extraversion may be less susceptible to FOR effects because it may have greater cross-situational stability. Yet in the Hunthausen et al. study, it was Extraversion that demonstrated the larger FOR effect, while the FOR effect was not demonstrated for Conscientiousness.

Interestingly, Hunthausen et al. posit a similar alternative as that described by Schmit et al., suggesting that it is perhaps Conscientiousness that is less susceptible to FOR effects. Given that it was suggested by both groups of researchers as an explanation for non-significant findings, the idea that certain latent traits may be more or less susceptible to FOR effects deserves further attention. A closer examination of the FORs used across the five factors when responding to personality inventories may shed some light on this issue.

A handful of other studies have been able to provide further evidence of the FOR effect. DeGroot and Klumper (2007) used a sample of retail store associates to examine whether providing FOR instructions would yield incremental validity beyond that of a situational interview. Results indicated that scores on “at work” measures of Extraversion, Conscientiousness, and Agreeableness accounted for an additional 7% of the variance in job performance. Criterion-related validities were also significantly higher for individuals who complete “at work” measures of Extraversion and Conscientiousness, although no significant difference was found for Agreeableness.

Small and Diefendorff (2006) conducted a very interesting study with undergraduates employed in a variety of positions. Participants provided responses on both non-contextualized and contextualized personality measures, and ratings of personality, task performance, and organizational citizenship behaviors (OCBs) were provided by both coworkers and supervisors. Results indicated significantly higher means for contextualized versus non-contextualized measures of Conscientiousness, Extraversion, and Emotional Stability (i.e., Neuroticism). Further, although

contextualized Conscientiousness and Extraversion were significantly related to task performance and OCBs, respectively, they did not show incremental validity in the prediction of these outcomes beyond that accounted for by general self-ratings. Emotional Stability did show incremental validity in the prediction of OCBs, however. These results are in contradiction to those of Bing et al. (2004) who were able to demonstrate that a contextualized measure of Conscientiousness had incremental validity beyond a general measure of Conscientiousness. Small and Diefendorff suggest that the putting the FOR manipulation in the instructions may not have made it salient enough, which would explain why their results differ from those of Bing et al. who inserted the “at work” statement in each item. Although Hunthausen et al. (2003) were able to demonstrate the FOR effect by manipulating FOR in the instructions, their between-subjects design did not allow them to test whether the contextualized personality measure had incremental validity beyond the non-contextualized measure.

This review of studies focused on demonstrating the criterion-related validity of contextualized personality measures has indicated that it is likely that contextualized personality measures can provide incremental variance in the prediction of performance beyond that of non-contextualized personality measures. More specifically, the studies by Schmit et al. (1995) and Bing et al. (2004) demonstrate that the validity of facet measures of Conscientiousness can be improved by contextualizing a measure. However, the results of Small and Diefendorff (2006) and Hunthausen et al. (2003) are not as favorable for Conscientiousness. With regard to the other factors, Schmit et al. found that Extraversion did not benefit from an FOR,

whereas Hunthausen et al. found that it did. Hunthausen et al., Small and Diefendorff, and Schmit et al. showed that a FOR can benefit measures of Neuroticism, but only Hunthausen et al. were able to demonstrate similar positive effects for Openness to Experience. Given these somewhat divergent findings, it would be desirable for future studies examining FORs to more closely examine the potential reasons for the differences in these findings.

FOR and the Psychometric Properties of the Big Five

As discussed previously, Schmit et al. (1995) evaluated the psychometric properties of the four Big Five factors utilized in their study across the four conditions and found strong evidence for measurement equivalence across the conditions. Since then, researchers have revisited the idea of measurement equivalence and conducted additional research to explore the ways in which the psychometric properties of the Big Five may or may not be altered by providing respondents with a FOR. These studies are reviewed below.

Smith, Hanges, and Dickson (2001) were interested in testing the idea that asking a person to respond to a personality test as an applicant versus responding in a generic context would alter the factor structure of the FFM. Using three samples (job incumbents, job applicants, and students) Smith et al. compared the factor structure using a multi-group confirmatory factor analysis. Results indicated support for measurement equivalence in the FFM across the three samples.

Holtz, Ployhart, and Dominguez (2005) used an organizational justice framework to examine whether providing a FOR and validity information for a test

would improve test-taker reactions. They utilized a 2 (test format: non-contextualized versus “at work” items) by 2 (validity information: non-contextualized versus job-related) between subjects design, and a series of multiple group confirmatory factor analyses were conducted to evaluate whether the error variances in the personality measures differed significantly across the four conditions. Similar to the results found by Schmit et al. (1995), error variances were not invariant across the four groups, with the largest errors being in the non-contextualized test format conditions. Further analysis suggested that factor variances were also largest when the test format was non-contextualized. Holtz et al. suggest that, given this pattern of error and factor variances, using a FOR should improve the criterion-related validity of personality measures by reducing error in prediction.

In addition to demonstrating the incremental validity of FOR personality measures beyond non-contextualized personality measures, Bing et al. (2004) were interested in exploring mechanisms by which FORs might improve validity. Bing et al. acknowledge the argument made by other FOR researchers, like Holtz and colleagues, concerning the decrease measurement error as one mechanism, but Bing et al. also argue that contextualizing items may result in greater predictor-criterion match. To test this hypothesis, Bing et al. re-analyzed the Schmit et al. (1995) data by constructing a general linear model with the validity of the Conscientiousness facets as the dependent variable and scale reliability as a covariate. The independent variables of instructions (general versus applicant) and item type (non-contextualized versus FOR) were still significant, and Bing et al. interpreted this result to mean that the

higher validities seen in the FOR condition were not solely due to increases in reliability, thus allowing room for additional explanatory mechanisms of the FOR effect. The researchers then replicated these analyses with their own data with the inclusion of cognitive ability as a covariate and obtained similar results.

Robie, Born, and Schmit (2001) undertook an interesting re-analysis of the Schmit et al. (1995) data by imposing a generalizability theory framework. Using generalizability theory (Shavelson & Webb, 1991) allowed the researchers to obtain estimates of the amount of variance in personality test responses that is due to the person (i.e., the actual respondent filling out the measure), the situation (i.e., “in general” or “at work”), the interaction between the person and the situation, and measurement error. The researchers expected to see a significant main effect for person, since people are likely to differ from another in the responses they provide. They also expected a significant situation main effect, with scale scores being highest for the “at work” responses. A significant person by situation interaction was also expected, meaning that the researchers anticipated differences in the extent to which individuals’ responses would vary across situations. The researchers anticipated that responses for the “in general” and “at work” conditions would be most different for the Neuroticism scale, followed by Conscientiousness and Agreeableness. They expected that responses would be most similar for the “in general” and “at work” conditions for the Extraversion scale. This hypothesis was based on the idea that the factors having the highest correlations with social desirability should be more apt to

change across environments, so the hypothesized order reflects the rank ordering of those correlations (see Ones, Viswesvaran, & Reiss, 1996).

Robie et al. (2001) found significant main effects for person, situation, and significant person by situation interactions in the hypothesized direction for all factors. With regard to the size of the person by situation interactions, the hypothesized order was not supported. The largest person by situation interaction was found for Neuroticism, followed by Extraversion, Conscientiousness, and Agreeableness. Exploratory analyses were also conducted to compare the results from the applicant and general instructions conditions. Interestingly, the person by situation interactions remained significant in the applicant instructions condition. This finding was somewhat unexpected by the researchers because they anticipated that asking respondents to respond as though they were applying for a job would induce a FOR-type effect. However, given that Schmit et al. (1995) found that contextualizing the items had a greater effect on item responses than did the instruction type, and remembering that these are the same data, this result should not be too surprising. Overall, the presentation of Schmit et al.'s data in a generalizability theory context is a useful reconfiguration because it highlights, and provides evidence for, the idea that a FOR may function differently across personality factors, and that there are individual differences in the extent to which people vary their behavior across situations.

Robie, Schmit, Ryan, and Zickar (2000) conducted a very thorough investigation of the measurement equivalence of the facets of Conscientiousness using both contextualized and non-contextualized items. One noteworthy aspect of this study

is that it was the first to examine contextualized items in an applicant sample. A second noteworthy aspect is that the researchers attempted to obtain a greater understanding of why contextualized measures might differ in their psychometric properties from non-contextualized measures by assessing the bandwidth of each item. In an initial study, the researchers administered a 48-item facet measure of Conscientiousness to a sample of undergraduate students and asked them to rate the context specificity of each item, ranging from extremely narrow to extremely broad. Researchers then standardized the responses within subjects using standard scores to account for the fact that some respondents would be more likely to consistently rate items as more broad or more narrow. The researchers then conducted multiple-group confirmatory factor analyses at both the item and facet levels and attempted to discern whether the context specificity rating might account for differences seen in the factor structures across the contextualized and non-contextualized groups.

Results from the multiple-group confirmatory factor analysis at the facet level indicated that factor loadings were equivalent across groups, but error variances were not, with the exception of the Achievement Striving facet. For the remaining five facets, error variances were higher for the group who completed the non-contextualized personality measure. This finding is consistent with the analysis conducted by Schmit et al. (1995). Regarding the context specificity ratings, non-contextualized items were rated as more generic than contextualized items, as would be expected. Results from the item-level multiple-group CFAs were more complicated. The first thing to note is that the single factor, uncorrelated errors model

that the researchers fit to five of the six facets did not fit well for the Deliberation facet. The researchers continued conducting equivalence tests for this facet, but note that the results should be interpreted cautiously. Factor loadings were equivalent across four of the six facets, and approximately half of the items in each facet demonstrated nonequivalent error variances across the two groups. For the Competence, Dutifulness, and Self-Discipline facets, error variances were higher in the non-contextualized group, but for the Order and Achievement Striving facets, error variances were not consistently higher for one group over another. Unfortunately, the context specificity ratings failed to shed any light on the pattern of error variances. Ideally, the ratings would have been higher (indicating more generic items) for items that had larger error variances. This was sometimes the case, but did not happen frequently enough to provide convincing evidence of the ratings as an explanatory mechanism.

The Robie et al. (2000) study is important for several reasons. First, it takes an item-level approach to the question of measurement equivalence in contextualized and non-contextualized measures. This is a significant contribution to the literature as it is unrealistic to assume that attempts to contextualize result in the same effect for all items. Second, even though it was ultimately uninformative regarding the issues it set out to examine, the notion of gathering information about the context specificity of an item is an important step in the right direction. What is really known about how broadly or narrowly respondents view personality items, on both contextualized and non-contextualized inventories? Is bandwidth solely a property of the item, or is it a

function of the person, or some combination of the two? Robie et al. acknowledge that there are within-person tendencies toward rating items as more or less context-specific, and this is why they standardized these ratings within individuals before creating an average specificity score for each item. One logical extension of this line of thought is, if there is within person variance in the extent to which people rate items as more or less context specific, there should also be within person variability in the extent to which people actually think of themselves in more or fewer contexts when responding to a personality measure. That is, a single person may think of themselves in only one context when responding to a given item, but for a different item they might think of themselves in a different context, or in multiple contexts. The idea of within-person variability in responding to contextualized and non-contextualized personality measures was explored in depth by Lievens, De Corte, and Schollaert (2008).

Lievens et al. (2008) challenged the traditional explanation that providing a FOR increases criterion-related validity by reducing between-person variability in responding, which in turn improves reliability. Lievens et al. posed an alternative explanation based largely on the fact that Cronbach's alpha is only slightly impacted by between-person variability, assuming that respondents are consistent in the FOR that they use, and that the reliabilities of the FORs used by different respondents are more or less the same. Thus, reductions in between-person variability would not explain increases in reliability. Within-person inconsistency, on the other hand, should have a considerable effect on reliability, which would in turn affect validity. This

might be best demonstrated with an example. Let us say a respondent completes a non-contextualized ten-item measure of Conscientiousness. For the first five items, the respondent thinks of themselves at work, and for the second five items, the respondent thinks of themselves at school. Cronbach's alpha for the ten-item scale will be based upon the covariance matrix for all 10 items. However, this use of alpha is inappropriate, because the first five items are really measuring something like "at work Conscientiousness" while the last five items are measuring something like "at school Conscientiousness." Lumping all ten items together should result in a lower reliability estimate, all else equal, than computing reliabilities for the different FORs. The lowered reliability of this full ten-item scale would then result in a lower validity for the ten-item scale than for the two, five-item scales considered separately.

Lievens et al. (2008) conducted two studies to test the hypothesis that within-person inconsistency affects both the reliability and validity of a contextualized personality measure. The first study was a between-subjects design with three conditions ("at work" FOR, "at school" FOR, no FOR) that demonstrated that reducing between-person variability by providing a FOR did not affect reliability. Reliabilities for the FOR scales were not significantly higher than that of the non-contextualized scales, and this result supports the alternative explanation posed by Lievens et al.: reliability is not affected by between-person variability as long as a respondent is consistent in the FOR that they use and that the various FORs have similar reliabilities. This first study also demonstrated that although reliabilities did not differ across the conditions, validities did, and this shows that (as suggested by

Bing et al., 2004) reducing between-person variability cannot be the only mechanism by which FORs improve validity. The last finding from the first study is that the reduction in between-person variability that occurred as a result of the FORs was only beneficial for the respondents who used the FOR (i.e., “at school”) that conceptually overlapped with the criterion (i.e., GPA).

The second study was a within-subjects design where participants responded to a broad measure of Conscientiousness and two facet measures of Conscientiousness using “at school” and “at work” FORs. The researchers were able to simulate within-person inconsistency in this design by randomly drawing samples that varied in the number of items that were rated using a specific FOR. For example, in one sample 90% of the items might have been rated using an “at school” FOR and the remaining 10% were rated using an “at work” FOR, while in another sample 60% of the items were rated using an “at work” FOR and 40% were rated using an “at school” FOR. Samples that approached the 50/50 mark, meaning that half of the items were rated using the “at school” FOR and the other half were rated using the “at work” FOR, represented the highest degree of within-person inconsistency. After drawing 1,000 of these kinds of samples and computing the validity coefficients for each, Lievens et al. concluded validities were highest when a single, “correct” FOR (meaning in this case, at school) was used to rate a large number of items. These results provide additional support for the idea that within-person consistency has substantial effects on reliability, and that validity is strongly impacted by a predictor that has conceptual overlap with the criterion.

The results from the Lievens et al. (2008) study represent several important contributions to the FOR literature. First, the studies show that FORs function by reducing within-person inconsistency in responding, rather than reducing between-person variability. Regarding validity, the results indicate that reducing both between-person variability and within-person inconsistency will result in increased validity. Notably, though, the researchers also demonstrated that the FOR must conceptually overlap with the criterion. Indeed, they were able to demonstrate that using an “incorrect” FOR (i.e., one that does not overlap with the criterion) actually decreases validity.

The Present Studies

Although the Lievens et al. (2008) study, along with the rest of the FOR literature that has been reviewed, have demonstrated that the reliability and validity of personality measures can be improved by providing respondents with additional context, this stream of research may have charged ahead without the proper footing. Virtually all of the FOR studies are variants on the same basic design: ask respondents to complete a non-contextualized personality inventory, ask respondents to complete a contextualized personality inventory, then compare the psychometric properties and validities of the two conditions. FOR researchers have put a good deal of effort into understanding the psychometric properties and important elements of the contextualized inventories, while to some extent overlooking the fact that “non-contextualized” inventories are not certainly not context-free. The present studies seek to obtain a greater understanding of the context(s) that are implied by a non-

contextualized personality inventory, and in doing so further inform the use of personality measures in research and practice. The two studies utilize qualitative and quantitative methods to begin addressing questions surrounding the situations/contexts that are implied by non-contextualized personality inventories.

Chapter 4

Development of Research Questions for Study 1

What happens as a person completes a non-contextualized personality inventory? What information are they relying upon to inform their responses to each item? As suggested by the literature reviewed thus far, context plays an important role in the process. As noted by Johns (2006), the word “context” can take on a variety of meanings, depending of course, on the context. The same issue applies to the term “context effect,” which is defined in the measurement literature as “any influence or interpretation that an item may acquire purely as a result of its relationship to the other items making up a specific test” (Wainer & Kiely, 1987, p. 187). Many studies examining context effects concern themselves primarily with the order in which items or stimuli are presented. Although the examination of this type of context effect is not the focus of the studies in this dissertation, a brief review of this literature is provided in the interest of presenting a complete picture of the various elements that influence responses to personality items.

Harrison and McLaughlin (1993) describe how context effects of this variety are operationalized in various research areas. For example, studies in cognitive psychology demonstrate that the presentation of certain stimuli can prime respondents to activate certain response processes or retrieve certain types of information (e.g., Tversky, 1977). These findings have been logically extended to the performance appraisal literature (e.g., Kravitz & Balzer, 1992) where it has been shown that ratings can differ based on the order in which raters (e.g., supervisors) rate multiple targets

(e.g., subordinates). Assimilation effects, for example, occur when ratings of a particular target tend to follow those of the previous targets, whereas contrast effects occur when ratings of a particular target are highly divergent from those of the previous targets. In a study of context effects in work attitude measures, Harrison and McLaughlin found that placing items containing neutral words (e.g., “impartial,” “fair,” “objective”) in a block of positive items resulted in positive responses to the neutral items, whereas placing the same items in a block of negative items resulted in negative responses to the neutral items. Knowles (1988) examined context effects in four personality measures and found that items appearing at the higher end of a scale tended to have higher correlations with the rest of the test items than items appearing at the beginning of a scale. Knowles’ explanation of how self-schemas influenced this finding has relevance to the proposed study.

Knowles (1988) argues that respondents are continually accessing a self-schema in order to inform their responses to personality items (Higgins, King, & Marvin, 1982). As a respondent progresses through a list of personality items, this self-schema becomes more accessible and consistent. Thus, respondents are evaluating items that appear at the end of a scale on the basis of a self-schema that is more available to them than it was when they were responding to items at the beginning of a scale, thereby producing greater consistency in responses as they progress through the personality measure. Indeed, the FOR literature, particularly the results found by Lievens et al. (2008) seems to support the idea that holding a consistent self-schema has beneficial effects on the reliability and validity of personality measures.

Hogan (1991) also discusses the role of the self-schema in responding to personality measures. He argues that a person's view of themselves is influenced both by their social reputation and their own "inner nature" (p. 875). Social reputation refers to the way in which a person is viewed by others in their lives (e.g., friends, family, coworkers). As Hogan points out, this concept of the self is a public one. One's inner nature, however, is a much more private self. One's inner nature cannot be directly observed, rather, it refers to the internal structures of one's personality that cause certain behavioral tendencies. Both of these understandings of the self will influence the responses that a person provides to a personality test. In responding to each item, the person will consider both the public image of they believe they portray, and the private self that remains set away from the outside world, in order to form a view of themselves that they will rely upon when completing the measure.

One question that stems from Knowles's (1988) findings and Hogan's (1991) postulations, however, is *how* the self-schema is developed as a person progresses through a personality measure. Self-schemas are analogous to the self-concept described in McCrae and Costa's (1996) meta-theoretical framework. As described earlier, self-concepts are a specific kind of characteristic adaptation and, as such, are affected by external influences. CAPS theory (Mischel & Shoda, 1995) also supports this view of the self as being in part determined by external influences. As described by McCrae and Costa, and Mischel and Shoda, this process is occurring at a relatively macro-level, that is, the system is functioning across situations and other relatively broad contexts as it continues to shape and define one's general pattern of behavior.

However, as previously described, external influences also exist at the level of the mesosystem (e.g., the roles an individual holds) and the microsystem (e.g., situational constraints and social cues). Thus, according to the model presented by McCrae and Costa, environmental influences such as the roles held by individuals affect the development of the self-concept. In theory, then, individuals should rely on these various environmental influences while responding to a personality inventory. This assumption is inherent in the FOR literature, but it has yet to be explicitly examined.

Research Question 1: Do people think of themselves in particular environments or contexts when responding to a non-contextualized personality inventory? In other words, do individuals use FORs when responding to a non-contextualized personality inventory?

If the assumption being explored in Research Question 1 is shown to be true, the next logical step would be to obtain an understanding of those contexts.

Research Question 2: In which contexts do people see themselves when responding to a non-contextualized personality inventory?

Study 1 addresses these research questions by utilizing verbal protocol analysis methods (Ericsson & Simon, 1993) that encourage participants to think aloud while performing a task. An overview of this study and its results is presented in the next chapter.

Chapter 5

Study 1

Overview and Rationale

As described in the review of FOR research, a number of quantitative studies have examined how a single FOR, provided by the researchers, may improve the criterion-related validity of a personality measure and yield smaller error variances than a non-contextualized personality measure (e.g., Bing et al., 2004; Schmit et al., 1995). However, in these studies, the FOR is manipulated by the researchers. Further, the decision about how the FOR was chosen is typically given little or no explanation. In most cases, it seems to be based on assumptions by the researchers of the relevance of a particular FOR (e.g., “at work”) to the sample being studied. Although this strategy is by no means unreasonable, it is an experimental manipulation based on implicit assumptions about how FORs are utilized by respondents when completing a personality measure. These assumptions that have not been sufficiently explored. The most basic assumption is that individuals do, in fact, use FORs in some way when responding to non-contextualized personality measures. A second assumption is that the use of FORs in non-contextualized measures differs across individuals. Another assumption, particularly illustrated by the Lieven’s et al. (2008) study, is that a single individual uses more than one FOR when responding to a non-contextualized inventory. Despite the fact that these assumptions form the very core of FOR research, they have never been explored.

The primary goal of the Study 1 was to explore the assumption that individuals make use of FORs when responding to non-contextualized personality inventories. The method chosen to address this issue is a qualitative technique referred to as Verbal Protocol Analysis (VPA; Ericsson & Simon, 1993). VPA studies, also known as “think-aloud” studies, involve asking participants to verbally describe their mental processes while completing a task. There are many examples of VPA studies in other disciplines, but the only VPA study I am aware of in personality research is by Robie, Brown, and Beaty (2007). Robie et al. (2007) used think aloud methods to garner insights into faking behavior in personality tests. This methodology is influenced both by the approach taken by Robie et al. and by recommendations from Ericsson and Simon.

VPA methodology was chosen for Study 1 because it allows for the analysis of individuals’ “live” responses to a personality measure. In VPA studies, the role of the researcher is that of passive observer. There is no attempt by the researcher to influence responses. Their role is only to encourage the participants to verbalize their thought processes. By following this method and allowing participants to freely express what comes to their minds while completing a personality measure, a relatively unobstructed view of the thought processes involved in responding to a personality inventory should be revealed.

Although VPA is believed to be the most appropriate method for exploring the assumption of the use of FORs and for addressing the first research question of the current studies, it is not without criticism. For example, some have likened think-aloud

methods to the introspective techniques that have been all but abandoned since the early 20th century. VPA has its roots in the introspective method (Pritchard, 1990) but has undergone various transformations over the past 100 years that render it a distinct methodology. For example, in the late 1800s Titchener used introspection in order to study the “elemental ingredients of [human] experience” (Hergenhahn, 2001, p. 242). Titchener’s subjects had to be highly trained with regard to which types of information they were to report. How subjects should be trained and what information should be reported was often a point of contention among researchers using introspective techniques (Hergenhahn). VPA, on the other hand, does not require training. No training is required because the goal is for the subject to simply express their thought processes out loud. Further, introspective techniques required subjects to describe past experiences that were essentially memories of an event or stimulus. In VPA studies, verbal reports are collected from subjects as the thought processes are taking place, which eliminates the need to rely on memories and, as such, avoids the possibility of retrospective bias. It is worth noting that some VPA studies have used retrospective, as opposed to concurrent, verbal reports, and the retrospective method has been found to yield less useful data (Kuusela & Paul, 2000). Thus, despite the similarities between VPA and introspective techniques, the VPA method is distinct in ways that allow it to produce more meaningful and valid data.

Another limitation of the VPA method is that requiring a participant to provide a verbal report may change the information that the participant attends to, and it may also use up some of the cognitive resources needed to complete a task (Wilson &

Schooler, 1991). If either of these were to take place, it could change the way that participants respond to the task at hand. However, these potential limitations may be diminished by allowing the participant to take as much time as necessary when responding and by minimizing any demand characteristics that may cause them to alter their natural behavior.

Despite these potential limitations, VPA methodology was determined to be the most appropriate for Study 1 because it allows for the exploration of the thought processes that are involved in completing a personality measure. The method served to examine commonly held, but not yet explored, assumptions about the use of FORs when responding to personality measures, and it did so as unobtrusively as possible in the interest of obtaining the most authentic data.

Method

Determining Necessary Sample Size

Determining sample size *a priori* in a qualitative study is somewhat difficult and can be quite arbitrary. According to Morse (1991), the concept of *saturation* dictates when data collection ends. Saturation is defined as the adequacy of data and is operationalized as the act of collecting data until no new information is obtained (Morse, 1995). However, there are no guidelines or objective tests of saturation akin to power analysis in quantitative research. Moreover, Morse (1995) argues that it is not the quantity of data that has the most bearing on saturation, but the quality or the richness of the data. Although Morse emphasizes that there is no precise formula for predicting when saturation will occur, she does discuss at least two strategies that

might accelerate the process. The first suggestion is to select a sample that is both culturally cohesive and likely to demonstrate consistency with regard to the research topic. But cohesive samples will of course limit the generalizability of the results. The second suggestion is to employ a purposeful (AKA, theoretical) sampling strategy, which is described next. When selecting the sample for Study 1, I balanced considerations of cohesiveness and generalizability, and I utilized a purposeful sampling strategy in order to improve the likelihood of achieving saturation.

Morse (1991) describes four sampling strategies for qualitative research. The sampling strategy for the pilot study was a combination of two strategies that Morse refers to as *volunteer sampling* and *purposeful sampling*. Volunteer sampling was the primary strategy, as all participants were undergraduate students who volunteered to participate in the study in exchange for extra credit. A purposeful sampling strategy involves selecting participants based on the needs of the study. For Study 1, it was important that all participants had a sufficient amount of work experience and that they had actually applied for a sufficient number of jobs. It was also highly desirable to obtain as diverse a sample as possible in terms of gender, ethnicity, and age. Thus, an initial screening was conducted in an attempt to obtain participants with an adequate amount of work experience, as well as a reasonable amount of diversity. Although a demographically diverse sample is less likely to be cohesive, it seemed reasonable to assume that the sample would contain some degree of cohesiveness given that they were being recruited from the same department within the same university, and in some cases even the same classes.

As stated earlier, it is impossible to determine sample size for a qualitative study *a priori*. However, a consideration of pragmatics and available resources is a good place to start. In addition, although Study 1 is a critical component of the dissertation as a whole, it is not the exclusive focus and, as such, the data collection phase could not continue indefinitely. Thus, in conjunction with my committee, I decided that a sample size of 20-30 individuals would be acceptable, barring any additional unforeseen issues that arose during data collection. In particular, it was agreed that if 20-30 individuals were not enough to achieve a reasonable degree of saturation, data collection would continue.

Participant Recruitment, Screening and Selection

Participants were recruited from two undergraduate psychology courses. Recruitment involved giving a brief description of the study, along with its intended benefits and any potential consequences of participation. Potential participants were provided with a link to a short online screening survey where they were asked to provide their demographic information (e.g., gender, age, ethnicity, work experience, number of jobs held; see Appendix A) along with an email address. In the survey, potential participants were also asked to indicate from a list of provided times when they would be available to participate in an interview. Potential participants had one week to complete this short screening survey. A total of 65 individuals completed the screening survey, resulting in a 65% response rate. Of these 65 individuals, 30 were selected to participate in interviews and the remaining 35 were given an alternative assignment to complete in order to have the opportunity to earn extra credit.

The selection process for the interviews began by eliminating individuals from the pool who did not have any work experience ($n = 1$). Descriptive statistics were then calculated in order to determine the demographic make-up of the sample. Sixty-eight percent of the sample were female and the average age was 25.73 ($SD = 6.76$). The majority of the sample (82%) identified themselves as White, while the remaining 18% of the sample was comprised of Blacks (3%), Native Americans (6%), Hispanics (9%), and Asians (9%). Also, the majority of the sample (85%) were born in the United States. Sixty-two percent of the sample were currently employed and worked an average of 20.62 hours per week ($SD = 12.78$). Of these individuals, 18% held more than one job. The total pool ($N = 64$) had an average of 7.73 years of work experience ($SD = 6.86$). The majority of the sample (69%) were single, had not been previously married, and were not responsible for the care of minors or elders. Thus, a typical demographic profile of an individual drawn from this sample would be a single White female in her early to mid twenties with a part time job and no children.

The goal of using the purposeful sampling strategy described above was to try to ensure that the participants who were selected for the interviews did not all have the same demographic profile, but rather that the final sample be sufficiently diverse. The exact determination of “sufficiently diverse” is difficult to specify, but the approach I used was to determine each individual’s demographic profile, relative to the other individuals in the sample, and to try to select the outliers. The demographic profiles for each individual were determined as follows. The continuous variables (age, hours worked per week, years of work experience, number of minors for which you are

responsible, hours spent giving care to elders per week) were standardized and individuals with a z-score greater than zero, indicating that they fell above the mean of the variable in question, were given a value of “1” for that variable, and all others were given a value of “0.” For the categorical variables (e.g., ethnicity, whether born in the United States, marital status, whether previously married) the ‘majority’ response was identified. For example, the majority response for ethnicity was White. Individuals who selected the majority response were given a value of “0” and those who selected any of the remaining responses were given a value of “1.” In general, the recoded values were meant to identify those individuals who were most dissimilar from the overall demographic profile of the participant pool. However, individuals who indicated that they were currently employed were given a value of “1” for that variable because it was desirable to have employed people serve as participants. The recoded values for each variable were then summed to create an overall ‘diversity score’, which had a maximum of ten and a minimum of zero (actual scores for the sample ranged from zero to eight). Individuals were then rank-ordered on the basis of this score with the intention of scheduling interviews with 35 of them. Thirty individuals were easily selected for interviews using this method, but the thirty-first through the forty-sixth individuals all had diversity scores of two. At this point I simply chose the remaining five individuals based on the order in which they had completed the screening survey.

Once this sample of 35 individuals was identified, emails were sent to them asking if they were able to participate in an interview at a specified date and time.

Emails were sent on a Thursday afternoon and participants were asked to respond by Monday evening. Thirty of the 35 selected individuals responded to confirm their availability. My intention was to go back to the participant pool and select individuals to fill these unoccupied slots, but none of the remaining individuals had indicated that they were available during the scheduled time slots. Due to these scheduling conflicts, a total of 30 interviews were scheduled. Only two participants failed to arrive for their interviews.

The demographic make-up of the final sample of 28 individuals was 64% female with an average age of 29.64 years ($SD = 8.33$). The majority of the sample (71%) identified themselves as White, while the remainder of the sample was comprised of Blacks (7%), Native Americans (7%), Hispanics (10%), and Asians (13%); 25% of the sample were born in the United States. Sixty percent of the sample were currently employed and worked an average of 19.30 hours per week ($SD = 12.13$). The sample had an average of 10.92 years of work experience ($SD = 8.73$). Exactly half of the sample (50%) were single, but of those that were single, 35% had been previously married and had children. Finally, the majority of the sample (65%) had either child or elder care responsibilities, and 14% had both child and elder care responsibilities.

Interviewer Training and Materials

Two research assistants and myself served as interviewers. In order to prepare the research assistants, a two-hour training session was conducted. During the training session, the interviewers became familiar with all of the materials that they would

need to use during the interviews. These materials included: an audio mp3 recorder, informed consents, demographic information provided by each participant, the personality inventory, follow-up questions to the personality inventory, and a “cheat sheet” for the personality inventory. The cheat sheet (Appendix B) was a tool I prepared that defined some of the words in the personality inventory that might be unknown or confusing to participants (e.g., ‘shirk’). After conducting the first nine interviews, the cheat sheet was refined for the benefit of the research assistants. The research assistants were trained on how to use the audio recorder, how to administer the informed consents, and how they should behave during the interview. Specifically, a primary component of the training included instruction in the VPA methodology with an emphasis on making sure participants were thinking aloud. The research assistants were also instructed *not* to interpret the items for the participants, but to provide definitions of words or colloquial phrases if necessary (e.g., “get by”). After they became comfortable with the materials, mock interviews were conducted, one with each research assistant, where I played the role of participant. The final component of the training was to observe an actual interview that I conducted.

Procedure

Upon the participant’s arrival, the participant and the interviewer entered the room where the interview was to be conducted. This was a private space with a closed door so that disturbances and interruptions would be avoided. The script that the interviewer read to participants is provided in Appendix C. Participants were first asked to read and sign the informed consent. Then, participants were asked to review a

summary of the demographic information they provided during the screening survey and confirm the accuracy of this information. Next, the interviewer provided the participants with the personality inventory. The script includes the standard instructions that accompany the personality measure chosen for use in this study, supplemented by the instructions specific to the think aloud methodology. Once the participant understood the instructions and what they were being asked to do, the digital recorder was turned on and the interview began. After completing the personality inventory, participants were asked three follow-up questions (see Appendix C). Once the interview was completed, participants were thanked and reminded that they would be receiving extra credit for their participation. The average interview lasted 21 min, with the shortest taking just 7 min, and the longest lasting 48 min.

Measures

Participants completed the International Personality Item Pool (IPIP: Goldberg et al., 2006) personality measure modeled after the FFM. The measure (see Appendix D) consists of a total of 60 items, with ten items measuring each factor. There are five negatively worded and five positively worded items for each factor. Reliabilities for the factors are reported by Goldberg et al. as follows: Extraversion ($\alpha = .86$), Openness to Experience ($\alpha = .82$), Neuroticism ($\alpha = .86$), Conscientiousness ($\alpha = .81$), and Agreeableness ($\alpha = .77$). Participants are asked to indicate how descriptive each item is of them using a 7 point Likert scale ranging from 1 (*very inaccurate*) to 7 (*very accurate*).

Analytical Approach

I listened to the audio recording of each interview three times. The first round consisted of listening to each interview without taking any notes. The purpose of this was to become familiar with the nature of the interviews and obtain a holistic sense of whether the assumption that people use frames of reference when responding to non-contextualized inventories is valid. The act of listening during this first round was not particularly detail-oriented and was not focused on identifying the exact frames of reference that were being used (that was the goal of rounds two and three). I listened to each interview all the way through without pausing in order to try to understand what the experience had been like for both the interviewer and the participant.

The second round of analysis revolved around detailed note-taking. Each interview was essentially dissected, with the recording being paused dozens of times in order to note the participant's responses. The notes included both my paraphrasing of the participant's comments, as well as verbatim participant comments when it seemed more appropriate. Once the second round of analysis was complete, I compared the notes across all of the interviews with the goal of establishing a kind of coding scheme. When a working coding scheme had been established, I began the third and final round of analysis, which involved coding participant responses according to the established scheme. A second rater also coded participant responses and interrater agreement statistics were calculated.

Results

Outcomes of First Round of Analysis

As described above, one of the main goals of the first round of analysis was to get a sense for the nature and flow of the interviews. Overall, each interview went as expected. In general, participants seemed very comfortable and at ease. Although they would occasionally request feedback (e.g., “Is this what you’re looking for?” “Am I doing this right?”) for the most part they followed the think aloud methodology and provided candid explanations for their responses. A key point, however, is that the think aloud method did not come naturally to the majority of participants. This was evidenced by the frequency with which interviewers had to prompt them to think aloud. Furthermore, it became clear during the first few interviews that stringent enforcement of the think aloud methodology was not comfortable for the interviewer or the participant. Repeatedly asking a participant to “tell me what you’re thinking” after every single unexplained response felt strained, and thus the interviewers would generally only prompt after participants had responded to two or three consecutive questions without an explanation.

A key outcome of the first round of analysis was to address Research Question 1 and explore the assumption that individuals rely on context when completing non-contextualized personality inventories. The brief answer is a resounding yes. Every single participant indicated that they were considering context when choosing their answers to the personality inventory. In addition, the nature of the contexts used when responding varied within individuals, indicating that a single individual was using multiple FORs. Moreover, on average participants considered contextual information when responding to 22 of the 50 items, although it should be noted that the relatively

large standard deviation (15) around this mean indicates that there is potentially a significant level of individual variation in the extent to which context influences responding.

Outcomes of Second Round of Analysis

The goal of the second round of analysis was to address Research Question 2 and begin to identify the FORs used by individuals when responding to the non-contextualized personality inventory. In order to do this, I created a relatively structured coding scheme by which the FORs used by participants could be compared and grouped together. This analysis yielded 9 categories that are described in more detail below and summarized in Table 3. One important point to note is that the categories are not, for the most part, mutually exclusive. That is, participants could have used (and did use) more than one FOR when responding to a single item. The exception here is the “no elaboration” category. If a response was coded as no elaboration, no other categories would have accompanied it.

No elaboration. This aptly named category indicates that participants did not elaborate on their reasons for choosing a particular response option. If a participant did not voluntarily elaborate on an item and was also not prompted by the interviewer to do so, a rating of no elaboration was given.

Non-contextual. Non-contextual classifications indicate that some or all of the explanation for an item did not include a reference to a specific FOR. Some examples of a non-contextual response include, “That’s just the way I am,” “I was brought up to believe that...,” and “I played a lot of sports as a kid and it taught me that...”. A

number of different types of responses populated this relatively broad classification. Participants frequently referred to upbringing (e.g., “I was raised in a military family where being neat was really important”), world-views and religious beliefs (e.g., “Treating others with respect is an important part of my faith”), and values (“Having a career is really important to me”).

Specific situation. Occasionally, participants described themselves in a relatively specific situations when responding to an item. Examples here include, “when I’m driving,” and “when I’m on vacation/traveling.”

School FOR. When participants referred to themselves “at school,” “in class,” or “at PSU” while responding to an item, it was categorized as a school FOR.

Work FOR. When participants described themselves as “at work,” “on the clock,” “dealing with customers,” or interacting with their bosses or coworkers while responding to an item, the work FOR category was applied.

Family FOR. The family FOR category was applied when participants described themselves “with my family,” “with my kids,” or “with my parents.” The distinction between the family and home FORs is less apparent for people who live with their families, given that the one is so highly associated with the other. However, the distinction becomes greater, and perhaps more important, for those who do not live with family.

Friends FOR. Anytime participants described themselves “with my friends” the item was categorized with the friends FOR.

Home FOR. If participants described themselves “at home,” “doing housework/chores,” or “watching TV” when answering a question, the home FOR category was used.

Significant other FOR. When participants described themselves with their spouse, boyfriend/girlfriend, partner, or fiancée, the significant other category was applied. Although this category was not applied frequently, the fact that participants specifically described the way they behaved when in the presence of their significant other seemed to justify considering it as a separate category. The other categories that it may have been lumped into (family or friends) seemed inappropriate depending on the role of the significant other. In other words, lumping a spouse into the friends FOR seemed as inappropriate, as did lumping boyfriend/girlfriend into the family FOR.

Outcomes of the Third Round of Analysis

Once the categories and coding scheme had been established, I coded all of the interview responses according to the established scheme. Afterwards, a research assistant served as a second coder. In order to allow for inaccuracies in the first coding scheme, the research assistant was instructed to code any responses of which he was unsure as Other. In order to estimate reliability of ratings between the two raters, Cohen’s kappa was computed for each of the 10 categories within an item. Thus, a total of 500 kappas were calculated (10 kappas per item X 50 items). This extremely high number of estimates was required because the categories within an item were not mutually exclusive. If the categories *were* mutually exclusive, only 50 kappas (one per item) would have been required because the two raters would only have been able to

choose to indicate one of the ten categories. However, because participants could and did use more than one category in responding, it was necessary to evaluate whether or not the raters agreed regarding the use of each category within each item.

The kappas ranged from .65 to 1.00, with 96% of the kappas having a value of 1.00. Three percent of the kappas had a value of .65, and this resulted exclusively from differences in coding of the Other category. Specifically, rater 2 used the Other category and rater 1 did not, so differences here were anticipated. I reviewed the responses coded as Other first to determine whether I had neglected to include any situations in the first two rounds of coding and determined that was not the case. The occurrences in which rater 2 had chosen Other could all be fit into the existing categories. The remaining 1% of kappas had a value of .76, and upon reviewing these inconsistencies it was easily determined that they reflected errors in the coding process (e.g., accidentally coding School when the participant had clearly mentioned work and not school). Thus, all disagreements between raters were evaluated and resolved.

The final round of analysis involved coding each response according to the scheme. There are many ways to present the data that resulted from this coding process, but ultimately I opted to arrange the findings by first determining the frequency of the categories within each interview participant, and then averaging these results across participants. Analyses focusing on the frequency of FORs within each factor were also conducted. The results for each FOR are reported in two ways, using two different denominators. The first method involves the use of a denominator that reflects the sum of all coded responses, including the non-contextual, no elaboration,

and all FOR categories. The second method uses a denominator that reflects the sum of *only the coded responses that included a FOR*. In other words, responses that were coded as non-contextual or no elaboration were not included in this denominator. The goal of reporting results using these two methods was to demonstrate 1) how frequently particularly contextual responses are used in relation to all other responses (both contextual and non-contextual), and 2) to demonstrate the relative use of each context *when context is mentioned*.

Analyses by FOR

Non-contextual. Across all participants, approximately 61% of the responses were coded as Non-contextual, meaning that participants did not appear to be relying on a FOR when responding, but rather were thinking of themselves as they are generally. This figured ranged from 23% to 87% across respondents indicating a considerable amount of individual variation in the tendency to rely on a general image of oneself when responding. Of the five factors, the conscientiousness and extraversion items tended to receive the highest number of Non-contextual designations.

No elaboration. Across participants, approximately 16% of the responses to items were coded as No Elaboration. This number varied pretty dramatically across participants, with two participants receiving 0 no elaboration codes, meaning that they thought aloud for each of the 50 items, and one particularly reticent participant receiving 37 No Elaboration codes. The number of No Elaboration codes seemed to vary both as a function of the individual participant, as some were simply less vocal

than others and did not respond as well to the prompting, and of the interview conditions. For example, the last interview on days where the interviewer had more than two interviews scheduled tended to have a high number of No Elaboration codes, indicating that the interviewer simply was not prompting the participant to think aloud, potentially due to fatigue.

Family FOR. Approximately 3% of all responses were coded as utilizing a Family FOR. When considering just the contextual responses, the Family FOR accounted for approximately 12% of responses. The use of this FOR was relatively consistent across individuals, though it tended to be more predominant for participants with children. In addition, an interesting and somewhat unexpected finding was that participants who repeatedly used a Family FOR almost never used a Friend FOR, and the same is true of the reverse (i.e., those who frequently used a Friend FOR almost never used a Family FOR).

Friends FOR. Approximately 3% of the responses were coded with a Friends FOR. When considering just the contextual responses, the Friends FOR accounted for approximately 12% of responses. The frequency of use for this particular FOR varied somewhat across individuals, and seemed to occur more frequently with younger participants. In addition, there were 2 participants that relied almost exclusively on the Friends FOR when responding to items, a phenomenon not replicated with any other FOR.

Home FOR. Approximately 3% of all of the coded responses were considered Home FOR. When considering just the contextual responses, the Home FOR

accounted for approximately 12% of responses. I had originally assumed that there would be a substantial degree of overlap between usage of Home FOR and Family FOR, with participants tending to use one or the other, but this assumption was not borne out. Participants routinely used both Home and Family FORs, indicating that these are two psychologically distinct situations.

School FOR. Approximately 6% of all responses, on average, were coded as utilizing a School FOR. When considering just the contextual responses, the School FOR accounted for approximately 24% of responses. This should not come as a surprise given that the sample was comprised of college students and took place on a college campus. Only two participants did not use a School FOR in responding to one or more of the items, and for a few participants it accounted for more than 15% of their total responses and more than 60% of their contextual responses.

Work FOR. Approximately 5% of all responses were coded as using a Work FOR. When considering just the contextual responses, the Work FOR accounted for approximately 20% of responses. This figure varied to a great extent across individuals with approximately one-third of participants not using it all while it accounted for more than 40% of the contextual responses for 3 participants. Although this FOR was not as universal as the school FOR, it was still predominant for most participants.

Significant other FOR. As mentioned earlier, the Significant Other FOR was used infrequently (it only accounted for only 1% of all responses and 4% of the contextual responses), but it seemed to be a very important FOR for the few

individuals that used it. In fact, for those individuals, the Significant Other FOR accounted for approximately 20% of their contextual responses.

Specific situation. This category made up approximately 4% of all responses and 16% of contextual responses. The tendency to describe oneself in terms of more specific situations varied across individuals, with some participants not describing themselves in terms of any specific situations, while others referred to 3 or 4 specific situations across the items. One important point to keep in mind when interpreting these results is that each instance of a participant mentioning a specific situation was coded as 1. This means that a participant who received a total of 4 Specific Situation codes could have mentioned 4 different specific situations, or they could have mentioned the same specific situation 4 times. The latter occurred much more frequently than the former. For example, one participant repeatedly considered their behavior in different cultures when responding to items. Thus, the data indicate that this participant relied on specific situations more frequently than the other participants, but the same specific situation was used throughout the interview.

Analyses by Factor

Table 4 presents the frequency of FOR responses within each factor. For all five factors, the Non-contextual code is the most frequent, followed by the No Elaboration code. This is expected based on the analyses by FOR presented earlier. Therefore, in these analyses by factor, I will only discuss the Specific Situation, Work, School, Family, Friends, Significant Other, and Home FORs.

Agreeableness. The Family FOR appeared the most frequently for the agreeableness items, followed by Friends, Work, Home, Specific Situations, School, and Significant Other. The relatively higher occurrence of the Family and Friends FORs might have emerged because these are the situations in which people may see themselves as behaving in ways that are particularly illustrative of their general agreeableness tendencies, or in ways that contradict their general agreeableness tendencies. For example, when responding to the item “Respect others,” one participant indicated that they were generally respectful, but not when interacting with their parents, while another participant indicated that they were especially respectful when interacting with their grandparents. Thus the Family and Friends FORs may have been mentioned more frequently because these are situations in which the greatest range of agreeable behaviors occur. In contrast, the School FOR may not have been mentioned as frequently because the school environment is not typically one that would generate extreme displays of agreeable or disagreeable behaviors. The relatively low occurrence of the Home FOR also makes sense because agreeable behaviors often require interaction (e.g., Have a good word for everyone) and not all people experience a great deal of interaction at home, especially students who either live by themselves or with a roommate.

Conscientiousness. The School FOR was used most frequently in responding to conscientiousness items, followed by Work, Home, Family Specific Situations, Friends, and Significant Other. Given that participants were students and that conscientious behaviors in an academic context are typically discretionary, high

frequencies for the School FOR are logical. It also makes sense that conscientiousness items elicited Work FORs because of the relevance of conscientious behaviors in that context. The occurrence of the Home FOR for conscientiousness might seem somewhat less likely, but when examining the actual items (e.g., Get chores done right away, Waste my time) it becomes obvious how the content would lend itself to thinking of oneself at home.

Extraversion. The Friend and School FORs were mentioned with the greatest frequency for the extraversion items, followed by Specific Situations (e.g., at a bar, at a party), Work, Family, Significant Other, and Home. As with agreeableness, extraverted behaviors typically imply interaction, and thus it makes sense that the Home FOR was mentioned least frequently, while the Friend and School FORs were mentioned most frequently. In addition, extraverted behaviors are appropriate both while with one's friends and while at school, and both situations allow for fluctuations in extraverted behavior. The work context is similar in that fluctuations in extraverted behavior are expected, particularly when considering the nature of the job, but it may have been endorsed less frequently than school simply because most participants were full time students with part time jobs, rather than part time students with full time jobs.

Neuroticism. There did not appear to be meaningful differences between the frequencies of FORs within neuroticism. While the No Elaboration code accounted for approximately 16% of all responses, it accounted for 24% of the responses within neuroticism. Thus, with 57% of neuroticism responses accounted for by the Non-contextual category and 24% accounted for by No Elaboration, the remaining 19% of

responses were spread relatively evenly across the FORs. The sensitive nature of the neuroticism items (e.g., Often feel blue, Dislike myself) may have caused participants to want to say less about them. In fact, due to the sensitive nature of the items interviewers were coached to not pressure participants into elaborating on their responses so as to reduce potential discomfort experienced by the participants.

Openness to experience. The School FOR appeared with the greatest frequency for the openness to experience items. In fact, it appeared more than twice as often as any other FOR. Following the School FOR were the Specific Situation, Family, Work, Friends, Home, and Significant Other FORs. The frequent occurrence of the School FOR is expected not only because the participants are students, but because the nature of several of the items seems to imply an academic context (e.g., Am interested in hearing new ideas, Avoid philosophical discussions). Moreover, the university setting is where students are exposed to a range of different types of people, often quite different from themselves. The occurrence of the Specific Situation FOR is interesting in the context of openness to experience because certain items basically provide a context (e.g., Do not enjoy going to art museums, Tend to vote for liberal political candidates). Thus, the specific situations mentioned in relation to the openness to experience items were typically contexts implied by the item (e.g., No, I disagree because when I'm at an art museum, I usually enjoy it).

Discussion

The goal of Study 1 was to explore the assumption that individuals rely on context when responding to non-contextual personality inventories, and the results of

the Study indicated that assumption is valid. Participants in Study 1 stated that they were thinking of themselves in the following contexts: at work, at home, at school, with friends, with family, and with significant others. In addition, participants also relied on the use of more specific contexts (e.g., “in my car while driving”). Previous FOR studies (e.g., Hunthausen et al., 2003; Lievens et al., 2008) have focused almost exclusively on Work and School FORs, but Study 1 demonstrates that other contexts such as home and family play a role in the response process for non-contextualized inventories.

It is also important to note that although it became clear that FORs were integral to respondents’ thought processes, participants also relied on non-contextual explanations for their responses. In other words, participants appeared to be thinking of themselves as they generally behave, as well as how they behave in certain contexts, when responding to items. This finding should not be surprising because it is these more general tendencies and patterns of behavior that most personality inventories are designed to measure. High frequencies for the non-contextual category, then, would be expected. Yet given that most personality inventories are intended to be non-contextual, the fact that FORs were mentioned in approximately one-quarter of all responses is particularly important.

Study 1 also demonstrates that the way participants use the aforementioned FORs in responding to non-contextual personality inventories is far more complex than any of the published research has indicated. Not only did participants rely on FORs beyond those considered in previous studies (i.e., work and school) across

items, but they used multiple FORs in responding to a single item. This is a phenomenon that has not even been considered in previous research. In addition, Study 1 demonstrates that there is notable variation in the ways in which FORs are utilized across the Big Five factors. For example, when comparing the use of the Friend FOR across factors, it appeared the most frequently for the extraversion and agreeableness items, whereas the Home FOR appeared the most frequently for the conscientiousness items than it did for other factors. A more detailed analysis of these patterns should be addressed by future studies.

Saturation

As discussed earlier, Morse (1995) defines saturation as the adequacy of data and is operationalized as the act of collecting data until no new information is obtained. Using this operationalization, I have confidence that saturation was achieved *for this population* of working students. In fact, after the tenth participant (out of 28) no new broad FORs were generated. In other words, by the tenth interview the categories of Work, School, Home, Family, Friends, and Significant Other had all been used by participants. The remaining 18 interviews did contain occasional examples of more specific contexts, but the broad categories were established early-on in the data collection. A logical explanation for this finding is that the contexts that have been identified are the most salient for this population and, consequently, appeared early and often.

Morse (1995) also points out that the mere *quantity* of data is not enough to guarantee saturation, but rather that the *quality* of the data is the key determinant of

saturation. The interview data certainly exemplified this point. A handful of participants seemed to be quite reticent and did not adapt well to the VPA methodology. Even when prompted by the interviewer, this group of participants gave very little elaboration on their responses. This behavior then influenced the interviewer such that he or she was less likely to continue to prompt the participant to elaborate. Even with interview data from dozens of these such participants, I doubt that saturation could have been reached. On the other hand, after completing the analysis it possible to look back at the data and identify a subset of interviews (approximately 12) that were rich and informative and would have provided all the data that was needed. Given these two extremes I feel confident that, even with a limited sample size of 28, saturation was achieved for this population.

Lessons Learned

I learned a number of things about how the process of responding to personality inventories while conducting these interviews and analyzing the data that I feel are important to address, even though they are not specifically tied to any research questions or gaps in the literature. This section contains these lessons learned.

Study 1 provided interesting insights regarding the process used by participants in identifying an item as situational. Participants would typically begin by following the instructions provided to them (i.e., Please consider yourself as you generally are). They would say things such as, “Yes/No, for the most part, I...” or “Well, I was raised to believe that...” or “Yes/No, I like to think that I...” At this point one of three things would happen: 1) They would decide that the item was a reasonably

accurate/inaccurate description of their behavior and move onto the next; 2) they would say things such as, “For example, when I’m with my friends/at work/at school I...” or; 3) they would pause and say things such as, “But sometimes I...” or “Except when...” or “But it depends on whether...” Most often these hesitations would occur when they thought of a specific instance or situation in which their behavior was/is discordant with how they generally think of themselves. This is an interesting finding because it implies that not only do participants think of particular contexts when responding to certain items because perhaps that situation is a salient example of their behavior, but also because they may behave differently in specific situations than how they would generally describe themselves. This is an important point because it implies that process of *using* a FOR when responding is somewhat complex. For example, if I am responding to the item, “Have a good word for everyone” and I mention my family, it could mean one of four things: 1) Generally I have a good word for everyone, and this is *especially true* when I am with my family, 2) Generally I have a good word for everyone, but this is *not always the case* when I am with my family, 3) Generally I *do not* have a good word for everyone, and this is *especially true* when I am with my family, 4) Generally I *do not* have a good word for everyone, but this is *not always the case* when I am with my family.

Perhaps the most interesting lesson in relation to the goal of Study 1 is how participants responded once they identified an item as situational. That is, once the thought process began to involve the use of one or more FORs, the response process often became very interesting. For example, when responding to the agreeableness

item “Have a good word for everyone” some respondents stated that while they generally try to have a good word for everyone, when they were with friends they often found themselves gossiping or saying mean things about people. Although they did not seem to think the “Very Accurate” response was appropriate, they were torn regarding which option was in fact the most accurate description. Another example involves the extraversion item “Keep in the background.” One participant struggled over this item because their job required them to be noticed and get people to pay attention to them. However, the person felt that outside of work, this was an accurate description. The participant went so far as to try and weight the appropriate response based on the amount of time spent at work. In other words, they reasoned that if they spent 25% of their total waking hours at work where the appropriate response would be “Very Accurate” and the rest of their time in situations where the appropriate response would be “Very Inaccurate” they should probably choose “Somewhat Inaccurate.” However, this level of analysis did not occur with other participants. In general, if participants found themselves torn between how they behave in one FOR versus how they behave in general, or how they behave in one FOR versus another FOR, they would almost without exception choose the midpoint of the scale. This implies that in many cases choosing the neutral response option is a potentially a reflection of the fact that a participant’s behavior may vary across contexts.

Another important lesson is that negatively worded items are *very* confusing for participants. This is not by any means a new or revolutionary finding, in fact many researchers avoid using negatively worded items at all because of the additional

cognitive load they place on respondents. However, it was eye-opening for me to see it firsthand. In almost every interview, a participant would encounter a negatively worded item (e.g., “Rarely get irritated”), decide that the item was inaccurate, and then proceed to select the “Very Accurate” response option. This seemed to occur most frequently for negatively worded items following positively worded items.

Occasionally, participants would notice their own mistake and go back and correct it, but often they did not. This behavior was particularly striking because, in theory at least, the interviewer had the full, undivided attention of the participant. The participants should have been, and seemed to be, completely focused on the items.

And yet even given this level of attention, they were still tripped up by the negatively worded items. If this is the case with highly attentive participants, it must be assumed that in cases where participants may not be as focused on the instrument (e.g., when completing it online) negatively worded items present an even greater problem.

Another important lesson that I learned related specifically to the emotional stability/neuroticism items (Appendix D). While responding to these items, a handful of participants indicated that they had been treated for depression and, as a result, did not believe many of the items to be an accurate description of their current behavior, although they may have been before they sought treatment. One participant was even visibly upset as she recalled her past experience. Although the IPIP is not a diagnostic tool, it cannot be ignored that these types of questions do apparently elicit thought processes in individuals that could result in negative emotions or feelings of distress. This is often listed as a potential consequence on informed consent forms, but actually

witnessing it reinforced that it is not an issue to be taken lightly. This also reinforces the important of voluntary responding; if providing an answer to an item is going to cause distress, it is important that participants not be required to do so.

Limitations

Although Study 1 made important contributions to the literature by exploring the assumption that respondents rely upon FORs when completing non-contextualized personality inventories, and by identifying a range of FORs used, it also has limitations. The primary limitation may be in the VPA methodology itself. As described earlier, VPA has been criticized for being too similar to introspection and because it has the potential to change the information that participants attend to when completing a task. However, these critiques were minimized in the present study by having untrained participants provide real time descriptions of their thought processes, and by minimizing demand characteristics. Participants were simply asked to explain their thought process and the only demand placed on them was to verbalize their thoughts.

Another limitation of the current study was that there was a learning curve for the interviewers (including myself). The most significant challenge encountered by the interviewers was continuing to ask participants to elaborate on their experiences. At times, not asking for elaboration seemed like the appropriate action. For example, if participants did not voluntarily respond to the item “Seldom feel blue” interviewers would not often press them to elaborate given the sensitive nature of the question. At other times, the interviewer simply missed an opportunity to ask for elaboration. Once

the research assistants began to conduct interviews, I listened to each interview and sent feedback as quickly as possible in order to encourage them to ask for elaboration more frequently. However, I did not advise them to insist on elaboration for every item as this would have likely caused more of a disruption to the thought process of the individual and made the interview more uncomfortable for them. Although it would have been desirable to have more elaboration occur naturally, I am satisfied that an appropriate balance was struck between obtaining information from participants and keeping them at ease during the interview with this approach.

Another limitation involving the interviewer learning curve involved the role of the interviewer in helping respondents interpret items. In an effort to make the process of completing the personality inventory using the think aloud methodology as similar as possible to completing it on paper, interviewers were instructed not to interpret items for participants. They were allowed to provide definitions for certain words or colloquial phrases if participants did not know them (e.g., very few participants knew the definition of the word “shirk”) but they were not allowed to provide an interpretation of the overall item. In one incident the respondent was unsure of the meaning of the phrase “get by” as in, “You do just enough to get by.” This was the first (and only) occasion where a participant asked for clarification of this phrase, so a definition had not been supplied on the cheat sheet. In attempting to provide a definition, the interviewer used an example of figuring out what was necessary to get a “C” in a course, and only do that amount of work. Obviously, the participant’s response to this item could not be considered because the interviewer

contaminated it by providing a FOR for the individual. However, occasional mistakes made by interviewers are an inevitable part of qualitative research, and because none of the mistakes were recurring they do not constitute significant limitations.

Finally, Study 1 considered a sample of working students. Thus, it must be assumed that while the results of this study will hopefully generalize to other working students, they may not generalize to other populations such as full time employees. It is possible that other populations may fulfill different roles and, as such, rely on different FORs when responding to non-contextualized personality inventories. Perhaps most obviously, full time employees that are not enrolled in classes are extremely unlikely to use the school FOR. And it is also likely that the different roles held by full time employees, such as that of a community volunteer, may lead to differences in the FORs used. Yet although there may be these differences, Study 1 has demonstrated an important principle that is very likely to generalize to different populations, that is, people rely on a variety of FORs when responding to non-contextualized personality inventories.

Despite these limitations, Study 1 was an important contribution to the FOR literature because it provided evidence of the validity of the assumption that participants do rely on context when responding to non-contextualized inventories, and it also provided an understanding of the range of those contexts. In addition, it helped generate a number of additional questions that should be addressed by future research and many of these are discussed in the next section.

Chapter 6

Development of Research Questions for Study 2

Based on the results of Study 1, one issue that deserves further exploration relates to whether certain items tend to elicit more or fewer FORs than other items. Data from in Study 1 indicated that certain FORs seemed to be used with greater frequency for items within certain factors (e.g., home FOR being associated with conscientiousness items). Such a trend warrants further investigation. The question of whether certain items elicit more or fewer FORs was addressed in an indirect way by Robie et al. (2000), a study described earlier. Robie et al. attempted to measure the context specificity of conscientiousness items by asking undergraduate students to rate the degree to which they viewed each item as being relatively broad or narrow. The goal was to use this information to explain differences in the factor structure of a non-contextualized versus a contextualized personality measure. The results were not nearly as informative as the authors had hoped, likely because the operationalization of the context specificity variable was flawed. For one, the authors used a convenience sample of undergraduates who had no advanced knowledge of personality theory or measurement. Thus, the brief definitions for “broad” versus “narrow” were likely inadequate to convey the full meaning of the information the researchers were attempting to gather. In addition, and perhaps more importantly, the sample that provided the context specificity ratings was not the sample that completed the personality inventories. Thus, the perceptions of item context specificity from a non-expert student sample were imposed onto a separate student sample. A more direct and

appropriate operationalization of the notion of context specificity might be to directly ask respondents to indicate the contexts in which they think of themselves when responding to items. This would provide a direct method for ascertaining the number of FORs a respondent uses for particular items.

Research Question 3: Do certain items/factors elicit more or fewer FORs than others?

The results from Study 1 indicated that certain FORs seemed to be more frequently used when responding to items within a particular factor (e.g., an “at work” FOR was often used for conscientiousness items), and Study 2 will extend those findings by taking a precise, quantitative approach to addressing this question by statistically comparing the number of FORs endorsed across items/factors.

A topic relevant to the quantity of FORs elicited by various items is the qualitative nature of the FORs. Research Question 3 will address how many FORs are endorsed for each item/factor, but a logical follow-up question involves examining which FORs are endorsed most frequently for each factor. For example, is a work FOR used more frequently for conscientiousness items than a family FOR? While Lievens et al. (2008) demonstrated that reliability and validity are adversely affected when multiple FORs are used, basic questions concerning 1) which items/factors might tend to elicit more FORs (i.e., Research Question 3), and 2) which FORs are used most frequently for each factor, have not been addressed. Answering these questions would provide useful information about how the items are generally interpreted by respondents.

Research Question 4: Are there significant differences in the FORs used within factors when responding to a non-contextualized personality measure?

Study 2 will be the primary vehicle for addressing this research question. The systematic approach taken in this quantitative study will provide an understanding of the differences in FORs that are used within factors when completing a non-contextualized measure.

The research questions posed thus far have been primarily centered around properties of the items (i.e., do certain items elicit more or fewer FORs, are there significant differences in the FORs elicited by certain factors?). The next two research questions are focused on the role the person plays in responding to non-contextualized personality inventories.

Bem and Allen's (1974) research indicates that there are individual differences in the extent to which people vary their behavior across situations. In a very interesting but simple study, the researchers asked respondents to indicate the extent to which their behavior on a particular dimension (e.g., friendliness, conscientiousness) varied from one situation to another. Results indicated that the degree to which people reported their behavior as varying across situations predicted the cross-situational stability of their behavior. Simply put, people who said they tended to behave differently in different situations did in fact appear to do so. Bem and Allen conclude that their results support an increased focus on the role of the person in trying to understand cross-situational variability in behavior.

Bem and Allen's (1974) findings are highly relevant to the proposed study. Primarily, they highlight the fact that individual differences do factor into the degree of cross-situational stability of behavior. In addition to the role played by items and the situational cues they contain, people also differ in the extent to which they are likely to see their behavior as variable across situations. Indeed, this is a key finding of the self-monitoring literature (Snyder & Gangestad, 1986). High self-monitors are described as 'social chameleons' and tend to alter their behavior to fit the situation, whereas low self-monitors tend to behave more consistently across contexts. Given these findings, it would be worthwhile to examine which of the situations/FORs identified in Study 1 yield the greatest degree of variability in endorsement, and also to explore which individual differences might serve to explain some of that variance.

Research Question 5: Which FORs exhibit the greatest degree of variability in endorsement, and what are the individual differences that are related to endorsement of each FOR/response option?

Research Questions five (above) and six (below) will be addressed in Study 2. The approach taken in this study will allow the number of contexts in which people see themselves to be quantified, and this variable can then be used as a moderator of factor structure.

As described in the FOR chapter, several studies have examined the question of whether the factor structure is different across contextualized versus non-contextualized personality inventories (e.g., Bing et al., 2004; Robie et al., 2001; Schmit et al., 1995). The general consensus that emerges from these studies is that

latent factor structure, loadings, and correlations are invariant, but error variances are larger in the non-contextualized measures. However, the approach taken by these studies is one that is focused on the item, not the person. A contextualized inventory imposes situational cues on the items by asking respondents to describe themselves “at work” or “at school.” And in this regard, the factor structures seem to be invariant. However, as discussed in the context of Bem and Allen’s (1974) findings, there are important individual differences that can affect the ways in which people respond to personality measures. Specifically, people may tend to see themselves in more or fewer contexts when responding to personality items. Thus, it is possible that individual variability in the tendency to rely upon more or fewer FORs may result in non-equivalent factor structures for individuals occupying these groups.

Research Question 6: Does the factor structure of the FFM differ for people who are relatively less or more cross-situationally consistent? More specifically, are the factor structure, loadings, correlations, and error variances for the factors that were found to be less context specific (i.e., as discussed in Research Question 3) non-equivalent across the two groups?

Method

Participants

Participants were 463 students recruited from advanced business and psychology courses at Portland State University. The results of a power analysis indicated that in order to achieve power of .90 for the multi-sample SEM in Research Question 6, at least 402 participants would be needed. Thus, the total participant pool

was large, consisting of 1,335 students from 21 courses (response rate = 35%). Of the 462 students who completed the survey, 20 were removed due to lack of work experience (4% of the total sample). Although it was not a requirement to be currently employed to participate, respondents needed to at least have prior work experience in order to make the “At work” FOR relevant. This step also increases the likelihood that the results will generalize to the working population. An additional 24 participants were removed from the dataset due to anomalous response patterns (described below).

The remaining 419 participants were majority White (72.5%) with American Indians (2.8%), Asians (15.5%) African Americans (3.5%), Hispanics (8.0%), and Pacific Islanders (2.3%) constituting the remainder of the sample. Females made up 68.8% of the sample and the mean age was 24.59 (SD = 6.69). Most participants were college juniors or seniors (70.2%). Nearly three-fourths (73.6%) of the sample were currently employed, and the most commonly held positions were in the sales or service industries (35.5%). On average participants worked 26.14 hours per week (SD = 12.99) and most held just one job (76.5%). In terms of home life, most participants were single (67.4%) and less than 20% of the sample had children for whom they were either partially or fully responsible. The majority of the sample (68.3%) has two living parents/step-parents, but does not spend any time caring for them (52.9% of the sample spent 0 hours per week caring for parents). Hours spent caring for parents was not significantly related to either participant age or marital status, indicating that younger, unmarried participants may also be assuming elder care responsibilities.

Procedure

Participants were recruited during class time. Recruitment consisted of providing potential participants with an overview of the study, informing them that it was expected to be low risk, and that they would likely receive extra credit for participating (only one instructor did not offer extra credit for participation). Participants were then provided with the information necessary to complete the survey online on their own time (e.g., survey link, deadline). Upon typing in the URL, participants were randomly assigned to one of three versions of the survey (see below). The informed consent was presented on the first page of the survey, and participants were asked to decide at that point if they would like to participate. The actual survey for Study 2 (see Appendix E) began by having participants complete a non-contextualized personality inventory. Afterwards participants were presented with the personality items again and asked to indicate for each item, using a “choose all that apply” format, which contexts they were considering when responding to the non-contextualized personality inventory. The self-monitoring scale appeared in the next section, and respondents were asked to provide demographic and work experience information at the end of the survey.

Survey Versions

Primarily due to concerns over fatigue affecting the quality of the data provided toward the end of the survey, three versions of the survey were created. The content of all three versions was identical, but the order in which the personality items were presented was not. Version 1 presented items in the same order that they appeared in for Study 1. Version 2 presented all of the positively worded items for

each factor first, followed by the negatively worded items. Version 3 presented all of the negatively worded items for each factor first, followed by the positively worded items. The presentation of the items for each survey version is provided in Table 5.

Measures

Participants completed the IPIP personality inventory described in Study 1. The inventory consisted of 50 items, with ten items representing each factor. Within each factor, five items are positively worded and five items are negatively worded. Based on the results for Study 1, a small edit was made to one of the conscientiousness items. The item that originally read, “Shirk my duties” was changed to “Try to avoid my duties.” All but two participants in Study 1 asked for clarification of the meaning of the word “shirk,” thus an appropriate synonym was used in Study 2 to avoid this confusion. Internal consistency estimates were computed for each factor and are as follows: agreeableness ($\alpha = .81$), conscientiousness ($\alpha = .86$), extraversion ($\alpha = .90$), neuroticism ($\alpha = .88$), and openness to experience ($\alpha = .82$). The internal consistency estimate for Snyder and Gangestad’s (1986) 18-item measure of self-monitoring was also acceptable ($\alpha = .71$).

After completing the standard IPIP inventory, participants were asked to indicate the contexts in which they were thinking of themselves when responding to each item. The contexts provided to participants are those that emerged in Study 1. More specifically, the response options were: at work, at school, at home, with friends, with family (including significant other), in general, and “other”. These response options were presented using a “check all that apply format” so participants

had the ability to indicate that they were thinking of themselves in more than one context or, for example, that they were thinking of themselves in general and at work. For all 50 items, the maximum number of contexts endorsed was 7 and the minimum was 0. The mean number of contexts endorsed for all 50 items is presented in Table 6.

Data Cleaning

Data were examined to determine unusual patterns of responding associated with context endorsement across all 50 personality items. A variable reflecting the total number of contexts (i.e., all FORs plus the “In General” response option) endorsed across all 50 items was created with values ranging from 0 to 280, and the maximum possible value was: $(50 \text{ items}) \times (7 \text{ contexts}) = 350$. *Z* scores were then computed for this variable, and 7 cases that had *Z* scores greater than 3 were removed due to having endorsed an exceptionally high number of contexts. In addition, 17 cases were removed due to having endorsed zero contexts. This latter decision may be somewhat controversial because it is possible that perhaps these participants were not thinking of any of the provided FORs. However, in that case the reasonable responses would have been “In General” or “Other.” Based on Study 1 data, it is unlikely that if participants were fully attentive that they would not have endorsed any contexts. Thus, not endorsing any contexts for any of the items was considered an anomalous response pattern and cases that fit this criterion were removed from the analyses. After performing this step, the number of participants not endorsing any response options for a given item ranged from 0 to 3% across all 50 items.

Across the non-contextualized IPIP inventory and the self-monitoring scale, missing values for each item ranged from 0-4. Given the relatively few instances of missing data for these items, a simple mean imputation strategy was used. Although the mean imputation strategy for replacing missing data has been criticized because it can result in reduced variability (Kline, 2005) replacing no more than 4 missing values (out of 419 total) with the mean for an item did not impact variability.

Analysis Strategy

Research Questions 3 and 4 were addressed with a series of one-way repeated measures ANOVAs. Omnibus tests across items and factors were conducted and follow-up pairwise comparisons correcting for Type I error were performed. Analysis of Research Question 5 involved calculating the coefficient of variation for the endorsement of each response option (i.e., each FOR plus the “In General” and “Other” response option), and correlating a number of individual difference variables with the frequency of endorsement of each response option. Finally, a multiple-sample SEM was performed to address Research Question 6.

Results

Research Question 3

Research Question 3 was focused on determining whether certain items and/or factors elicit more or fewer FORs than others. This first step toward answering this question was to create variables reflecting the total number of contexts each participant endorsed for each item. Thus, one variable was created for each of the 50 personality items that indicated how many contexts a participant had endorsed for

each item. This variable *did not* include the “In General” option because selecting the In General option indicated that a participant was thinking of themselves as they generally behave, rather than how they behave in any particular context. The values of these variables ranged from 0 (did not endorse any contexts) to 6 (endorsed all contexts). After creating these variables for each personality item, the values of these variables within each factor was summed to create a variable reflecting the frequency of contexts endorsed across all items within a factor. Theoretically, the highest possible value for this variable was 60 (10 items X 6 contexts), but the observed maximum value was 50. Descriptive statistics for each of these variables is presented in Table 6. Fifteen of the 50 items had means of less than 1.00, indicating that on average respondents were endorsing either one or no FORs for these items. Rather, for these items, which were primarily from neuroticism and openness to experience, participants were relying heavily on the “In General” response option. Eight items had means greater than 2.00, indicating that for these items participants were endorsing, on average, at least 2 FORs.

Factor level comparisons. The first analysis focused on comparing the frequency of context endorsement across the five personality factors. The one-way repeated measures ANOVA was significant, $F(3.85, 1.610.80) = 101.08, p < .05$, partial $\eta^2 = .20$.¹ Descriptively, agreeableness had the highest mean number of contexts endorsed (mean = 16.32, SD = 10.77), followed by conscientiousness (mean = 15.45, SD = 8.62), extraversion (mean = 14.53, SD = 9.49), neuroticism (mean = 11.61, SD = 9.75), and openness to experience (mean = 9.59, SD = 8.97). Follow-up

pairwise comparisons using the Bonferroni correction for Type I error indicated that the mean difference between agreeableness and conscientiousness was not significant, and neither was the difference between conscientiousness and extraversion (see Table 7). However, the means for agreeableness and conscientiousness were significantly higher than those of neuroticism and openness to experience, indicating that participants used FORs more frequently when responding to agreeableness and conscientiousness items than when responding to neuroticism or openness to experience items.

Agreeableness. A one-way repeated measures ANOVA was conducted for agreeableness with the ten items serving as the levels of the within-subjects factor, thus the number of FORs endorsed for each item were the values being compared. The ANOVA was significant, $F(3.16, 1320.60) = 90.40, p < .05, \text{partial } \eta^2 = .18$. The mean number of FORs endorsed ranged from 2.45 ('Respect others') to 0.77 ('Get back at others; see Table 8 for all means). Follow-up pairwise comparisons using the Bonferroni method are also presented in Table 8. Interestingly, positively worded items tended to elicit relatively more FORs than negatively worded items.

Conscientiousness. A one-way repeated measures ANOVA with ten levels was conducted for conscientiousness, with the number of FORs endorsed for each of the ten items serving as the cell values. The ANOVA was significant, $F(3.62, 1515.03) = 65.89, p < .05, \text{partial } \eta^2 = .14$. The mean number of FORs endorsed ranged from 2.16 ('Carry out my plans') to 1.00 ('Don't see things through; see Table 9 for all means). Follow-up pairwise comparisons using the Bonferroni method are also presented in

Table 9. As with the agreeableness items, positively worded items tended to elicit relatively more FORs than negatively worded items.

Extraversion. A one-way repeated measures ANOVA was conducted for extraversion with the ten items serving as the levels of the within-subjects factor, thus the number of FORs endorsed for each item were the values being compared. The ANOVA was significant, $F(3.83, 1601.79) = 54.01, p < .05, \text{partial } \eta^2 = .11$. The mean number of FORs endorsed ranged from 2.28 ('Feel comfortable around people') to 0.94 ('Would describe my experiences as somewhat dull; see Table 10 for all means). Follow-up pairwise comparisons using the Bonferroni correction are also presented in Table 10. Again, positively worded items tended to elicit relatively more FORs than negatively worded items.

Neuroticism. A one-way repeated measures ANOVA with ten levels was conducted for neuroticism, with the number of FORs endorsed for each of the ten items serving as the cell values. The ANOVA was significant, $F(3.82, 1598.20) = 48.90, p < .05, \text{partial } \eta^2 = .11$. The mean number of FORs endorsed ranged from 1.77 ('Feel comfortable with myself') to 0.63 ('Am often down in the dumps; see Table 11 for all means). Follow-up pairwise comparisons using the Bonferroni correction are also presented in Table 11. For the neuroticism items, the negatively keyed items elicited relatively more FORs than the positively keyed items. However, for the neuroticism items the negatively keyed items are the items with positive connotations, and may also be associated with greater socially desirable responding.

Openness to experience. A one-way repeated measures ANOVA with ten levels was conducted for openness to experience, with the number of FORs endorsed for each of the ten items serving as the levels of the within-subjects factor. The ANOVA was significant, $F(4.23, 1766.43) = 76.96, p < .05, \text{partial } \eta^2 = .16$. The mean number of FORs endorsed ranged from 1.92 ('Enjoy hearing new ideas) to 0.39 ('Tend to vote for conservative political candidates;' see Table 12 for all means). Follow-up pairwise comparisons using the Bonferroni correction are also presented in Table 12. The openness to experience items followed the same general pattern as the items within the other factors, that is, with positively keyed items eliciting more FORs. The one exception for openness is the item 'Tend to vote for liberal political candidates.' Although this item is positively keyed, its mean (0.66) was the seventh lowest.

Research Question 4

The goal of Research Question 4 is to determine whether there are significant differences in the FORs used when responding to non-contextualized items/factors. In other words, do respondents tend to use one FOR with greater frequency than other available FORs or the In General option? While Research Question 3 addressed the question of the frequency with which *any* FORs were used in responding, Research Question 4 is aimed at understanding *which* FORs were used.

The descriptive statistics presented in Table 6 provide some interesting descriptive information related to Research Question 4. In examining the FORs that were endorsed most frequently in association with each item some interesting patterns

emerged. First, more than 50% of respondents endorsed the “With Friends” FOR for the positively worded agreeableness items. In addition, more than 50% of respondents endorsed both the “At School” and “At Work” FORs for the conscientiousness items “Am always prepared” and “Pay attention to details.” Regarding extraversion, more than 50% of respondents endorsed the “With Friends” FOR for the items “Feel comfortable around people,” “Am skilled in handling social situations,” and “Am the life of the party.” Interestingly, none of the neuroticism or openness to experience items had FORs that were endorsed by more than 50% of respondents. Finally, it is interesting to note that the “In General” response option was endorsed by more than half of the participants for only 22 of the 50 items.

Repeated measures ANOVAs were used to address Research Question 4. The first step involved computing a variable that reflected the number of times each FOR had been endorsed for items within each factor. For example, the first variable that was computed was a sum of the number of times the “With Family” FOR was endorsed for all ten agreeableness items. This process was repeated for each response option, and then again for each of the four remaining factors, yielding 35 new variables. Each of these variables ranged from 0 (the response option was not endorsed for any of the ten items within the factor) to 10 (the response option was endorsed for all ten items within the factor). An omnibus 5 (factors) X 7 (response options) repeated measures ANOVA was performed to determine whether there were main effects due to factors and response options, and whether there was an interaction between factors and response options. The “In General” response option was endorsed

the most frequently across factors, and the Agreeableness items had the highest number of response options endorsed. There were significant main effects for factor, $F(3.81, 1592.99) = 74.80, p < .01, \text{partial } \eta^2 = .15$, and response option, $F(2.57, 1074.02) = 234.67, p < .01, \text{partial } \eta^2 = .36$, and there was a significant interaction between factor and FOR, $F(14.32, 5983.70) = 92.33, p < .01, \text{partial } \eta^2 = .18$. An graph of the interaction is presented in Figure 2. Follow-up pairwise comparisons for the factor main effect using the Bonferroni correction for Type I error indicated significant differences in endorsement across all factors except conscientiousness and extraversion. Follow-up pairwise comparisons for the response option main effect indicated significant differences between all response options except “With Friends” and “At Work”, and “At Home” and “At Work.”

Of particular interest with respect to Research Question 4 was the interaction between factor and response options. In order to determine significant differences within each level of factor and response option, a series of one-way repeated measures ANOVAs were first conducted within factors (e.g., comparing the frequency with which each FOR was endorsed within agreeableness) and then within response options (e.g., comparing the relative endorsement of the “With Family” FOR across agreeableness, conscientiousness, extraversion, neuroticism, and openness to experience).

Follow-Up Analyses by Factor

Agreeableness. A one-way repeated measures ANOVA with seven levels (one for each response option including In General) was conducted to determine which

response option was endorsed with the greatest frequency across agreeableness items. The ANOVA was significant, $F(3.54, 1478.19) = 173.15, p < .05$, partial $\eta^2 = .29$. Not surprisingly, the “In General” response option was endorsed with the greatest frequency (mean = 5.13, SD = 2.71). The “With Friends” response option was endorsed with the second greatest frequency (mean = 3.99, SD = 2.77). Means for all response options are provided in Table 13. Follow-up pairwise comparisons using the Bonferroni correction for Type I error were performed, and results indicated that the mean for the “With Friends” FOR was significantly higher than the mean for all other FORs (not including the “In General” option). Table 13 also presents the mean differences between response options.

Conscientiousness. A one-way repeated measures ANOVA with seven levels (one for each response option including In General) was conducted to determine which response option was endorsed with the greatest frequency across conscientiousness items. The ANOVA was significant, $F(3.93, 1640.98) = 153.45, p < .05$, partial $\eta^2 = .27$. The “At School” FOR was endorsed with the greatest frequency (mean = 4.12, SD = 2.60) and the “At Home” response option was endorsed with the second greatest frequency (mean = 3.53, SD = 2.77). Means for all response options are provided in Table 14. Follow-up pairwise comparisons using the Bonferroni correction for Type I error were performed, and results indicated that the mean for the “At School” FOR was significantly higher than the mean for all other response options. Table 14 also presents the mean differences between response options.

Extraversion. A one-way repeated measures ANOVA with seven levels (one for each response option) was conducted to determine which response option was endorsed with the greatest frequency across extraversion items. The ANOVA was significant, $F(3.64, 1523.09) = 151.28, p < .05$, partial $\eta^2 = .27$. As with agreeableness, the “In General” response option was endorsed with the greatest frequency (mean = 4.40, SD = 2.91). The “At School” response option was endorsed with the second greatest frequency (mean = 3.76, SD = 2.57). Means for all response options are provided in Table 15. Follow-up pairwise comparisons using the Bonferroni correction for Type I error were performed, and results indicated that the mean for the “At School” FOR was significantly higher than the mean for all other FORs, but not for the “In General” response option. Table 15 also presents the mean differences between response options.

Neuroticism. A one-way repeated measures ANOVA with seven levels (one for each response option) was conducted to determine which response option was endorsed with the greatest frequency across neuroticism items. The ANOVA was significant, $F(2.63, 1097.95) = 149.69, p < .05$, partial $\eta^2 = .26$. As with agreeableness and extraversion, the “In General” response option was endorsed with the greatest frequency (mean = 5.02, SD = 3.52). The “At Home” response option was endorsed with the second greatest frequency (mean = 2.77, SD = 2.69). Means for all response options are provided in Table 16. Follow-up pairwise comparisons using the Bonferroni correction for Type I error were performed, and results indicated that the mean for the “At Home” FOR was significantly higher than the mean for all other

FORs, and that the “In General” response option was endorsed significantly more than any other response option. Table 16 also presents the mean differences between response options.

Openness to experience. A one-way repeated measures ANOVA with seven levels (one for each response option) was conducted to determine which response option was endorsed with the greatest frequency across openness to experience items. The ANOVA was significant, $F(2.63, 1099.54) = 269.44, p < .05, \text{partial } \eta^2 = .39$. As with agreeableness, extraversion, and neuroticism, the “In General” response option was endorsed with the greatest frequency (mean = 5.54, SD = 3.07). The “At School” response option was endorsed with the second greatest frequency (mean = 2.18, SD = 2.27). Means for all response options are provided in Table 17. Follow-up pairwise comparisons using the Bonferroni correction for Type I error were performed, and results indicated that the mean for the “At School” FOR was significantly higher than the mean for all other FORs, and that the “In General” response option was endorsed significantly more than any other response option. Table 17 also presents the mean differences between response options.

Follow-Up Analyses by Response Option

In general. Although the “In General” response option is not a FOR, it is still informative to know which factors tend to receive the highest endorsement of the “In General” response option. A one-way repeated measures ANOVA with five levels, one for each personality factor, was significant, $F(3.68, 1539.61) = 75.22, p < .05, \text{partial } \eta^2 = .15$. The openness to experience items had the highest endorsement of the

“In General” response option (mean = 5.54, SD = 3.07), followed by agreeableness, neuroticism, extraversion, and conscientiousness. Means for each factor are presented in Table 18. Follow-up pairwise comparisons using the Bonferroni method (presented in Table 18) indicate significant mean differences across all factors except agreeableness and neuroticism. In terms of practical significance, this result is among the most compelling because it demonstrates that, on average, participants are thinking of themselves in general on about 5 openness to experience items, whereas participants are thinking of themselves in general on only 3 conscientiousness items.

With family FOR. The next set of analyses involved conducting one-way repeated measures ANOVAs to compare the relative endorsement of each FOR across the five personality factors. The ANOVA for the “With Family” FOR was significant, $F(3.89, 1627.11) = 71.96, p < .05, \text{partial } \eta^2 = .15$. Descriptively, agreeableness items had the highest endorsement of the “With Family” FOR (mean = 3.19, SD = 2.46), followed by extraversion, neuroticism, conscientiousness, and openness. Means for each factor are presented in Table 19. Follow-up pairwise comparisons using the Bonferroni method (presented in Table 19) indicate that the mean endorsement for agreeableness items is significantly higher than for any other factor. In fact, mean differences between all factors are significant, with the exception of the difference between conscientiousness and neuroticism.

With friends FOR. A one-way repeated measures ANOVA with five levels, one for each personality factor, was significant, $F(3.68, 1536.00) = 118.26, p < .05, \text{partial } \eta^2 = .22$. As with the “With Family” FOR, agreeableness items had the highest

endorsement (mean = 3.99, SD = 2.77), followed by extraversion, conscientiousness, neuroticism, and openness. Means for each factor are presented in Table 20. Follow-up pairwise comparisons using the Bonferroni method (presented in Table 20) indicate that the mean endorsement for agreeableness items is significantly higher than for any other factor. Mean differences between conscientiousness, neuroticism, and openness to experience are not significant.

At home FOR. A one-way repeated measures ANOVA with five levels, one for each personality factor, was significant, $F(3.59, 1501.81) = 90.82, p < .05$, partial $\eta^2 = .18$. The conscientiousness items had the highest endorsement of the “At Home” FOR (mean = 3.53, SD = 2.44), followed by neuroticism, agreeableness, extraversion, and openness to experience. Means for each factor are presented in Table 21. Follow-up pairwise comparisons using the Bonferroni method (presented in Table 21) indicate that the mean endorsement for conscientiousness items is significantly greater than endorsement for any other factor.

At school FOR. A one-way repeated measures ANOVA with five levels, one for each personality factor, was significant, $F(3.79, 1582.14) = 110.15, p < .05$, partial $\eta^2 = .21$. The conscientiousness items had the highest endorsement of the “At School” FOR (mean = 4.12, SD = 2.60), followed by extraversion, agreeableness, neuroticism, and openness to experience. Means for each factor are presented in Table 22. Follow-up pairwise comparisons using the Bonferroni method (presented in Table 22) indicate significant differences in mean endorsement of the “At School” FOR across all factors

excepting the comparison between conscientiousness and extraversion, and between neuroticism and openness to experience.

At work FOR. A one-way repeated measures ANOVA with five levels, one for each personality factor, was significant, $F(3.83, 1604.57) = 87.29, p < .05$, partial $\eta^2 = .17$. The conscientiousness items had the highest endorsement of the “At Work” FOR (mean = 3.22, SD = 2.54), followed by agreeableness, extraversion, neuroticism, and openness to experience. Means for each factor are presented in Table 23. Follow-up pairwise comparisons using the Bonferroni method (presented in Table 23) indicate that the difference between mean endorsement for conscientiousness and agreeableness items is not significant. In addition, the difference between conscientiousness and extraversion is significant, but the difference between agreeableness and extraversion is not.

Other FOR. The “Other” FOR was provided to allow participants to indicate that they were thinking of a situation or context other than those provided. It was endorsed very infrequently, but conducting an ANOVA to determine which factor had the highest endorsement of this response option is potentially informative because it can provide some indication of the adequacy of the response options provided for each factor. A one-way repeated measures ANOVA with five levels, one for each personality factor, was significant, $F(1, 418) = 11.33, p < .05$, partial $\eta^2 = .03$. The agreeableness and openness items had the highest endorsement of the “Other” FOR (means = 0.75, SDs = 1.60 and 1.71, respectively), followed by neuroticism, extraversion, and conscientiousness. Means for each factor are presented in Table 24.

Follow-up pairwise comparisons using the Bonferroni method (also presented in Table 24) indicated very few significant differences in mean endorsement across factors.

Research Question 5

Research Question 5 addresses the variability in FOR endorsement for each FOR/response option, and the extent to which individual differences can be identified that explain this variance. The first step, calculating and comparing the variance associated with the frequency of endorsement for each response option, involved computing the coefficient of variation for each response option. The coefficient of variation is a way of comparing standard deviations for distributions with quite different means (Howell, 2002). Considering that mean endorsement for each response option ranged from 3.34 (Other) to 23.53 (In General) it was necessary to scale the standard deviations by their respective means in order to appropriately compare them. Table 25 presents the mean, standard deviation, and coefficient of variation, along with a 5 number distribution (minimum, 25th percentile, 50th percentile, 75th percentile, and maximum values), for each response option. Using the coefficient of variation statistic, the distribution of the Other response option has the greatest dispersion about its mean with a value of 2.04, indicating that the standard deviation is slightly more than twice the size of the mean. This is the result of the relatively infrequent endorsement of the Other category, evidenced by not only its relatively small mean (3.34) but by the fact that the median for this response option is 0.00. The In General response option had the least dispersion about its mean (CV = .52). Among the FORs the CVs were relatively similar, with the With Family FOR having the largest (CV =

.76), followed by the At Work FOR (CV = .75), At Home (CV = .74), With Friends (CV = .66), and At School (CV = .62).

Exploratory analyses examining the correlations between the endorsement of each response option and several individual difference variables were also conducted in order to address Research Question 5. The results of these analyses are presented in Table 26. Participant age demonstrated a statistically significant, negative relationship with the endorsement of the With Family, With Friends, and At School FORs, and a statistically significant, positive relationship with the In General and Other response options. Gender was not significantly related to endorsement of any of the response options. Years of work experience was significantly negatively related to endorsement of the With Family, With Friends, and At School FORs, although it was not significantly related to endorsement of the At Work FOR. Employment status was, however, significantly related to endorsement of the At Work FOR, such that participants who were currently employed endorsed the At Work FOR more frequently than those who were not currently employed.

In terms of relationships between personality and the endorsement of FORs, self-monitoring was significantly related to endorsement of the With Family, With Friends, and At Work FORs such that high self-monitors had higher levels of endorsement of these FORs than low self-monitors. There were very few statistically significant relationships between the FFM personality factors and FOR endorsement. The relationship between extraversion and endorsement of the In General response option was significant, indicating that more extraverted individuals were less likely to

rely endorse the In General response option than were introverts. Perhaps more interesting was the significant relationship between conscientiousness and endorsement of the At Work FOR which indicated that conscientious individuals were more likely to endorse an At Work FOR than were less conscientious participants.

Research Question 6

The final analysis in Study 2 addressed invariance in factor structure across individuals who vary in their cross-situational consistency. Snyder and Gangestad's (1986) 18-item measure of self-monitoring was administered to participants and its correlations with the context specificity variables created in Research Question 3 were tested with the intention of establishing some initial construct validity evidence for the context specificity variables. The 5 context specificity variables reflect the number of FORs that were endorsed for each factor. For the purposes of Research Question 6, a sixth context specificity variable (overall context specificity) was created that reflects the total number of FORs endorsed across all factors.

Correlations between the self-monitoring scale and the context-specificity variables were weak. Correlations between self-monitoring and the overall context specificity variable ($r = .11, p = .02$), context specificity for neuroticism ($r = .13, p = .01$), and context specificity for openness ($r = .14, p = .01$) were significant, but correlations between self-monitoring and context specificity for agreeableness ($r = .07, p > .05$), context specificity for conscientiousness ($r = .07, p > .05$), and context specificity for extraversion ($r = .07, p > .05$) were not significant.

To determine whether the Big 5 factor structure differed based on context specificity, six multiple sample CFAs were conducted. The analyses followed tests for measurement invariance as described by Vandenberg and Lance (2000) and were conducted using AMOS 7.0. The process involves examining the fit of a model that places no equality constraints across groups to models that place increasing levels of constraints (i.e., constrained factor loadings and covariances, constrained residual covariances). The first CFA focused on differences in Big 5 structure based on the overall context specificity variable. A median split was performed that divided the sample into two groups based on the overall context specificity variable. The initial fit of the unconstrained model was poor, $\chi^2(2330) = 5323.19$, CFI = .68, RMSEA = .05.

Modification indices for both groups were consulted to determine paths that could be added to improve fit. Modification indices were very similar for both groups and suggested that the most significant gains in fit would be achieved by allowing correlations among error terms for certain items. Suggested modifications were sorted by modification index values and then a rational process was used to determine which correlations among error terms should be added. For example, a path was added between the openness to experience items “Tend to vote for liberal political candidates” and “Tend to vote for conservative political candidates” because it is likely that observed values for these items share a source of variance that is not accounted for by openness (e.g., political affiliation). It was necessary to correlate 29 error variances in order to obtain a reasonable fit for the unconstrained model, $\chi^2(2372) = 3845.84$, CFI = .93, RMSEA = .04.

The unconstrained model was then compared to three progressively more restrictive models, the first of which constrained factor loadings to be equal, the second of which constrained factor loadings and covariances to be equal, and the third of which constrained the covariances among residuals. Table 27 presents the results of these comparisons. Using the chi-square difference test, the differences between the unconstrained model and models 1 and 2 were not significant, but model 3 fit significantly worse than the unconstrained model. This indicates that the assumption of equal covariances among residuals between groups is not supported.

Although the CFA containing all Big 5 factors was somewhat informative in that it indicates that the covariance structure of the residuals differs based on context specificity, the model is somewhat lacking in sensitivity to differences in FOR use across factors. Indeed, the results of Research Question 3 demonstrated that there are significant differences in context specificity across factors. Therefore, it seemed more appropriate to use the context specificity variables for each factor to split the sample within each factor, and then examine the potential differences in factor structure by examining each factor on its own. For example, a median split was performed using context specificity variable for agreeableness, and then differences in the factor structure of *just agreeableness* were examined. Aside from the potential to be a more theoretically meaningful analysis, reducing the number of parameters in each CFA results in a more reasonable ratio of observations to parameters (approximately 4:1) than in the CFA containing all five factors (approximately 1.2:1).

Initial fit for the models containing each factor separately were much better than for the model containing all five factors, but did not achieve levels of acceptable fit for most indices. Modification indices were used in the same way as described above to attempt to improve fit. In addition, modification indices for the individual factor CFAs were compared to modification indices for the overall Big 5 CFA. Because the modification indices suggested by the individual factor CFAs were very similar to the modification indices suggested by the overall Big 5 CFA, the same paths (i.e., correlated errors) were added to the individual factor CFAs for consistency.

Results of the model comparisons for each factor can be found in Table 27. In summary, there were no significant differences between the models tested for agreeableness and extraversion. However, for conscientiousness, neuroticism, and openness to experience, the fit of the model with constrained residuals was significantly worse than the unconstrained model. For these three factors, then, error variance is higher for participants who rely on more FORs when responding than it is for participants who rely on relatively fewer FORs.

Supplementary Analyses

A supplementary set of analyses was conducted to determine whether the actual situation participants were in when responding to the survey was related to FOR usage. At the end of the survey, participants were asked to indicate (using a check all that apply format) which of the following situations they were in when responding: at home, at work, at school, with friends, with family, or in another situation not listed. Each of these contexts served as the independent variable for a series of t-tests for

independent means. The dependent variables in these analyses were the frequency of endorsement for each FOR across all Big 5 factors. The first set of *t*-tests, then, compared the mean endorsement for each FOR (i.e., total use of the “At Home,” “At Work,” “At School,” “With Friends,” “With Family,” and “Other” FORs across all factors) for participants who had indicated that they were “At Home” when filling out the survey. The second set used the same dependent variables, but the independent variable split the sample into two groups based on whether or not they were “At School” when responding. A total of 36 *t*-tests for independent means were conducted using this approach, and none of them indicated statistically significant differences in FOR use based on the actual context participants were in when responding. Table 28 presents the means, SDs, and *t*-test results for each of these comparisons.

Discussion

Study 2 was aimed at taking a quantitative approach to understanding the use of context in non-contextualized personality inventories. Analyses for Research Question 3 demonstrate that there are significant differences in the number of FORs used when responding to items within a factor. Results for Research Question 4 demonstrate that there are significant differences in the endorsement of certain FORs both within and across factors. Analyses for Research Question 5 indicate that there are differences in the variability associated with endorsement of each FOR/response option, and that some of this variation may be explained by individual differences. The degree of individual variation in the use of FORs does not appear to affect the factor structure of the Big 5, as evidenced by analyses for Research Question 6.

Analyses for Research Question 3 involved taking a more in-depth look at the FORs used in responding to individual items, and some interesting differences in the number of contexts endorsed across items within each factor were revealed. For eight of the 50 items, the mean number of FORs endorsed was greater than 2.00. These eight items provide further support for Study 1 findings in that they indicate that participants are, on average, thinking of themselves in at least 2 contexts when responding to certain items. Examining the percentage of respondents that endorsed each FOR for these items yields additional insight. Four of the eight items with means greater than 2.00 are agreeableness items. As would be expected based on Research Question 4 results, the “With Friends” FOR was heavily endorsed by participants, with more than 50% of respondents referring to it when responding to the 4 items with means greater than 2.00. Thus, it is reasonable to conclude that people are likely to be considering how they behave when with friends as they respond to these items. However, identifying the second FOR that is most likely to be used in association with these items is somewhat more difficult. For example, for the item, “Respect others” 49% of participants endorsed the “With Family” FOR and 49% of participants endorsed the “At School” FOR. These descriptive statistics are useful in attempting to make educated guesses about the FORs that are most likely to be used in association with certain items, but they certainly do not provide definitive answers.

The 15 items that had means of less than 1.00 are also interesting because the implication could be that these are items that may approximate the condition of “context-free.” That is, because participants did not tend to endorse FORs for these

items, they may be items for which participants think of themselves generally, rather than within any specific context. It is also worth noting that all of the negatively keyed items for neuroticism and openness to experience had means of less than 1.00, indicating that as a whole, these two factors may be less context-dependent than agreeableness, conscientiousness, and extraversion. On the other hand, rather than being “context-free”, these items may simply be an amalgamation of many contexts in the minds of test-takers, in which case, it is unclear how much one context is weighed relative to another when a particular individual is responding to an item. In other words, rather than being context-free, these items may present another source of error across test-takers.

Another interesting pattern of results discerned through Research Question 3 was that positively worded items tended to elicit endorsement of a higher number of FORs than negatively worded items. One possible reason may simply refer back to what was learned in Study 1, that is, respondents tend to struggle with negatively worded items. Because participants needed to devote more resources to understanding and choosing the accurate response for these items, they may be less cognizant of context while responding than they were when responding to positively worded items. Another possibility has to do with the social desirability of the positively worded items. Positively worded items, generally speaking, imply socially desirable behaviors (e.g., Respect others, Get chores done right away, Carry the conversation to a higher level). It is possible that participants simply like to think of themselves as engaging in these socially desirable behaviors across many situations and contexts. It is also

possible that the number of contexts in which participants engage in these more socially desirable (i.e., positively worded) behaviors is actually higher than the number of contexts in which they engage in less socially desirable (i.e., negatively worded) connoted behaviors (e.g., Have a sharp tongue, Find it difficult to get down to work; Have frequent mood swings). And this, in turn, is possibly a byproduct of the tendency to engage in socially undesirable behaviors less frequently.

The final point to highlight regarding Research Question 3 involves the endorsement of the “In General” response option, which was endorsed by more than 50% of respondents for only 22 of the 50 items. In theory, if participants were following the instructions provided, all participants should have endorsed the “In General” option for all items because the instructions specifically state that they should think of themselves “as they generally are.” However, the highest rate of endorsement for the “In General” option for any item was 80%. There are a few ways to interpret this finding. First, the simple explanation here is that participants were not fully attentive to the instructions or that they simply did not bother to repeatedly select the “In General” response. Second, another possible explanation is that while participants may have been subtly considering themselves in general, one or more of the FOR response options resonated with them, and they opted to select the option that was most salient. Study 1 findings indicate that this is a possibility because while participants typically began the response process by thinking of how they generally are, they would sometimes latch onto an examination of their behavior in a specific context. It would make sense that if Study 2 participants were following this same

process, they would be more likely to select the response option that was the most salient in their thought process. A third possibility, which is important to constructing personality items, is to consider that certain dimensions, and certain items within certain dimensions, relate much more to certain contexts than to others, that is, that certain items and contexts may elicit a certain context from a test-taker regardless of what instructions they are given.

Research Question 4 analyses focusing on differences in the use of FORs within factors generally supported the descriptive statistics calculated for Study 1. For agreeableness, Study 2 indicated that participants most frequently endorsed the “In General” response option, followed by the “With Friends” option. In Study 1, the non-contextual response appeared the most frequently as well, followed by family and friends. Results for openness were also consistent across Studies 1 and 2, that is, after the “In General” response, the “At School” response was endorsed with the greatest frequency. The pattern of FOR endorsement for extraversion items was also consistent across Studies 1 and 2, with the “In General” option being endorsed most frequently, followed by “At School” and “With Friends.”

The Research Question 4 results for conscientiousness were interesting in that it was the only factor for which the “In General” response option was *not* the most frequently endorsed in Study 2. Instead, “At School” was endorsed the most frequently, followed by “At Home” and then by “In General.” In Study 1, the non-contextual code was followed in frequency by school and then home. The Study 2 results are particularly interesting because they imply that because respondents were

more likely to be thinking of themselves in one or two specific situations than they were to be thinking of themselves in general, conscientious behavior may be particularly prone to situational specificity when compared to the other four factors. Also, given the interest in conscientiousness in personnel selection contexts (Barrick & Mount, 1991) and research showing that a Work FOR increases validity for selection (e.g., Hunthausen et al., 2003), it seems important to note that the “default” for these participants was not a Work FOR.

Research Question 4 results for neuroticism were also intriguing because while Study 1 did not reveal any discernable pattern in the use of FORs for neuroticism, Study 2 indicated that the “At Home” FOR was used significantly more than any other FOR excepting the “In General” response option. One possible reason for this difference could simply be the differences in how FORs were indicated. Because Study 2 provided the option to indicate that respondents were thinking of themselves at home, they chose it, whereas Study 1 participants were not given this option and simply had to describe their thought process. Another possible reason is the additional privacy afforded by Study 2. As mentioned in Study 1 results, the neuroticism items had a particularly high occurrence of no elaboration codes when compared to the other factors, possibly due to the sensitive nature of the items. However, Study 2 participants were afforded a degree of anonymity in responding not offered to Study 1 participants, and therefore may have been more inclined to offer additional information for these sensitive items. Finally, it is likely that neuroticism is a characteristic that pervades all areas of a person’s life, hence respondents would tend

to think about themselves in general, rather than in any specific FOR. And perhaps the reason for the frequent “At Home” FOR is because this context allows for the reflection required for people to have the negative thoughts included under neuroticism.

The results examining the use of each FOR across factors were in line with expectations based on Study 1, and also with what might be expected on the basis of the analyses focusing on FOR endorsement within factors. The results for the “At Work” FOR are interesting because participants indicated that they were thinking of themselves at work when responding to, on average, 3 items each from the conscientiousness, agreeableness and extraversion factors. This speaks to the relevance of an “At Work” FOR for these factors.

Research Question 5 involved examining the variability associated with each FOR/response option and then identifying which individual differences might be related to endorsement of each FOR/response option. Calculating the coefficient of variation (CV) for each response option indicated that there were notable differences in the dispersion of values around the mean when comparing the “In General” and “Other” categories, such that the “Other” category had a CV nearly 4 times as large as the “In General” category. As mentioned in the results section, this is due to the infrequent endorsement of the “Other” category which created rather extreme positive skew for the “Other” response distribution. Perhaps more interesting were the relatively similar sizes of the CVs for the FORs. The CVs for the FORs ranged from .62 (“At School”) to .76 (“With Family”), indicating that, on average, the standard

deviations for each of these distributions were a little more than half of the value of the mean. The relatively similar sizes of the CVs demonstrates that the variability associated with each FOR is comparable, indicating that while there may be differences in the mean endorsement of each FOR (as evidenced by the main effect of response option in Research Question 4 analyses) the dispersion of endorsement around the mean values are comparable.

Research Question 5 also involved an exploratory analysis aimed at identifying individual difference variables that might be related to endorsement for each of the response options. Of the demographic variables included, age yielded negative correlations with endorsement of the “With Family,” “With Friends,” and “At School” FORs and positive correlations with the “In General” and “Other” response options. When interpreting these findings, it is important to bear in mind that mean age for the sample was 24.52 and that 26 marks the 75th percentile for the sample. Finding that relatively older participants in the sample tended to rely less on the use of FORs than did their younger counterparts, and more on “In General” assessments of themselves could speak to the stabilization of personality that many researchers believe occurs as one approaches 30 (e.g., Pervin & John, 2010). As self-awareness develops and personality stabilizes, it is possible that people may come to see their behavior as less situationally dependent. However, given the small effect sizes for the relationships, these findings should be interpreted cautiously.

Another interesting finding for the demographic individual difference variables was the relationship between work experience and endorsement of the “At Work”

FOR. Somewhat surprisingly, the relationship between years of work experience and the “At Work” FOR was not significant, but the relationship between current employment status (employed/not employed) and the “At Work” FOR was significant. It is possible that the relationship between years of work experience and endorsement of the “At Work” FOR may not be linear, that is, simply having any work experience may lead to the use of the “At Work” FOR, and endorsement may not become more frequent as years of work experience increases. The more proximal variable in relation to the “At Work” FOR appears to be whether or not one is currently employed, rather than how much or little work experience one has.

Regarding the association between personality factors and FOR endorsement, interesting results were found with respect to self-monitoring and conscientiousness. Participants with relatively high levels of self-monitoring were significantly more likely to endorse the “With Family,” “With Friends,” and “At Work” FORs than were participants with lower levels of self-monitoring. These results are particularly interesting when compared to the findings of Research Question 6 because they indicate that while high self-monitors may not endorse a greater number of FORs when compared to low self-monitors (i.e., Research Question 6) high self-monitors do appear to use these three FORs more frequently than do low self-monitors. This finding indicates that there may be underlying associations between self-monitoring and FOR use that should be further explored.

Results of Research Question 5 also indicated that conscientiousness is significantly associated with endorsement of the “At Work” FOR, such that more

conscientious participants used the “At Work” FOR more frequently. This finding could potentially shed some light on why non-contextualized measures of conscientiousness tend to consistently predict work performance. If more conscientious individuals tend to use a Work FOR more frequently than do less conscientious individuals, this could have the effect of creating conceptual overlap between predictor and criterion for those high in conscientiousness. This is of course just speculation, and one point that needs to be considered is that conscientiousness was also significantly and positively related to endorsement of the “Other” response option. It could be argued that more conscientious individuals were more likely to follow instructions and thus were more attentive about endorsing the “Other” response option when they were thinking of themselves in other situations than were less conscientious individuals.

Research Question 6 examined the extent to which factor structure was affected by the use of relatively more or fewer FORs when responding. The first part of the analysis involved evaluating the relationships between the variables representing the number of FORs endorsed for each factor and self-monitoring. The correlations among these variables were very weak, and only two (self-monitoring with FOR endorsement for neuroticism and FOR endorsement for self-monitoring with openness) were statistically significant. In theory, high self-monitors tailor their behavior to fit different situations and so should be more likely to rely on a range of FORs when responding, thus yielding a positive correlation with FOR endorsement. Yet this relationship seems to only hold (weakly) for the neuroticism and openness to

experience factors. One tentative explanation that might be offered is that the relationship between FOR endorsement and self-monitoring only emerged for neuroticism and openness to experience because these two factors appeared to be the most “context-free.” The high level of FOR endorsement for the other three factors indicated that they are relatively context-dependent, and thus the very nature of the items/factors may have implied the relevance of various of FORs. Thus, for agreeableness, conscientiousness, and extraversion the relationship between FOR endorsement and self-monitoring may not have emerged because the use of FORs was more strongly related to the items themselves than it was to differences among individuals. However, it is possible that, because neuroticism and openness to experience items appeared to be relatively less context-dependent, the use of FORs was perhaps somewhat more willful. For neuroticism and openness to experience, then, the participants who endorsed more FORs were individuals who were accustomed to thinking of themselves in multiple contexts and changing their behavior across situations (e.g., self-monitors). Again, given the weak effect sizes, this is offered as a tentative explanation and future studies would need to determine whether the relationships can be replicated. It is also possible that another individual difference such as adaptability (e.g., Wang et al., in press) could be more strongly associated with flexibility in the use of FORs than self-monitoring, or that some measure of interpersonal skill (e.g., extraversion) is associated with such flexibility. In addition, the findings related to Research Question 5 indicate that self-monitoring may be associated with greater use of *certain* FORs (e.g., “with Family,” “With Friends,” and

“At Work”) rather than any/all FOR. Future studies may consider whether other individual differences share a stronger relationship with the use of FORs, and whether self-monitoring exhibits a greater influence in particular situations/contexts.

The central analysis for Research Question 6 involved conducting a series of multi-sample CFAs. Samples were split in half using the overall FOR endorsement variable (e.g., reflecting the total FOR endorsement across factors) and then the FOR endorsement variables for each factor. Models representing no parameter constraints between the two samples were compared with a series of models that added progressively more constraints. The agreeableness and extraversion models showed no significant differences in fit between the tested models, thus it appears that it is reasonable to assume measurement invariance across the two samples for these factors. For conscientiousness, neuroticism, and openness to experience, the most restrictive model in which residual covariances were constrained fit significantly worse than the unconstrained model.

To some extent, these results should not come as a surprise because previous studies examining the factor structure of the Big 5 using contextualized versus non-contextualized inventories have only found differences in error variances and not in factor loadings or covariances (e.g., Bing et al., 2004). Indeed, the final model (constraining residual covariances) was tested to provide a comparison between the results of this study and existing FOR studies, and the results indicate that error variances increase when more FORs are used. It is interesting to note, then, that the question regarding differences in factor structures that result from contextualization of

items has been asked in a number of ways, including the traditional FOR/no FOR comparison groups, the approach used by Robie et al, and the approach used in the present study. And yet, truly meaningful differences in factor structure have yet to emerge. In addition, it is worth pointing out that many researchers (e.g., Bagozzi & Edwards, 1998; Steenkamp & Baumgartner, 1998; Vandenburg & Lance, 2000) consider equality of residual covariances to be a very strict assumption and assert that measurement invariance may be assumed even when it is violated. Thus, while it is informative to know that the use of FORs may affect measurement error, it is possible that this line of research has served its purpose and studies that continue in this vein are unlikely to yield additional insights.

Potential Limitations

Study 2 was intended as a follow-up to Study 1 and set out to address the research questions using a quantitative approach. Despite the insights gained there are also some limitations, the first of which involves retrospective bias. Study 2 participants completed the non-contextualized IPIP personality inventory in its entirety, and afterwards they were asked to indicate which contexts they were referring to while responding. Structuring the survey in this way introduces the possibility of retrospective bias, that is, because participants had to reflect on what they were thinking, their responses may not be a completely accurate representation of their thought process. The alternative to structuring the survey this way was to have the participant respond to an item, and then immediately afterwards indicate the relevant contexts. This follows the methodology used by Bem and Allen (1974), so there is

some precedent for it. However, I chose not to take this approach for the following reason. The concern with setting up the survey this way was that it could actually alter the process of responding. If participants were prompted to endorse relevant contexts after the first item, they would be more likely to be influenced by those contexts when considering the next item, and this pattern would be continually reinforced as they proceeded through the personality inventory. Instead, it seemed preferable for participants to complete the inventory as intended, and then to ask them for the context information retrospectively. The benefits of having participants experience responding to the inventory as intended seemed to outweigh the costs associated with retrospective bias.

Another potentially significant limitation is that demand characteristics could have been placed on participants in the process of endorsing context. It is possible that participants were not considering any of the contexts as they responded to the item(s) but they felt obligated to choose a context given the demands of the experiment. However, this concern figured prominently in the planning of the survey and was addressed in a few ways. First, participants could choose to skip any question they did not want to answer. Therefore, if they did not feel that any of the response options provided were relevant for a particular item, they could simply move to the next. In fact, it does seem that participants may have done this to some extent, given that the percentage of respondents who did not endorse any response options ranged from 0-3% across the 50 items. Secondly, participants were provided with the “In General” response option because it was anticipated based on Study 1 results that they would

have been thinking of themselves in general when responding to a number of the items. Thus, even if participants did not feel that any of the FORs provided were relevant, they had the option to indicate that they were thinking of themselves in general. Finally, participants were also provided with an “Other” option so that they could indicate when none of the contexts provided were relevant to how they were thinking of themselves for a particular item. Although it is possible that demand characteristics still affected responding, several attempts were made to mitigate this influence.

Another limitation of Study 2 involves the use of a student sample. This was also a limitation of Study 1, but efforts were made to increase the generalizability of Study 1 findings by selecting participants who held multiple roles and did not fit the traditional student stereotype. Given the need for a larger sample in Study 2, participants were not screened on the basis of any criteria other than previous work experience (those with none were removed). However, nearly 75% of participants were currently employed at least part-time, about 1/3 were married, and just under half spent time each week caring for parents or children, indicating that this sample is far more diverse in terms of demographics than the traditional student sample.

Chapter 7

General Discussion

The two studies conducted as part of this dissertation examine the role of context used by test-takers in non-contextualized personality inventories. These issues were studied both qualitatively (Study 1) and, based on these results, quantitatively (Study 2). As discussion specific to the results and limitations of each study has been already presented, this chapter will integrate the findings of the two studies, focus on their theoretical and practical implications, and provide directions for future research.

Implications for Personality Research

The two studies have several implications for personality theory and measurement. While it is generally known that external influences and situations affect the self-concept and behavior (see Figure 1), these studies illustrate that it is reasonable to assume this relationship is a relatively conscious component of the process of responding to a personality inventory. Study 1 provided a great deal of insight into the nature of this phenomenon, and Study 2 particularly reinforced the role of context in making evaluations of oneself. As described earlier, Study 1 revealed that participants typically began the response process by first assessing the extent to which the item was an accurate/inaccurate reflection of the way they generally behaved. McCrae and Costa (1996) would describe this process as one of explicitly considering the self-concept (i.e., knowledge, views, and evaluations of the self) and how it translates into behavior. The high frequency of endorsement of the “In General”

response option in Study 2 provides supporting evidence for the central role that the self-concept plays in responding.

After considering the degree to which the item was an accurate reflection of their self-concept and behavior in general, Study 1 participants often began to consider the extent to which their behavior varied across situations. Study 2 results demonstrating the frequent co-occurrence of the “In General” response option and one or more of the FORs imply that this process is quite typical. Thus, while participants seemed to be relying on their self-concept in forming an initial reaction to the item (e.g., “I like to think I’m that way), a complete assessment of their actual *behavior* involved consideration of situations in which their behavior was an accurate or inaccurate reflection of their self-concept. While personality theorists acknowledge that behavior represents an interaction between person and situation, Studies 1 and 2 demonstrate that the interplay between these two forces is consciously considered when individuals respond to a personality inventory.

In addition to demonstrating that person by situation interactions affect responses to non-contextualized personality inventories, together the two studies provide information regarding what situations are considered by test-takers, and to some extent, what the person by situation interactions look like. The VPA methodology used in Study 1 was critical to obtaining an understanding of the range of contexts in which participants thought of themselves while responding. Further, responses from Study 1 participants indicated that certain situations enhanced elements of their self-concept (e.g., I’m usually pretty talkative anyway, but when I’m

with my friends I'm never quiet) whereas in other situations they found themselves behaving in discordance with their self-concept (e.g., I have a strong work ethic, but there are certain parts of my job that I really don't like and will try to get out of doing).

The interplay between person and situation in the response process as it occurred for participants in the two studies is what would be expected on the basis of CAPS theory (Mischel & Shoda, 1997). In CAPS theory terminology, the self-concept is comprised of key interactions among cognitive-affective mediating units that interact with one another in stable and predictable ways. The interactions between cognitive-affective units are different, but predictably so, across situations. Thus, the "essence" of the person is the same; it is only the connections between the units that change. This explains why participants were able to speak about their general behavior, while simultaneously acknowledging that there are situations in which they deviate from that behavior. General behavior and the self-concept is a reflection of key interactions among cognitive-affective units, and changes to that behavior in certain situations reflects differences in the interactions between units.

Yet despite the fact that the response process is compatible with CAPS theory, CAPS theory does little in the way of helping to explain the use of specific FORs in the two studies. According to CAPS theory, predicting behavior is dependent on identifying the psychologically relevant features of situations. If there were a comprehensive taxonomy describing these features, the FORs could be categorized on this basis and predictions could be made. Without that tool, however, this is not

possible. Instead, I believe that trait activation theory (Tett & Guterman, 2000) is a useful lens through which the results regarding the use of specific FORs can be viewed.

Trait activation theory (Tett & Guterman, 2000) is an extension of the basic strong versus weak situation argument. Trait activation theory argues that while situational strength (the extent to which the situation allows for individual variability in behavior) does impact behavior, the more important element is trait relevance. Trait relevance, simply described, is the extent to which a particular trait is relevant to the situation at hand. A strong situation might be created when one is punished for tardiness by being fired. Conversely, a weak situation might be one where there are no extrinsic rewards for punctuality or consequences for tardiness. These situations, whether weak or strong, are highly relevant to the trait of punctuality. The situations are far less relevant to the trait of extraversion, for example. Thus, although it is expected that there will be less variability in behavior in strong situations than in weak situations, if the behavior/trait of interest is not relevant to the situation (be it strong or weak) it is unlikely to be displayed at all, let alone exhibit variability.

Trait activation is helpful in understanding differences in endorsement of the FORs for each factor. Agreeableness serves as a particularly good example here. In both studies, the “With Family” and “With Friends” FOR was used most frequently (after the “In General” response). Trait activation theory can help us understand why. First assume that, according to CAPS theory, participants have considered their behavior in terms of their self-concept and are now considering situations that are

either especially strongly connected to that self-concept, or situations in which their behavior differs. Now consider the strength of the situations reflected in the FORs. In a relative sense, the “At School” and “At Work” FORs are stronger than the “At Home,” “With Friends,” and “With Family” FORs because most people are socialized to display agreeable behaviors in public places like school and work. Thus, it is likely that because agreeable behaviors are strongly tied to public behavior, and because most of one’s life is spent in public, these situations are actually reflected in one’s self-concept. Let us consider the trait relevance of agreeableness to the remaining three situations. Agreeable behaviors (e.g., Have a kind word for everyone) typically require interaction with others. The “At Home” FOR does not specifically imply interaction (as one can be at home alone, or live alone) whereas the “With Friends” and “With Family” FORs do. Thus, by combining both CAPS theory and trait activation theory, it is possible to make better predictions about which situations will be considered in association with certain factors.

Practical Implications

The two studies have a number of practical implications. The primary implication of the findings is that contextualized measures should be used over non-contextualized measures. Results from Study 1 indicate that a consideration of context plays a key role in determining responses to non-contextualized items, and Study 2 further demonstrates that respondents will often consider multiple contexts, across which their behaviors may vary. CAPS theory implies that these differences in behavior are the not the result of chance but rather are stable, predictable patterns.

Thus, the studies provide further support for the importance of providing an “at work” FOR when trying to use personality measures to predict behavior at work.

The implications of the current studies for I/O practitioners who elect not to use a work-specific measure of personality are perhaps even greater than for those who do. The results of the Study 2 are particularly informative here. Participants relied on the “At Work” FOR most frequently when responding to conscientiousness items, followed by agreeableness, extraversion, neuroticism, and openness to experience items. It should come as no surprise that participants were relying heavily on the “At Work” FOR while responding to conscientiousness items; this may explain why research has repeatedly demonstrated that conscientiousness is generally the strongest predictor of work performance of the Big 5 (e.g., Barrick & Mount, 1991). In fact, one could argue that the consistent finding that conscientiousness is the best predictor of work performance among the Big Five factors is less a function of the construct itself than the fact that it is the factor most associated with the “At Work” FOR. Yet the results demonstrated that participants were not thinking exclusively of work when responding, and this warrants further examination even of conscientiousness-based instruments.

Practitioners using neuroticism or openness to experience items should proceed with caution. The results of these studies indicate that respondents are generally not considering their behavior at work when responding to these items, and therefore it is likely inappropriate to be using them for the purposes of selection. Given that respondents were most often simply thinking of themselves “In General” when

responding to neuroticism and openness to experience items, and that we know based on CAPS theory that responses to these items can change in meaningful ways depending on the situation, practitioners should be hesitant to use general tendencies to predict behavior in specific situations.

A related implication points to the importance of incorporating job analytic information when using personality assessments for selection. Job analytic information describes the tasks performed on the job and the necessary knowledge, skills, abilities, and other characteristics that incumbents need to perform those tasks. This information should provide the foundation for establishing the content of a selection tool, and is particularly important when using personality tests. Because we know that behavior can change in meaningful ways depending on the situation, job analytic data should be used to inform to the fullest extent possible the situations that applicants would be facing on the job and personality items should be contextualized to represent those situations. This process should result not only in increase face validity of the measures, but in increased predictive validity as well.

The present studies also have implications for test vendors selling “off-the-shelf” personality assessments. First, these studies imply that vendors need to consider the extent to which their assessments may be asking participants to describe their behavior across a broad range of situations versus a specific situation. The results indicate that agreeableness items may be a good starting point as these items tended to result in greater range of FOR endorsements than other factors. Secondly, vendors would be encouraged to take an in-depth look at the ways in which openness to

experience and neuroticism items are being used. If these scales are being used to predict work performance, care should be taken to contextualize the items because it is unlikely that participants will be considering their behavior in a work context when responding. Third, vendors may want to tailor their tests for specific types of jobs or job contexts, for example, helping others (customer service), solving problems (supervision), or working in teams (firefighting or product development.)

Future Research

The present studies provide a rich basis for future research. Perhaps most importantly, the approach used in Study 2 should be implemented with a sample of regular employees, and the impact of using multiple FORs should be examined. This is critical because although the results of the present studies demonstrate significant differences between FOR use across and within factors, the impact on validity is still unknown. Similarly, one could examine these issues with job applicants, the population of greatest interest among many I/O personality researchers. However, this might be more of a challenge since applicants might feel compelled to respond to research questioning in a socially desirable way. Another interesting future study might involve manipulating the instructions such that participants are told to respond as if applying for a job. Comparing the results of instructions manipulated in this way to the results of the present studies would provide an indication of the extent to which variability in FOR use is a concern when applying for a job. More specifically, if participants repeatedly endorsed the “At Work” FOR, this would diminish concerns around FOR use in non-contextualized inventories being used for selection. In

addition, future studies should address the use of FORs among other non-contextualized personality inventories. Although the IPIP is commonly used among researchers, there are a number of other instruments based on the Big 5 and future research should address how FORs are used when responding to those inventories.

Another potentially fruitful direction would be to examine the interaction between certain personality factors and FOR usage. For example, do extraverts tend to refer more frequently to their behavior “With Friends,” and does the “At Home” FOR have greater resonance with introverts? Heller, Ferris, Brown, and Watson (2009) have demonstrated that people display significant differences in personality based on their roles (e.g., significantly more conscientious at work than home). A study addressing whether these demonstrated differences in behavior translate to FOR usage when responding to personality items would be very informative.

A related study might address the relationship between self-monitoring and FOR usage. An attempt was made in the present studies to correlate self-monitoring and FOR usage, but the relationship was weak. Yet constructs like self-monitoring are likely the most theoretically related construct to FOR usage that is currently available. Another example would be personality characteristics such as adaptability (e.g., Wang et al., in press), or more specifically interpersonal adaptability, which measures the tendency to be flexible and open-minded when dealing with others (Ployhart & Bliese, 2006). Future studies should investigate the potential for empirically demonstrating this relationship.

Another worthwhile line of future research would involve investigating the effect of role salience on the use of FORs. According to Katz and Kahn (1978) every individual holds multiple roles, but certain roles are more salient than others. Role salience reflects the importance of the identity associated with a particular life role (e.g., family, work). With regard to FORs, role salience could be an important predictor of FOR usage. FORs that are associated with salient roles (e.g., work FOR and career role salience, family FOR and family role salience) may be used with greater frequency than FORs that are not associated with those roles.

The correlations between age and FOR use could yield another interesting line of research. One implication of the negative correlations between age and FOR use is that as age increases, FOR use may diminish. Taking these findings a step further, it might be suggested that concerns around utilizing multiple FORs, and the potential negative effects this would have on reliability and validity, might be more problematic for younger individuals than for older individuals. Future studies should more rigorously investigate the relationship between age and FOR usage to determine whether stabilization of personality through the lifespan tends to result in fewer contextualized responses.

Finally, I believe there is much to be gained by working toward the development of a taxonomic structure of work situations. The situations used in the present studies were derived largely based on roles (e.g., “At School” in my role as a student). While there is a solid base of research to support this approach, and researchers such as Heller and colleagues continue to advance it, I would argue that

greater predictive power would be possible if we had a framework for analyzing work situations. This assertion would be supported by both CAPS theory and trait activation theory because both acknowledge that while individuals do have tendencies toward certain behaviors, the extent to which those behaviors are demonstrated depends on whether the situation the individual is in activates features that are psychologically relevant to the trait. A taxonomy of the “psychologically relevant features” of work situations would thus be key to predicting whether an extravert actually displays extraverted behaviors, or whether a neurotic individual suppresses those tendencies.

Meyer (2010) has begun the arduous task of creating a taxonomy of work situations. Using a creative variation on daily diary methodology, Meyer established two orthogonal dimensions of work situations, formality-informality and maintenance-development. He then named each of the four quadrants represented in the 2 X 2 matrix and classified the sample of work situations into one of the four quadrants (see Figure 3). Although these are the results of just the first study and the structure needs to be tested in other samples to determine whether it will replicate, the study represents an exciting beginning. The results of the present studies could be nicely combined with the initial taxonomic structure and used as the basis for the development of a highly predictive personality assessment for the workplace.

Contributions

The present studies contribute to the literature in at least four important ways. First, the present studies provide a greater understanding of the contexts, or FORs, used by respondents when completing a non-contextualized personality inventory. The

present studies represent the first attempt at addressing this basic research question, and the combination of qualitative and quantitative methods provide even greater clarity in addressing this issue. As pointed out by Lievens et al. (2008), despite the growing number of FOR studies, none of these studies addresses the question of how participants respond to a non-contextualized personality inventory. Study 1 represents a true first in this area of research by identifying common FORs used by participants when completing a standard measure of the Big 5. Prior to Study 1, the assumption that FORs do play a role in responding to non-contextualized inventories had not yet been tested. Not only did Study 1 validate this assumption, but it provided a number of key insights regarding how context is used by participants when responding. Specifically, results indicated that respondents think of themselves in general, at home, work, school, when with friends and family, and occasionally in other more specific situations. In addition, Study 1 results demonstrated that the ways in which participants use context in responding vary. For example, sometimes participants mentioned a particular context because it served as a prime example of their typical behavior, whereas at other times a particular context was mentioned because it served as an exception to what they considered to be their general behavior. Thus, Study 1 provides a level of insight into the response process that does not currently exist in the literature.

The second contribution of the present studies involves a deeper understanding of the FORs used in non-contextualized personality inventories and the potentially meaningful ways these FORs differ across items, and across and within individuals.

For example, Study 2 demonstrated that while participants frequently reference their general behavior when responding, their behavior at school and at work is particularly relevant when responding to conscientiousness items. Results also demonstrated that neuroticism and openness to experience items may be less dependent on context than other factors. Not only do the present studies contribute to the literature by providing insight into the process of responding to a non-contextualized inventory, but they expand on those initial findings by understanding the role that context plays for each Big 5 factor. Because no other studies have sought to understand the role of FORs in non-contextual inventories, there are no other studies that can provide the level of detail around FOR use that is achieved in Study 2. In turn, this level of detail has the potential to provide powerful insights for both researchers and practitioners. For example, knowing that conscientiousness items are significantly more likely to elicit an “At Work” FOR than are openness to experience or neuroticism items not only informs future work, but may also shed light on past findings, such as the consistently positive relationship between conscientiousness and work performance.

Third, the present studies utilized multiple methodologies in order to answer the research questions. Personality research in I/O psychology has long been criticized for a reliance on self-report questionnaires (e.g., Hough & Furnham, 2003). In contrast, the present studies relied on qualitative methods in addition to a self-report methodology that allowed participants to elaborate on their responses. Thus the present studies represent an innovative approach to personality research methodology.

Finally, and perhaps most importantly, the present studies represent a person-centric approach to research as advocated by Weiss and Rupp (2010). According to Weiss and Rupp, a person-centric approach to research “will want to take account of the nature of the experience itself, the “what is it like” to work or to experience certain events,” (p. 13) and “a person-centric work is of necessity a work psychology imbued with the self; not the self-as-object but instead the self as source of perspective, the self as ever-present agent” (p. 14). The qualitative component of this dissertation in particular provided a unique contribution to understanding the experience of taking a personality inventory. As one small example, no past work has pointed out that neuroticism items might actually induce anxiety or worry in some test-takers, as was found in Study 1. Moreover, this dissertation demonstrates the different ways that respondents interpret test items. This in turn has implications not only for test-takers themselves, but also for the psychometric properties of such variables. The present studies have demonstrated the value that can be obtained by taking a person-centric approach and considering the individual to be a valuable source of information in and of themselves, rather than simply as one of countless other data points.

Summary

Recently, Heggstad and Gordon (2008) suggested that a FOR should always be provided for personality measures that will be used in a work context. In fact, in the original version of that article the authors had gone so far as to state that providing an “At Work” FOR personality items should be considered a best practice, but the editor requested that the statement be removed. Clearly, then, there is still some hesitation

regarding the use of FORs in personality measures. Perhaps this is because the FOR literature is not in universal agreement regarding finding an increase in the validity of personality measures by providing an “at work” FOR (e.g., Small & Diefendorff, 2006). Further, it was only recently that Lievens et al. (2008) demonstrated that it is actually *within*-person inconsistencies in responding, rather than *between*-person inconsistencies as was previously thought, that drive down the reliability of non-contextualized measures.

The goal of these studies was to take a step back from the standard format of FOR research and try to try and understand the phenomenon from the ground up. To this end, it was quite successful. Study 1 provided insight into the response process involved in completely a non-contextualized personality inventory, provided a range of situations often used by respondents, and also shed light on how people think about variations in their behavior across contexts. Study 2 provided a more detailed and comprehensive picture of the frequency with which certain FORs are endorsed both within and across factors. Together the two studies represent a unique way of considering the FOR issue in personality assessment and have provided insights that will be useful to other personality researchers interested in understanding person by situation interactions.

Note

1. In each of the ANOVAs that follow, Mauchly's test of sphericity was significant, indicating that an adjustment to the degrees of freedom is necessary. Thus, degrees of freedom calculated using the Greenhouse-Geisser correction are reported for all analyses. However, it is worth noting that F ratios remain significant for all analyses regardless of the choice for correcting degrees of freedom.

Table 1. *Research Questions and Methodology for the Proposed Study*

Study Sequence	Study Design	Research Question
1	Verbal protocol analysis	1. Do people think of themselves in particular environments or contexts when responding to a non-contextualized personality inventory? In other words, do individuals use FORs when responding to a non-contextualized personality inventory?
1	Verbal protocol analysis	2. In which contexts do people see themselves when responding to a noncontextualized personality inventory?
2	Correlational	3. Do certain items/factors elicit more or fewer frames of reference than others?
2	Correlational	4. Are there significant differences in the FORs used when responding to non-contextualized items/factors?
2	Correlational	5. Which FORs exhibit the greatest degree of variability in endorsement, and what are the individual differences that are related to endorsement of each FOR/response option?
2	Correlational	6. Does the factor structure of the FFM differ for people who are relatively less or more cross-situationally consistent?

Table 2. *Labels and Definitions for the FFM Domains and Facets (Costa & McCrae, 1992)*

Label	Definition
Neuroticism	Individuals who are prone to psychological distress
N1: Anxiety	Level of free floating anxiety
N2: Angry Hostility	Tendency to experience anger and related states such as frustration and bitterness
N3: Depression	Tendency to experience feelings of guilt, sadness, despondency, and loneliness
N4: Self-Consciousness	Shyness or social anxiety
N5: Impulsiveness	Tendency to act on cravings and urges rather than reigning them in and delaying gratification
N6: Vulnerability	General susceptibility to stress
Extraversion	Quantity and intensity of energy directed outwards into the social world
E1: Warmth	Interest in and friendliness toward others
E2: Gregariousness	Preference for the company of others
E3: Assertiveness	Social ascendancy and forcefulness of expression
E4: Activity	Pace of living
E5: Excitement seeking	Need for environmental stimulation
E6: Positive Emotions	Tendency to experience positive emotions
Openness to Experience	The active seeking and appreciation of experiences for their own sake
O1: Fantasy	Receptivity to the inner world of imagination
O2: Aesthetics	Appreciation of art and beauty
O3: Feelings	Openness to inner feelings and emotions
O4: Actions	Openness to new experiences on a practical level
O5: Ideas	Intellectual curiosity
O6: Values	Readiness to re-examine own values and those of authority figures
Agreeableness	The kinds of interactions an individual prefers, from compassion to tough mindedness
A1: Trust	Belief in the sincerity and good intentions of others
A2: Straightforwardness	Frankness in expression
A3: Altruism	Active concern for the welfare of others
A4: Compliance	Tendency to avoid and quickly resolve interpersonal conflict
A5: Modesty	Tendency to down play one's own achievements and be humble
A6: Tender mindedness	Attitude of sympathy toward others

table continues

Conscientiousness	Degree of organization, persistence, control and motivation in goal-directed behavior
C1: Competence	Belief in one's own self-efficacy
C2: Order	Personal organization
C3: Dutifulness	Emphasis placed on importance of fulfilling moral obligations
C4: Achievement Striving	Need for personal achievement and sense of direction
C5: Self Discipline	Capacity to begin tasks and follow through to completion despite boredom or distractions
C6: Deliberation	Tendency to think things through before acting or speaking

Table 3. *Examples of Key Quotes for Each Study 1 FOR Category*

FOR Category	Key Quotes
No elaboration	Item: Panic easily. "That is very inaccurate."
Non-contextual	Item: Don't like to draw attention to myself. "Accurate, if I don't have to be noticed, I don't want to be noticed." Item: Am always prepared. "Very accurate. Do I have to explain that? Because I don't really know how. I just am."
Specific situation	Item: Am skilled in handling social situations. "Mostly accurate, but I don't really like public speaking." Item: Insult people. "I was driving and someone cut me off...I called them a name. That does happen sometimes."
School FOR	Item: Am not interested in abstract ideas. "I'm a psych major, so in class I pretty much have to be [interest in abstract ideas]." Item: Am always prepared. "It really depends. At school, yes definitely. Always. In family situations, hmm, not always. It really just depends on the context."
Work FOR	Item: Shirk my duties. "Sometimes at work if I can get away with not doing something I will." Item: Suspect hidden motives in others. "I work with at risk kids, so you always have to be on your toes."
Family FOR	Item: Would describe my experiences as somewhat dull. "I have lots of interesting experiences due to my family. My father was just diagnosed with prostate cancer, that's one example." Item: Avoid philosophical discussions. "Yeah for the most part, but my brother is a philosophy major, so we always get into these ridiculous discussions."
Friends FOR	Item: Am the life of the party. "Whenever my friends throw parties I'm always the bartender and the focus of attention." Item: Panic easily. "My friends tend to freak out easily, so I'm usually the one that keeps a cool head."
Home FOR	Item: Find it difficult to get down to work. "Not usually, but if I'm at home and there is a good TV show on, then sometimes it is." Item: Get my chores done right away. "That's accurate...when I think about my home life, anyway."

FOR Category	Key Quotes
Significant	Item: Carry out my plans. “For the most part, I do. The only thing that makes me hesitate is that my boyfriend isn’t very plan-oriented, so anytime I’m doing things with him it might not work out like I intend.”
Other FOR	Item: Often feel blue. “Very inaccurate. Even when I’m in very sad situations, like when I broke up with my boyfriend... I just thought, this could be a new opportunity for me, and I shouldn’t be too sad.”

Table 4. *Percent of FORs Appearing within Each Factor for Study 1*

FOR Category	<u>Agreeableness</u>		<u>Conscientiousness</u>		<u>Extraversion</u>		<u>Neuroticism</u>		<u>Openness to Experience</u>	
	All Coded		All Coded		All Coded		All Coded		All Coded	
	Responses as Denominator	Just FORs as Denominator	Responses as Denominator	Just FORs as Denominator	Responses as Denominator	Just FORs as Denominator	Responses as Denominator	Just FORs as Denominator	Responses as Denominator	Just FORs as Denominator
Non-contextualized	65%	--	60%	--	63%	--	57%	--	59%	--
No Elaboration	15%	--	12%	--	12%	--	24%	--	18%	--
Family	5%	25%	3%	10%	2%	7%	2%	9%	3%	14%
Friends	4%	21%	1%	4%	6%	23%	2%	8%	2%	8%
Home	3%	15%	3%	12%	1%	4%	4%	22%	2%	9%
School	1%	4%	11%	40%	6%	23%	3%	15%	8%	33%
Significant Other	1%	4%	1%	5%	2%	7%	1%	3%	1%	4%
Work	4%	21%	6%	22%	4%	16%	4%	21%	3%	15%
Specific Situation	2%	10%	2%	7%	5%	20%	4%	22%	4%	17%

Table 5. *Item Order for Each Survey Version*

Version 1	Version 2	Version 3
Have a good word for everyone.	Have a good word for everyone.	Don't see things through.
Believe that others have good intentions.	Believe that others have good intentions.	Don't like to draw attention to myself.
Respect others.	Respect others.	Don't talk a lot.
Accept people as they are.	Accept people as they are.	Am not easily bothered by things.
Make people feel at ease.	Make people feel at ease.	Am not interested in abstract ideas.
Have a sharp tongue.	Am always prepared.	Do not like art.
Cut others to pieces.	Pay attention to details.	Do not enjoy going to art museums.
Suspect hidden motives in others.	Get chores done right away.	Have a sharp tongue.
Get back at others.	Carry out my plans.	Cut others to pieces.
Insult people.	Make plans and stick to them.	Suspect hidden motives in others.
Am always prepared.	Feel comfortable around people.	Get back at others.
Pay attention to details.	Make friends easily.	Insult people.
Get chores done right away.	Am skilled in handling social situations.	Waste my time.
Carry out my plans.	Am the life of the party.	Find it difficult to get down to work.
Make plans and stick to them.	Know how to captivate people.	Do just enough work to get by.
Waste my time.	Often feel blue.	Try to avoid my duties.
Find it difficult to get down to work.	Dislike myself.	Have little to say.
Do just enough work to get by.	Am often down in the dumps.	Keep in the background.
Don't see things through.	Have frequent mood swings.	Would describe my experiences as somewhat dull
Try to avoid my duties.	Panic easily.	Avoid philosophical discussions.
Feel comfortable around people.	Believe in the importance of art.	Tend to vote for conservative political candidates.
Make friends easily.	Have a vivid imagination.	Rarely get irritated.
Am skilled in handling social situations.	Tend to vote for liberal political candidates.	Seldom feel blue.
Am the life of the party.	Carry the conversation to a higher level.	Feel comfortable with myself.
Know how to captivate people.	Enjoy hearing new ideas.	Am very pleased with myself.

table continues

Version 1	Version 2	Version 3
Have little to say.	Have a sharp tongue.	Have a good word for everyone.
Keep in the background.	Cut others to pieces.	Believe that others have good intentions.
Would describe my experiences as somewhat dull.	Suspect hidden motives in others.	Respect others.
Don't like to draw attention to myself.	Get back at others.	Accept people as they are.
Don't talk a lot.	Insult people.	Make people feel at ease.
Often feel blue.	Waste my time.	Am always prepared.
Dislike myself.	Find it difficult to get down to work.	Pay attention to details.
Am often down in the dumps.	Do just enough work to get by.	Get chores done right away.
Have frequent mood swings.	Try to avoid my duties.	Carry out my plans.
Panic easily.	Have little to say.	Make plans and stick to them.
Rarely get irritated.	Keep in the background.	Feel comfortable around people.
Seldom feel blue.	Would describe my experiences as somewhat dull.	Make friends easily.
Feel comfortable with myself.	Avoid philosophical discussions.	Am skilled in handling social situations.
Am not easily bothered by things.	Tend to vote for conservative political candidates.	Am the life of the party.
Am very pleased with myself.	Rarely get irritated.	Know how to captivate people.
Believe in the importance of art.	Seldom feel blue.	Often feel blue.
Have a vivid imagination.	Feel comfortable with myself.	Dislike myself.
Tend to vote for liberal political candidates.	Am very pleased with myself.	Am often down in the dumps.
Carry the conversation to a higher level.	Don't see things through.	Have frequent mood swings.
Enjoy hearing new ideas.	Don't like to draw attention to myself.	Panic easily.
Am not interested in abstract ideas.	Don't talk a lot.	Believe in the importance of art.
Do not like art.	Am not easily bothered by things.	Have a vivid imagination.
Avoid philosophical discussions.	Am not interested in abstract ideas.	Tend to vote for liberal political candidates.
Do not enjoy going to art museums.	Do not like art.	Carry the conversation to a higher level.
Tend to vote for conservative political candidates.	Do not enjoy going to art museums.	Enjoy hearing new ideas.

Table 6. Means, Standard Deviations, and Confidence Intervals for Each FOR within Each Item

Item Text	Mean	Std. Deviation	Percent of Ss Endorsing Each Response Option for Each Item						
			With Family	With Friends	At Home	At School	At Work	Other	In General
Agreeableness									
Have a good word for everyone.	2.18	1.90	43%	55%	38%	36%	41%	4%	63%
Believe that others have good intentions.	1.95	1.83	44%	50%	35%	33%	29%	5%	58%
Respect others.	2.45	2.27	49%	53%	44%	49%	45%	5%	80%
Accept people as they are.	2.32	2.22	46%	51%	42%	43%	43%	6%	76%
Make people feel at ease.	2.20	2.08	42%	55%	37%	39%	42%	5%	68%
Have a sharp tongue.	1.47	1.46	37%	40%	31%	13%	19%	7%	33%
Cut others to pieces.	0.90	1.27	16%	24%	13%	12%	14%	11%	32%
Suspect hidden motives in others.	1.24	1.34	14%	27%	9%	28%	37%	9%	35%
Get back at others.	0.77	1.19	14%	21%	10%	10%	12%	11%	34%
Insult people.	0.84	1.23	14%	23%	11%	11%	13%	13%	35%
Conscientiousness									
Am always prepared.	1.97	1.59	25%	26%	27%	61%	56%	2%	44%
Pay attention to details.	2.14	1.83	29%	34%	34%	58%	55%	3%	52%
Get chores done right away.	1.58	1.27	16%	13%	53%	31%	42%	3%	28%
Carry out my plans.	2.16	1.85	37%	47%	35%	49%	44%	4%	54%
Make plans and stick to them.	1.93	1.82	38%	43%	28%	42%	37%	5%	53%
Waste my time.	1.33	1.07	7%	17%	53%	32%	17%	6%	23%
Find it difficult to get down to work.	1.29	1.06	8%	13%	44%	41%	18%	5%	19%
Do just enough work to get by.	1.00	1.03	5%	6%	19%	41%	22%	7%	21%
Don't see things through.	1.00	1.19	11%	16%	24%	27%	14%	8%	28%
Try to avoid my duties.	1.06	1.15	10%	7%	35%	30%	17%	7%	21%

table continues

Item Text	Mean	Std. Deviation	Percent of Ss Endorsing Each Response Option for Each Item						
			With Family	With Friends	At Home	At School	At Work	Other	In General
Extraversion									
Feel comfortable around people.	2.28	1.98	46%	55%	42%	41%	41%	4%	64%
Make friends easily.	1.79	1.78	24%	42%	24%	45%	40%	5%	63%
Am skilled in handling social situations.	2.02	1.91	32%	51%	30%	43%	43%	3%	61%
Am the life of the party.	1.33	1.39	27%	51%	19%	14%	13%	8%	35%
Know how to captivate people.	1.65	1.77	30%	45%	23%	28%	33%	6%	54%
Have little to say.	1.17	1.23	16%	18%	10%	41%	24%	8%	29%
Keep in the background.	1.08	1.14	11%	16%	7%	47%	20%	7%	27%
Would describe my experiences as somewhat dull.	0.94	1.35	13%	14%	16%	24%	19%	8%	36%
Don't like to draw attention to myself.	1.16	1.21	9%	17%	7%	46%	29%	7%	37%
Don't talk a lot.	1.11	1.20	12%	13%	10%	48%	21%	7%	33%
Neuroticism									
Often feel blue.	0.88	1.18	12%	7%	31%	17%	11%	9%	39%
Dislike myself.	0.63	1.10	8%	8%	17%	11%	8%	10%	41%
Am often down in the dumps.	0.63	1.03	7%	5%	22%	12%	8%	9%	40%
Have frequent mood swings.	0.84	1.23	16%	11%	25%	11%	11%	10%	43%
Panic easily.	0.91	1.28	9%	8%	16%	29%	18%	10%	41%
Rarely get irritated.	1.61	1.74	27%	38%	34%	29%	28%	5%	50%
Seldom feel blue.	1.30	1.69	24%	29%	31%	22%	19%	5%	56%
Feel comfortable with myself.	1.77	2.02	34%	41%	39%	29%	30%	4%	69%
Am not easily bothered by things.	1.46	1.81	24%	32%	28%	28%	29%	5%	56%
Am very pleased with myself.	1.59	1.94	29%	32%	33%	30%	28%	6%	67%

table continues

Item Text	Mean	Std. Deviation	Percent of Ss Endorsing Each Response Option for Each Item						
			With Family	With Friends	At Home	At School	At Work	Other	In General
<u>Openness to Experience</u>									
Believe in the importance of art.	1.17	1.73	19%	23%	24%	30%	17%	4%	75%
Have a vivid imagination.	1.24	1.74	21%	24%	27%	27%	20%	5%	72%
Tend to vote for liberal political candidates.	0.66	1.44	10%	11%	13%	12%	10%	10%	68%
Carry the conversation to a higher level.	1.64	1.74	31%	44%	26%	33%	25%	5%	58%
Enjoy hearing new ideas.	1.92	2.10	33%	42%	32%	46%	36%	4%	76%
Am not interested in abstract ideas.	0.68	1.14	8%	11%	10%	19%	12%	8%	38%
Do not like art.	0.47	0.98	6%	6%	7%	11%	6%	11%	44%
Avoid philosophical discussions.	0.91	1.23	17%	19%	9%	25%	16%	6%	32%
Do not enjoy going to art museums.	0.50	1.05	9%	8%	6%	10%	6%	10%	45%
Tend to vote for conservative political candidates.	0.39	0.96	6%	4%	6%	5%	5%	13%	46%

Note. Means range from 0 – 6.

Table 7. *RQ3: Factor Level Mean Differences in Total FOR Endorsement*

Item Text	Est Marginal Mean	Standard Error	1	2	3	4	5
1 Agreeableness	16.32	0.53					
2 Conscientiousness	15.45	0.42	0.87				
3 Extraversion	14.53	0.46	1.79 *	0.92			
4 Neuroticism	11.61	0.48	4.71 *	3.84 *	2.92 *		
5 Openness to Experience	9.59	0.44	6.73 *	5.86 *	4.94 *	2.02 *	

Note. Means can range from 0 (did not endorse any contexts) to 60 (endorsed all contexts for all items within a factor).

* $p < .05$

Table 8. *RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Total FOR Endorsement across Agreeableness Items*

Item Text	Est Marginal		Standard									
	Mean	Error	1	2	3	4	5	6	7	8	9	
1 Have a good word for everyone	2.18	0.09										
2 Believe that others have	1.95	0.09	0.23									
3 Respect others	2.45	0.11	-0.27	-0.50 *								
4 Accept people as they are	2.32	0.11	-0.14	-0.37 *	0.14							
5 Make people feel at ease	2.20	0.10	-0.02	-0.25	0.25	0.12						
6 Have a sharp tongue	1.47	0.07	0.71 *	0.49 *	0.99 *	0.85 *	0.74 *					
7 Cut others to pieces	0.90	0.06	1.28 *	1.05 *	1.55 *	1.42 *	1.30 *	0.57 *				
8 Suspect hidden motives in others	1.24	0.07	0.94 *	0.72 *	1.22 *	1.08 *	0.96 *	0.23	-0.34 *			
9 Get back at others	0.77	0.06	1.41 *	1.18 *	1.68 *	1.54 *	1.43 *	0.69 *	0.13	0.46 *		
10 Insult people	0.84	0.06	1.34 *	1.11 *	1.61 *	1.48 *	1.36 *	0.62 *	0.06	0.39 *	-0.07	

Note. Means can range from 0 (did not endorse any context for an item) to 6 (endorsed all contexts for an item).

* $p < .05$

Table 9. RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Total FOR Endorsement across Conscientiousness Items

Item Text	Est Marginal		Standard								
	Mean	Error	1	2	3	4	5	6	7	8	9
1 Am always prepared.	1.97	0.08									
2 Pay attention to details.	2.14	0.09	-0.17								
3 Get chores done right away.	1.58	0.06	0.39 *	0.56 *							
4 Carry out my plans.	2.16	0.09	-0.20	-0.02	-0.59 *						
5 Make plans and stick to them.	1.93	0.09	0.04	0.21	-0.35 *	0.23 *					
6 Waste my time.	1.33	0.05	0.64 *	0.81 *	0.25	0.84 *	0.60 *				
7 Find it difficult to get down to work.	1.29	0.05	0.68 *	0.85 *	0.29 *	0.88 *	0.64 *	0.04			
8 Do just enough work to get by.	1.00	0.05	0.97 *	1.14 *	0.58 *	1.17 *	0.93 *	0.33 *	0.29 *		
9 Don't see things through.	1.00	0.06	0.97 *	1.14 *	0.58 *	1.17 *	0.93 *	0.33 *	0.29 *	0.00	
10 Try to avoid my duties.	1.06	0.06	0.91 *	1.08 *	0.52 *	1.11 *	0.87 *	0.27 *	0.23 *	-0.06	-0.06

Note. Means can range from 0 (did not endorse any context for an item) to 6 (endorsed all contexts for an item).

* $p < .05$

Table 10. RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Total FOR Endorsement across Extraversion Items

Item Text	Est Marginal		Standard									
	Mean	Error	1	2	3	4	5	6	7	8	9	
1 Feel comfortable around people.	2.28	0.10										
2 Make friends easily.	1.79	0.09	0.48 *									
3 Am skilled in handling social situations.	2.02	0.09	0.26 *	-0.22 *								
4 Am the life of the party.	1.33	0.07	0.95 *	0.46 *	0.69 *							
5 Know how to captivate people.	1.65	0.09	0.63 *	0.14	0.37 *	-0.32 *						
6 Have little to say.	1.17	0.06	1.11 *	0.62 *	0.85 *	0.16	0.48 *					
7 Keep in the background.	1.08	0.06	1.20 *	0.71 *	0.94 *	0.25	0.57 *	0.09				
8 Would describe my experiences as somewhat dull.	0.94	0.07	1.33 *	0.85 *	1.07 *	0.39 *	0.71 *	0.23 *	0.14			
9 Don't like to draw attention to myself.	1.16	0.06	1.12 *	0.64 *	0.86 *	0.17	0.49 *	0.01	-0.08	-0.21		
10 Don't talk a lot.	1.11	0.06	1.17 *	0.68 *	0.91 *	0.22	0.54 *	0.06	-0.03	-0.17	0.05	

Note. Means can range from 0 (did not endorse any context for an item) to 6 (endorsed all contexts for an item).

* $p < .05$

Table 11. RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Total FOR Endorsement across Neuroticism Items

Item Text	Est Marginal		Standard								
	Mean	Error	1	2	3	4	5	6	7	8	9
1 Often feel blue.	0.88	0.06									
2 Dislike myself.	0.63	0.05	0.25 *								
3 Am often down in the dumps.	0.63	0.05	0.25 *	0.00							
4 Have frequent mood swings.	0.84	0.06	0.04	-0.20 *	-0.21 *						
5 Panic easily.	0.91	0.06	-0.03	-0.28 *	-0.28 *	-0.08					
6 Rarely get irritated.	1.61	0.09	-0.74 *	-0.98 *	-0.99 *	-0.78 *	-0.70 *				
7 Seldom feel blue.	1.30	0.08	-0.42 *	-0.67 *	-0.67 *	-0.46 *	-0.39 *	0.32 *			
8 Feel comfortable with myself.	1.77	0.10	-0.89 *	-1.13 *	-1.14 *	-0.93 *	-0.85 *	-0.15	-0.47 *		
9 Am not easily bothered by things.	1.46	0.09	-0.58 *	-0.83 *	-0.83 *	-0.63 *	-0.55 *	0.15	-0.16	0.31	
10 Am very pleased with myself.	1.59	0.09	-0.71 *	-0.96 *	-0.96 *	-0.75 *	-0.68 *	0.03	-0.29 *	0.18 *	-0.13

Note. Means can range from 0 (did not endorse any context for an item) to 6 (endorsed all contexts for an item).

* $p < .05$

Table 12. *RQ3: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Total FOR Endorsement across Openness to Experience Items*

Item Text	Est Marginal	Standard										
	Mean	Error	1	2	3	4	5	6	7	8	9	
1 Believe in the importance of art.	1.17	1.73										
2 Have a vivid imagination.	1.24	1.74	-0.07									
3 Tend to vote for liberal political candidates.	0.66	1.44	0.51	0.58 *								
4 Carry the conversation to a higher level.	1.64	1.74	-0.47 *	-0.40 *	-0.98 *							
5 Enjoy hearing new ideas.	1.92	2.10	-0.75 *	-0.68 *	-1.26 *	-0.28 *						
6 Am not interested in abstract ideas.	0.68	1.14	0.49 *	0.56 *	-0.02	0.96 *	1.24 *					
7 Do not like art.	0.47	0.98	0.70 *	0.76 *	0.19	1.16 *	1.45 *	0.21 *				
8 Avoid philosophical discussions.	0.91	1.23	0.26	0.33 *	-0.25	0.73 *	1.01 *	-0.23 *	-0.44 *			
9 Do not enjoy going to art museums.	0.50	1.05	0.67 *	0.74 *	0.16	1.14 *	1.42 *	0.18	-0.03	0.41 *		
10 Tend to vote for conservative political candidates.	0.39	0.96	0.78 *	0.85 *	0.27 *	1.25 *	1.53 *	0.29 *	0.08	0.52 *	0.11	

Note. Means can range from 0 (did not endorse any context for an item) to 6 (endorsed all contexts for an item).

* $p < .05$

Table 13. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for FORs within Agreeableness

Response Option	Est Marginal Mean	Standard Error	1	2	3	4	5	6
1 With Friends	3.19	0.12						
2 With Family	3.99	0.14	-0.80 *					
3 In General	5.13	0.13	-1.94 *	-1.14 *				
4 At Home	2.69	0.12	0.49 *	1.30 *	2.43 *			
5 Other	0.75	0.08	2.43 *	3.24 *	4.38 *	1.94 *		
6 At School	2.75	0.12	0.44 *	1.24 *	2.38 *	-0.05	-2.00 *	
7 At Work	2.95	0.12	0.24	1.05 *	2.18 *	-0.25	-2.19 *	-0.20

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 14. *RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for FORs within Conscientiousness*

Response Option	Est Marginal Mean	Standard Error	1	2	3	4	5	6
1 With Friends	1.88	0.09						
2 With Family	2.21	0.10	-0.33 *					
3 In General	3.44	0.13	-1.56 *	-1.23 *				
4 At Home	3.53	0.12	-1.65 *	-1.32 *	-0.09			
5 Other	0.49	0.06	1.39 *	1.72 *	2.96 *	3.04 *		
6 At School	4.12	0.13	-2.24 *	-1.91 *	-0.68 *	-0.59 *	-3.63 *	
7 At Work	3.22	0.12	-1.34 *	-1.01 *	0.22	0.31	-2.74 *	0.90 *

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 15. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for FORs within Extraversion

Response Option	Est Marginal	Standard						
	Mean	Error	1	2	3	4	5	6
1 With Friends	2.21	0.11						
2 With Family	3.22	0.12	-1.01 *					
3 In General	4.40	0.14	-2.19 *	-1.17 *				
4 At Home	1.88	0.10	0.33 *	1.35 *	2.52 *			
5 Other	0.63	0.08	1.58 *	2.59 *	3.77 *	1.25 *		
6 At School	3.76	0.13	-1.55 *	-0.54 *	0.64 *	-1.88 *	-3.13 *	
7 At Work	2.82	0.12	-0.61 *	0.40 *	1.58 *	-0.95 *	-2.19 *	0.94 *

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 16. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for FORs within Neuroticism

Response Option	Est Marginal Mean	Standard Error	1	2	3	4	5	6
1 With Friends	1.91	0.10						
2 With Family	2.12	0.11	-0.21					
3 In General	5.02	0.17	-3.12 *	-2.90 *				
4 At Home	2.77	0.13	-0.87 *	-0.65 *	2.25 *			
5 Other	0.72	0.09	1.19 *	1.40 *	4.30 *	2.06 *		
6 At School	2.19	0.11	-0.28 *	-0.07	2.83 *	0.59 *	-1.47 *	
7 At Work	1.90	0.10	0.00	0.21	3.12 *	0.87 *	-1.19 *	0.28 *

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 17. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for FORs within Openness to Experience

Response Option	Est Marginal	Standard						
	Mean	Error	1	2	3	4	5	6
1 With Friends	1.59	0.09						
2 With Family	1.92	0.10	-0.33 *					
3 In General	5.54	0.15	-3.95 *	-3.62 *				
4 At Home	1.61	0.10	-0.02	0.31 *	3.93 *			
5 Other	0.75	0.08	0.84 *	1.17 *	4.79 *	0.86 *		
6 At School	2.18	0.11	-0.59 *	-0.26 *	3.36 *	-0.57 *	-1.43 *	
7 At Work	1.52	0.09	0.07	0.40 *	4.02 *	0.09	-0.77 *	0.66 *

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 18. *RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Factors within the In General Response Option*

Factor	Est Marginal	Standard				
	Mean	Error	1	2	3	4
1 Agreeableness	5.13	0.13				
2 Conscientiousness	3.44	0.13	1.69 *			
3 Extraversion	4.40	0.14	0.73 *	-0.96 *		
4 Neuroticism	5.02	0.17	0.11	-1.58 *	-0.62 *	
5 Openness	5.54	0.15	-0.41 *	-2.10 *	-1.14 *	-0.52 *

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 19. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Factors within the Family FOR

Factor	Est Marginal	Standard				
	Mean	Error	1	2	3	4
1 Agreeableness	3.19	0.12				
2 Conscientiousness	1.88	0.09	1.31 *			
3 Extraversion	2.21	0.11	0.97 *	-0.33 *		
4 Neuroticism	1.91	0.10	1.28 *	-0.03	0.31 *	
5 Openness	1.59	0.09	1.59 *	0.29 *	0.62 *	0.31 *

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 20. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Factors within the Friends FOR

Factor	Est Marginal	Standard				
	Mean	Error	1	2	3	4
1 Agreeableness	3.99	0.14				
2 Conscientiousness	2.21	0.10	1.78 *			
3 Extraversion	3.22	0.12	0.77 *	-1.01 *		
4 Neuroticism	2.12	0.11	1.87 *	0.09	1.11 *	
5 Openness	1.92	0.10	2.07 *	0.29	1.30 *	0.20

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 21. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Factors within the Home FOR

Factor	Est Marginal	Standard				
	Mean	Error	1	2	3	4
1 Agreeableness	2.69	0.12				
2 Conscientiousness	3.53	0.12	-0.84 *			
3 Extraversion	1.88	0.10	0.82 *	1.65 *		
4 Neuroticism	2.77	0.13	-0.08	0.76 *	-0.90 *	
5 Openness	1.61	0.10	1.08 *	1.92 *	0.27 *	1.16 *

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 22. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Factors within the School FOR

Factor	Est Marginal Mean	Standard Error	1	2	3	4
1 Agreeableness	2.75	0.12				
2 Conscientiousness	4.12	0.13	-1.37 *			
3 Extraversion	3.76	0.13	-1.01 *	0.36		
4 Neuroticism	2.19	0.11	0.56 *	1.93 *	1.57 *	
5 Openness	2.18	0.11	0.57 *	1.94 *	1.58 *	0.00

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 23. RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Factors within the Work FOR

Factor	Est Marginal	Standard				
	Mean	Error	1	2	3	4
1 Agreeableness	2.95	0.12				
2 Conscientiousness	3.22	0.12	-0.28			
3 Extraversion	2.82	0.12	0.12	0.40 *		
4 Neuroticism	1.90	0.10	1.04 *	1.32 *	0.92 *	
5 Openness	1.52	0.09	1.42 *	1.70 *	1.30 *	0.38 *

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 24. *RQ4: Estimated Marginal Mean, Standard Errors, and Pairwise Mean Differences for Factors within the Other Response Option*

Factor	Est Marginal Mean	Standard Error	1	2	3	4
1 Agreeableness	0.75	0.08				
2 Conscientiousness	0.49	0.06	0.27 *			
3 Extraversion	0.63	0.08	0.12	-0.14		
4 Neuroticism	0.72	0.09	0.04	-0.23 *	-0.09	
5 Openness	0.75	0.08	0.00	-0.27 *	-0.12	-0.03

Note. Means can range from 0 (did not endorse the response option for any item within the factor) to 10 (endorsed the response option for all items within the factor).

* $p < .05$

Table 25. Descriptive Statistics for Endorsement of Response Options

Response Option	Mean	Std. Deviation	Coefficient of Variation	Minimum	25th Percentile	50th percentile	75th Percentile	Maximum
Family	10.78	8.24	0.76	0.00	3.00	10.00	17.00	50.00
Friends	13.47	8.90	0.66	0.00	6.00	13.00	19.00	50.00
Home	12.49	9.29	0.74	0.00	5.00	11.00	19.00	50.00
School	15.00	9.34	0.62	0.00	8.00	15.00	21.00	49.00
Work	12.42	9.26	0.75	0.00	4.00	11.00	20.00	41.00
In General	23.53	12.34	0.52	0.00	15.00	21.00	33.00	50.00
Other	3.34	6.71	2.01	0.00	0.00	0.00	3.00	41.00

Note. Means can range from 0 (did not endorse the response option for any items) to 50 (endorsed the response option for all items).

Table 26. Correlations between Response Option Endorsement Frequency and Individual Differences

Variables	Family FOR endorsement	Friends FOR endorsement	Home FOR endorsement	School FOR endorsement	Work FOR endorsement	In General endorsement	Other endorsement
Age	-0.12 *	-0.21 **	-0.08	-0.15 **	-0.06	0.12 *	0.11 *
Gender	0.04	0.07	0.06	0.05	0.08	-0.04	0.06
Years of work experience	-0.11 *	-0.19 **	-0.09	-0.14 **	-0.04	0.07	0.06
Employed	-0.02	0.00	-0.06	-0.05	-0.14 **	0.01	0.11 *
Self-monitoring	0.13 *	0.13 **	0.08	0.09	0.12 *	-0.05	-0.08
Agreeableness	0.00	-0.02	0.03	0.00	0.04	0.09	0.00
Conscientiousness	0.06	0.04	0.06	-0.01	0.12 *	-0.02	0.11 *
Extraversion	0.06	0.05	0.03	-0.06	0.07	-0.13 **	0.08
Neuroticism	0.01	-0.04	-0.02	-0.02	-0.05	-0.03	0.01
Openness to experience	-0.03	-0.01	0.03	-0.03	-0.02	0.09	-0.04

Note. Gender is coded as male = 1, female = 0. Employed is coded as 1 = currently employed and 0 = not currently employed.

* $p < .05$

** $p < .01$

Table 27. Results of CFA models.

Model	χ^2	df	$\Delta\chi^2$	df $\Delta\chi^2$	p-value	CFI	RMSEA
<u>Agreeableness</u>							
Unconstrained model	137.36	57.00	---	---	---	0.93	0.06
Model 1: Factor loadings constrained	145.74	65.00	8.38	8.00	0.40	0.93	0.06
Model 2: Factor loadings and covariances constrained	146.59	66.00	9.23	9.00	0.42	0.93	0.05
Model 3: Variances and covariances of residuals constrained	164.20	83.00	26.85	26.00	0.42	0.93	0.05
<u>Conscientiousness</u>							
Unconstrained model	134.24	62.00	---	---	---	0.96	0.05
Model 1: Factor loadings constrained	138.42	71.00	4.18	9.00	0.90	0.96	0.05
Model 2: Factor loadings and covariances constrained	147.92	72.00	13.68	10.00	0.19	0.95	0.05
<u>Extraversion</u>							
Unconstrained model	149.28	54.00	---	---	---	0.95	0.07
Model 1: Factor loadings constrained	154.92	63.00	5.64	9.00	0.78	0.95	0.06
Model 2: Factor loadings and covariances constrained	155.36	64.00	6.08	10.00	0.81	0.95	0.06
Model 3: Variances and covariances of residuals constrained	184.86	82.00	35.58	28.00	0.15	0.95	0.06
<u>Neuroticism</u>							
Unconstrained model	201.93	56.00	---	---	---	0.93	0.08
Model 1: Factor loadings constrained	218.43	65.00	16.50	9.00	0.06	0.93	0.08
Model 2: Factor loadings and covariances constrained	218.73	66.00	16.80	10.00	0.08	0.93	0.07
Model 3: Variances and covariances of residuals constrained	252.01	83.00	50.08	27.00	0.00	0.92	0.07

table continues

Model	χ^2	df	$\Delta\chi^2$	df $\Delta\chi^2$	<i>p</i> -value	CFI	RMSEA
<u>Openness</u>							
Unconstrained model	120.59	62.00	---	---	---	0.96	0.05
Model 1: Factor loadings constrained	136.27	71.00	15.68	9.00	0.07	0.95	0.05
Model 2: Factor loadings and covariances constrained	136.83	72.00	16.24	10.00	0.09	0.95	0.05
Model 3: Variances and covariances of residuals constrained	158.66	83.00	38.07	21.00	0.01	0.95	0.05

Table 28. Means, SDs, and t-test Results for Differences in Mean FOR Endorsement Based on Actual Survey Response Context

Dependent Variables	Group	N	Mean	SD	t- test
<u>"With Family" while responding</u>					
"With Family" FOR endorsement	Not with family	395	10.63	8.02	$t(417) = -1.52, ns$
	With family	24	13.25	11.19	
"With Friends" FOR endorsement	Not with family	395	13.37	8.69	$t(417) = -0.87, ns$
	With family	24	15.00	11.95	
"At Home" FOR endorsement	Not with family	395	12.35	9.17	$t(417) = -1.27, ns$
	With family	24	14.83	11.00	
"Other" FOR endorsement	Not with family	395	3.30	6.70	$t(417) = -0.46, ns$
	With family	24	3.96	6.94	
"At School" FOR endorsement	Not with family	395	14.87	9.16	$t(417) = -1.19, ns$
	With family	24	17.21	11.87	
"At Work" FOR endorsement	Not with family	395	12.39	9.26	$t(417) = -0.29, ns$
	With family	24	12.96	9.41	
<u>"With Friends" while responding</u>					
"With Family" FOR endorsement	Not with friends	401	10.69	8.28	$t(417) = -1.11, ns$
	With friends	18	12.89	6.99	
"With Friends" FOR endorsement	Not with friends	401	13.32	8.91	$t(417) = -1.59, ns$
	With friends	18	16.72	8.07	
"At Home" FOR endorsement	Not with friends	401	12.41	9.33	$t(417) = -0.81, ns$
	With friends	18	14.22	8.32	
"Other" FOR endorsement	Not with friends	401	3.32	6.70	$t(417) = -0.32, ns$
	With friends	18	3.83	7.09	
"At School" FOR endorsement	Not with friends	401	14.91	9.38	$t(417) = -1.01, ns$
	With friends	18	17.17	8.19	
"At Work" FOR endorsement	Not with friends	401	12.34	9.29	$t(417) = -0.82, ns$
	With friends	18	14.17	8.52	
<u>"At Home" while responding</u>					
"With Family" FOR endorsement	Not at home	135	10.50	7.66	$t(417) = -0.49, ns$
	At home	284	10.92	8.51	
"With Friends" FOR endorsement	Not at home	135	13.39	8.67	$t(417) = -0.13, ns$
	At home	284	13.51	9.02	
"At Home" FOR endorsement	Not at home	135	12.18	9.12	$t(417) = -0.47, ns$
	At home	284	12.64	9.38	
"Other" FOR endorsement	Not at home	135	4.23	7.50	$t(417) = 1.87, ns$
	At home	284	2.92	6.27	
"At School" FOR endorsement	Not at home	135	14.71	9.33	$t(417) = -0.44, ns$
	At home	284	15.14	9.35	
"At Work" FOR endorsement	Not at home	135	11.86	8.88	$t(417) = -0.85, ns$
	At home	284	12.69	9.44	

table continues

<u>In "Other" situation while responding</u>					
"With Family" FOR endorsement	Not other situation	408	10.76	8.29	$t(417) = -0.24, ns$
	In other situation	11	11.36	5.90	
"With Friends" FOR endorsement	Not other situation	408	13.47	8.98	$t(417) = 0.07, ns$
	In other situation	11	13.27	5.27	
"At Home" FOR endorsement	Not other situation	408	12.40	9.33	$t(417) = -1.14, ns$
	In other situation	11	15.64	7.16	
"Other" FOR endorsement	Not other situation	408	3.30	6.73	$t(417) = -0.79, ns$
	In other situation	11	4.91	6.14	
"At School" FOR endorsement	Not other situation	408	14.95	9.36	$t(417) = -0.69, ns$
	In other situation	11	16.91	8.63	
"At Work" FOR endorsement	Not other situation	408	12.50	9.27	$t(417) = 1.11, ns$
	In other situation	11	9.36	8.57	
<u>"At School" while responding</u>					
"With Family" FOR endorsement	Not at school	308	10.65	8.47	$t(417) = -0.54, ns$
	At school	111	11.14	7.57	
"With Friends" FOR endorsement	Not at school	308	13.29	9.04	$t(417) = -0.68, ns$
	At school	111	13.96	8.52	
"At Home" FOR endorsement	Not at school	308	12.67	9.54	$t(417) = 0.69, ns$
	At school	111	11.99	8.58	
"Other" FOR endorsement	Not at school	308	2.76	6.10	$t(417) = -2.98, ns$
	At school	111	4.95	7.98	
"At School" FOR endorsement	Not at school	308	15.05	9.59	$t(417) = 0.16, ns$
	At school	111	14.88	8.65	
"At Work" FOR endorsement	Not at school	308	12.47	9.48	$t(417) = 0.20, ns$
	At school	111	12.27	8.64	
<u>"At Work" while responding</u>					
"With Family" FOR endorsement	Not at work	398	10.79	8.19	$t(417) = 0.15, ns$
	At work	21	10.52	9.29	
"With Friends" FOR endorsement	Not at work	398	13.50	8.91	$t(417) = 0.30, ns$
	At work	21	12.90	8.83	
"At Home" FOR endorsement	Not at work	398	12.46	9.14	$t(417) = -0.28, ns$
	At work	21	13.05	12.01	
"Other" FOR endorsement	Not at work	398	3.35	6.76	$t(417) = 0.17, ns$
	At work	21	3.10	5.87	
"At School" FOR endorsement	Not at work	398	15.01	9.22	$t(417) = 0.03, ns$
	At work	21	14.95	11.61	
"At Work" FOR endorsement	Not at work	398	12.37	9.23	$t(417) = -0.51, ns$
	At work	21	13.43	10.06	

Note. *ns* = nonsignificant, $p > .05$

Figure 1. McCrae and Costa's (1996) Meta-Theoretical Framework of Personality

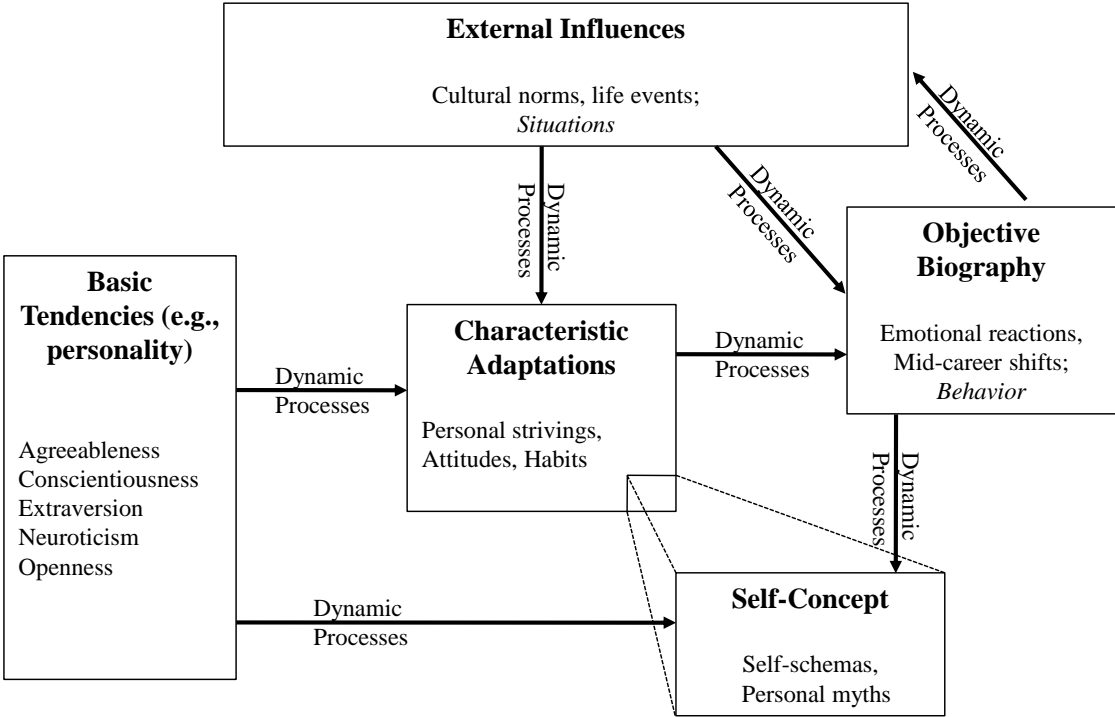
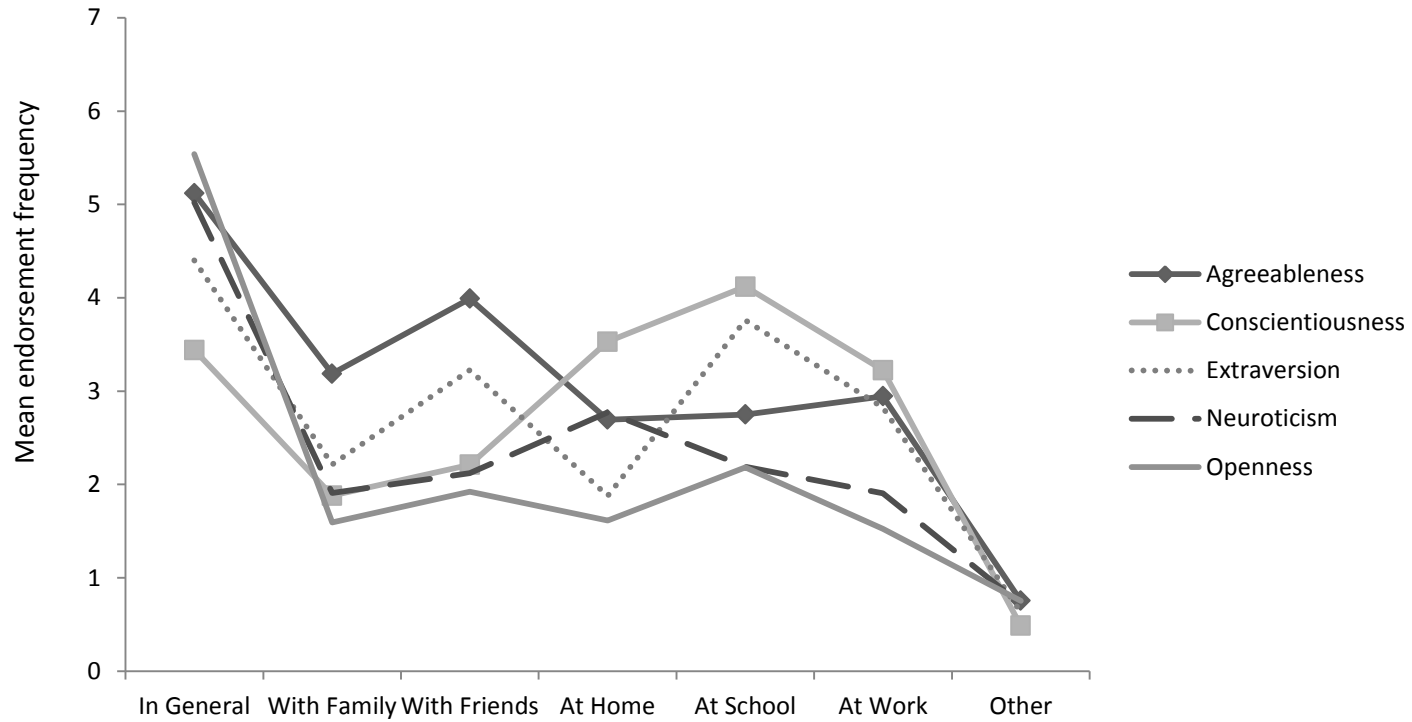
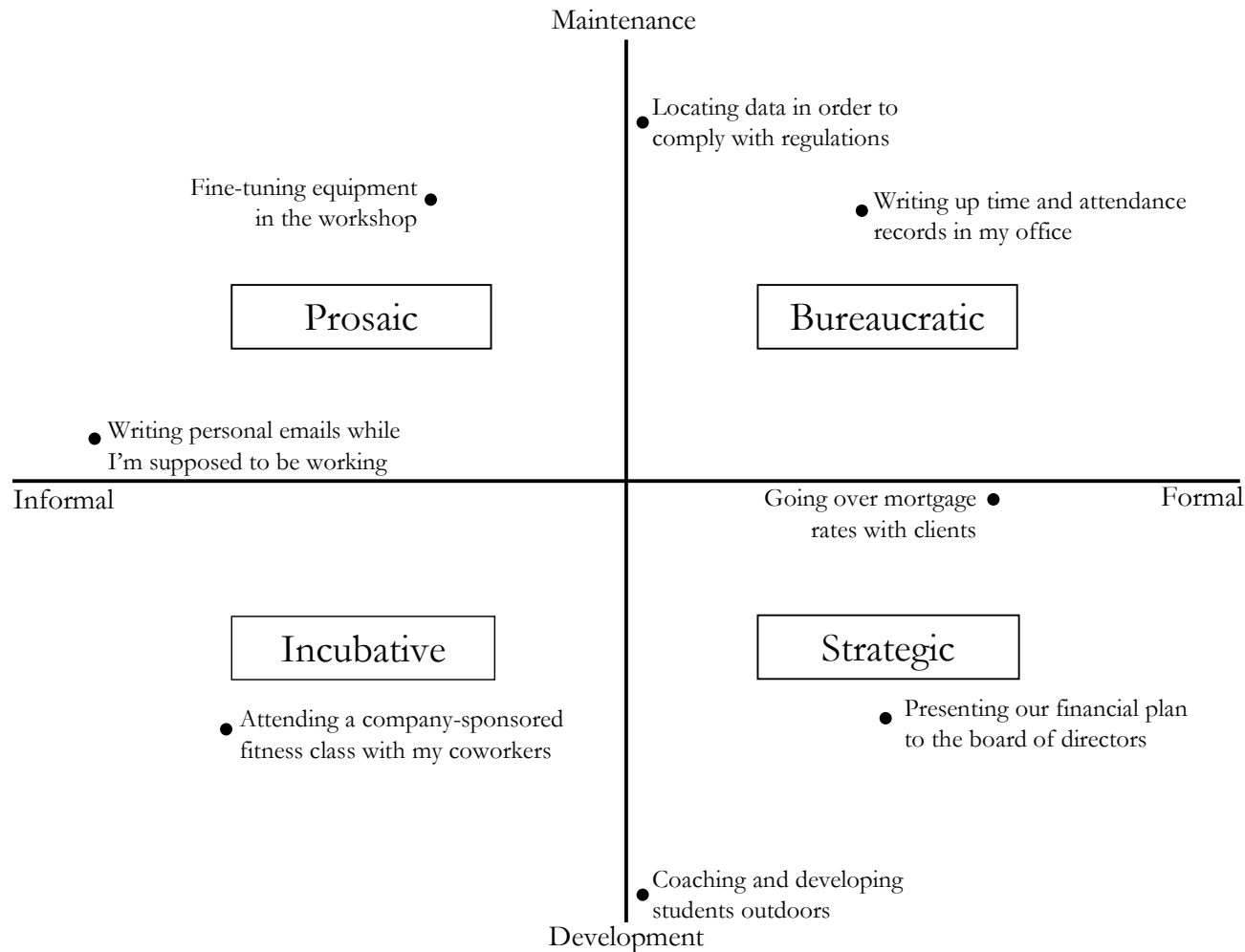


Figure 2. FOR Endorsement by Factor Interaction



Note. Y-axis ranges from 0 - 10.

Figure 3. Meyer (2010) Taxonomy of Work Situations



References

- Ahadi, S., & Diener, E. (1989). Multiple determinants and effects size. *Journal of Personality and Social Psychology, 56*, 398-406.
- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs, 47* (No. 211).
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., Devries, R. E., DiBlas, L., et al. (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*, 9-29.
- Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review, 81*, 506-520.
- Bing, M. N., Whanger, J. C., Davison, H. K., & VanHook, J. B. (2004). Incremental validity of the frame-of-reference effect in personality scale scores: A replication and extension. *Journal of Applied Psychology, 89*, 150-157.

- Briggs, S. R. (1989). The optimal level of measurement for personality constructs. In D. M. Buss, & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 246-260). New York: Springer-Verlag.
- Bronfenbrenner, U. (1979). Contexts of child-rearing. *American Psychologist*, *34*, 844-858.
- Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon (Ed.), *Handbook of child psychology* (Vol. 2, pp. 993-1028). Hoboken, NJ: Wiley.
- Cascio, W. F., & Aguinis, H. (2006). *Applied psychology in human resource management*. Upper Saddle River, NJ: Prentice Hall.
- Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, *38*, 476-506.
- Costa, P. T., & McCrae, R. R. (1980). Still stable after all these years: Personality as a key to some issues in adulthood and old age. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life span development and behavior* (Vol. 3, pp. 65-102). New York: Academic Press.
- Costa, P., & McCrae, R. (1992). *NEO-PI-R and NEO-FFI professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, *64*, 21-50.

- Costa, P. T., McCrae, R. R., & Zonderman, A. B. (1987). Environmental and dispositional influences on well-being: Longitudinal followup of an American national sample. *British Journal of Psychology*, *78*, 299-306.
- DeGroot, T., & Kluemper, D. (2007). Evidence of predictive and incremental validity of personality factors, vocal attractiveness and the situational interview. *International Journal of Selection and Assessment*, *15*, 30-39.
- 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*, 40-57.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.
- Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *Journal of Abnormal and Social Psychology*, *44*, 329-344.
- Galton, F. (1884). Measurement of character. *Fornightly Review*, *36*, 179-185.
- Gatewood, R. D., Feild, H. S., & Barrick (2007). *Human resource selection* (6th ed.). Orlando, FL: Dryden Press.
- Goldberg, L. R. (1997). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7-28). Tillburg, The Netherlands: Tillburg University Press.

- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96.
- Harrison, D. A., & McLaughlin, M. E. (1993). Cognitive processes in self-report responses: Tests of item context effects in work attitude measures. *Journal of Applied Psychology, 78*, 129-140.
- Heggestad, E. D., & Gordon, H. L. (2008). An argument for context-specific personality assessments. *Industrial and Organizational Psychology, 1*, 320-322.
- Heller, D., Ferris, D. L., Brown, D., & Watson, D. (2009). The influence of work personality on job satisfaction: Incremental validity and mediation effects. *Journal of Personality, 77*, 1051-1084.
- Hergenhahn, B. R. (2001). *An introduction to the history of psychology* (4th ed.). Belmont, CA: Wadsworth.
- Higgins, E. T., King, G. A., & Marvin, G. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology, 43*, 35-47.
- Hogan, R. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 873-919). Palo Alto, CA: Consulting Psychologists Press.

- Hogan, R., DeSoto, C. B., & Solano, C. (1975). Traits, tests, and personality research. *American Psychologist, 6*, 255-264.
- Holtz, B. C., Ployhart, R. E., & Dominguez, A. (2005). Testing the rules of justice: The effects of frame-of-reference and pre-test validity information on personality test responses and test perceptions. *International Journal of Selection and Assessment, 13*, 75-86.
- Hough, L. M. (1992). The “Big Five” personality variables-construct confusions: Description versus prediction. *Human Performance, 5*, 139-155.
- Hough, L. M., & Furnham, A. (2003). Use of personality variables in work settings. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology, Volume 12, Industrial and organizational psychology* (pp. 131-169).
- Hough, L. M., & Ones, D. S. (2001). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial work and organizational psychology*, (Vol. 1, pp. 233-277). London: Sage.
- Hough, L. M., Ones, D. S., & Viswesvaran, C. (1998). Personality correlates of managerial performance constructs. Paper presented at the 13th annual conference of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Howell, D. C. (2002). *Statistical methods for psychology*. Pacific Grove, CA: Thomson.

- Hunthausen, J. M., Truxillo, D. M., Bauer, T. N., & Hammer, L. B. (2003). A field study of frame-of-reference effects on personality test validity. *Journal of Applied Psychology, 88*, 545-551.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.
- Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*, 386-408.
- Kahn, R. L., Wolfe, D. M., Quinn, R., Snoek, J. D., & Rosenthal, R. A. (1964). *Organizational Stress*. New York: Wiley.
- Katz, D. & Kahn, R. (1978). *The social psychology of organizations* (2nd Ed.). New York: John Wiley & Sons.
- Kline, R. (2005). *Principles and practice of structural equation modeling*. New York: Guilford.
- Knowles, E. S. (1988). Item context effects on personality scales: Measuring changes the measure. *Journal of Personality and Social Psychology, 55*, 312-320.
- Kravitz, D. A., & Balzer, W. K. (1992). Context effects in performance appraisal: A methodological critique and empirical study. *Journal of Applied Psychology, 77*, 24-31.
- Kuunsela, H., & Paul, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *The American Journal of Psychology, 113*, 387-404.

- 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*, 268-279.
- McCrae, R. R., & Costa, P. T. (1996). Toward a new generation of personality theories: Theoretical contexts for the Five-Factor Model. In J. S. Wiggins (Ed.), *The Five-Factor Model of personality: Theoretical perspectives* (pp. 51-87). New York: Guilford Press.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Meyer, R. (2010). A taxonomy of work situations to help focus frame-of-reference personality tests. Paper presented at the annual meeting of the Society for industrial and Organizational Psychology, Atlanta, GA.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246-268.
- 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.
- Morse, J. M. (1991). Strategies for sampling. In J. M. Morse (Ed.) *Qualitative nursing research: A contemporary dialogue* (pp. 127-145). Newbury Park, CA: Sage.

- Morse, J. M. (1995). The significance of saturation. *Qualitative Health Research, 5*, 147-149.
- Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*, 145-165.
- Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology, 79*, 272-280.
- Murphy, K. R., & Lee, S. L. (1994). Personality variables related to integrity test scores: The role of conscientiousness. *Journal of Business and Psychology, 8*, 413-424.
- Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organization settings. *Personnel Psychology, 60*, 995-1027.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: A red herring. *Journal of Applied Psychology, 81*, 660-679.
- Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical employees. *Journal of Applied Psychology, 65*, 373-406.
- Pervin, L. A., & John, O. P. (2000). *Personality: Theory and research*. New York: Wiley.

- Ployhart, R. E., & Bliese, P. D. (2006). Individual ADAPTability (I-ADAPT) theory: Conceptualizing the antecedents, consequences, and measurement of individual differences in adaptability. In C. S. Burke, L. G. Pierce, & E. Salas (Eds.), *Understanding adaptability: A prerequisite for effective performance within complex environments* (pp. 3-40). Amsterdam, London: Elsevier.
- Pritchard, R. (1990). The evolution of introspective methodology and its implications for studying the reading process. *Reading Psychology: An International Quarterly*, *11*, 1 - 13.
- Pulakos, E. D., & Schmitt, N. (1995). Experience-based and situational interview questions: Studies of validity. *Personnel Psychology*, *48*, 289-308.
- 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*, 101-117.
- Robie, C., Brown, D. J., & Beaty, J. C. (2007). Do people fake on personality inventories? A verbal protocol analysis. *Journal of Business and Psychology*, *21*, 489-509.
- 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*, 348-365.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.

- Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology, 80*, 607-620.
- Shavelson, R. J., & Webb, M. N. (1991). *Generalizability theory: A primer*. Thousand Oaks, CA: Sage.
- Small, E. E., & Diefendorff, J. M. (2006). The impact of contextual self-ratings and observer ratings of personality on the personality-performance relationship. *Journal of Applied Social Psychology, 36*, 297-320.
- 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*, 304-315.
- Snyder, M., & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. *Journal of Personality and Social Psychology, 51*, 125-139.
- Steenkamp, J. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*, 78-90.
- Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology, 60*, 967-993.
- Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality, 34*, 397-423.

- Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality, 60*, 225-251. (Original work published in 1961)
- Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.
- Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.
- Vandenburg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-70.
- Wainer, H. & Kiely, G. L. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Journal of Education Measurement, 24*, 185-201.
- Wang, M., Zhan, J., McCune, E., & Truxillo, D. M. (in press). Understanding newcomers' adaptability and work-related outcomes: Testing the mediating role of perceived P-O fit. *Personnel Psychology*.
- Weiss, H. M., & Rupp, D. E. (2010). Experiencing work: An essay on person-centric work psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4*, 1-25.
- Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*, 181-192.
- Wright, J. C., & Mischel, W. (1987). A conditional approach to dispositional constructs: The local predictability of social behavior. *Journal of Personality and Social Psychology, 53*, 1159-1177.

Appendix A: Study 1 Screening Survey

1. Age: _____ years
2. Gender
 - a. Male
 - b. Female
3. Race
 - a. White/Caucasian
 - b. Black/African American
 - c. American Indian or Alaska Native
 - d. Asian
 - e. Hawaiian or Pacific Islander
 - f. Hispanic or Latino
 - g. Other: _____
4. Highest level of education
 - a. Grade school or less
 - b. Some high school
 - c. High school graduate
 - d. Some college
 - e. College graduate
 - f. Some graduate study
 - g. Graduate degree
5. Were you born in the United States?
6. What is your current status in school:
 - a. Freshman
 - b. Sophomore
 - c. Junior
 - d. Senior
 - e. Post-baccalaureate/graduate student
 - f. Other: _____
7. Are you currently employed?
 - a. Yes
 - b. No
8. If yes, how many paying jobs do you currently have?
9. How would you classify your current position? If you have more than one job, please check all that apply.
 - a. I am not currently employed.
 - b. Officials & Managerial
 - c. Professional
 - d. Technician
 - e. Sales
 - f. Office/Clerical
 - g. Craft Worker
 - h. Machine Operator

- i. Laborer
 - j. Service Worker
 - k. Other (please specify)
10. What is your current or most recent job title? If you have more than one job, please list them all.
 11. About how many hours per week do you work across all of your paying jobs?
 12. Please indicate how you are paid. If the method varies across the jobs you hold, please check all that apply.
 - a. Salaried
 - b. Paid by the hour
 - c. Paid by commission/tips
 - d. Paid by the hour and commission/tips
 - e. Other: _____
 13. How long have you worked for your current organization? If you have more than one job please provide this information for the job you have held the longest.
 14. How long have you been working in your field? If you work in more than one field, please provide this information for the field that you have worked in the longest.
 15. Please estimate the years and months of total work experience that you have, across all jobs that you have held.
 16. Please estimate the number of jobs you have applied for in your lifetime.
 17. Please describe your marital status:
 - a. Married
 - b. Not married, but living with a partner
 - c. Single
 18. Have you been married before?
 - a. Yes
 - b. No
 19. Please list the number of individuals under the age of 18 that live with you at least 3 days per week, and for whom you are primarily responsible.
 20. Please list the number of individuals under the age of 18 who do not live with you at least 3 days per week, but for whom you are at least partially responsible.
 21. How many living parents, including step-parents, do you have?
 22. If you are married or living with a partner: how many living parents, including step-parents, does your spouse or partner have?
 23. How many hours per week do you spend caring for your parents/step-parents or your spouse or partner's parents/step-parents? This could involve shopping, home maintenance, transportation, emotional support, financial management, making meals, bathing, etc.

Appendix B: Interviewer “Cheat Sheet”

Have a good word for everyone.	
Believe that others have good intentions.	
Respect others.	
Accept people as they are.	
Make people feel at ease.	
Have a sharp tongue.	‘Sharp tongue’: to say mean things
Cut others to pieces.	To be cruel
Suspect hidden motives in others.	
Get back at others.	‘Get back’: to take revenge
Insult people.	
Am always prepared.	
Pay attention to details.	
Get chores done right away.	
Carry out my plans.	
Make plans and stick to them.	
Waste my time.	
Find it difficult to get down to work.	
Do just enough work to get by.	
Don’t see things through.	
Shirk my duties.	‘Shirk’: to avoid, try and get out of
Make friends easily.	
Am skilled in handling social situations.	
Am the life of the party.	
Know how to captivate people.	
Have little to say.	
Keep in the background.	
Would describe my experiences as somewhat dull.	
Don’t like to draw attention to myself.	
Don’t talk a lot.	
Often feel blue.	
Dislike myself.	
Am often down in the dumps.	
Have frequent mood swings.	
Panic easily.	
Rarely get irritated.	
Seldom feel blue.	
Feel comfortable with myself.	
Am not easily bothered by things.	
Am very pleased with myself.	
Believe in the importance of art.	
Have a vivid imagination.	
Tend to vote for liberal political candidates.	Democrats are generally thought of as liberals.
Carry the conversation to a higher level.	
Enjoy hearing new ideas.	
Am not interested in abstract ideas.	‘Abstract’: theoretical, not applied or practical
Do not like art.	
Avoid philosophical discussions.	
Do not enjoy going to art museums.	
Tend to vote for conservative political candidates.	Republicans are generally thought of as conservatives.

Appendix C: Researcher Script and Follow-Up Questions for Study 1

Thank you for your willingness to participate in this study. I would first like you to review the informed consent. The informed consent will explain the purpose of this study and any potential benefits or problems you may experience as a consequence of participating. *[Review informed consent, if participant agrees have them sign.]* Here is a copy of the informed consent for you to keep. *[Give copy of IC to participant.]*

Here is a summary of the demographic information you provided in the screening survey. *[Provide them the sheet with their demographic information].* Please confirm that this information is correct, or correct any errors. *[Participant confirms demographic information, corrects any errors.]* Thank you.

We are about to begin the study. To protect your confidentiality, I am going to use a code instead of your name to identify you in the audio recording. When I turn the tape recorder on, I will state the code. Please refrain from saying your name once the recording has begun. Are you ready? *[Participant says yes, turn on recorder].*

[Into the recorder, say the following]. “This interview is for participant number ___”.

[Give participant the personality measure]. This a personality inventory that is frequently used in psychological research and practice. Please read the instructions and let me know if you have any questions. *[Answer participant questions, if any].*

I would like you to provide responses to each question on the inventory. As soon as you respond to EACH item, please think aloud and describe your reasoning process. Tell me *why* you are choosing a particular response. If you are silent for more than a few seconds, I will remind you to think aloud by saying: “Can you tell me what you

are thinking” or “Can you tell me why you chose that answer?” For the rest of the time I will just smile and nod.

Please read the statement out loud. Also, please say out loud which response option you have chosen.

Are you ready? Please begin.

[Participant completes inventory.]

Thank you. Now I just have a few follow-up questions for you.

- 1) As you responded to the statements, were you thinking of yourself in any particular places, environments or situations? If yes, please describe those environments. If no, proceed to #2.

- 2) Take a minute to briefly review the statements. Please describe any places, environments, or situations that might come to mind in re-reading the statements and determining how descriptive they are of you.

- 3) Would you describe yourself as a person that generally tends to behave in the same way regardless of the situation, or do you tend to adapt your behavior to fit the situation?

Appendix D: IPIP Personality Inventory

Agreeableness

Have a good word for everyone.
Believe that others have good intentions.
Respect others.
Accept people as they are.
Make people feel at ease.
Have a sharp tongue.
Cut others to pieces.
Suspect hidden motives in others.
Get back at others.
Insult people.

Conscientiousness

Am always prepared.
Pay attention to details.
Get chores done right away.
Carry out my plans.
Make plans and stick to them.
Waste my time.
Find it difficult to get down to work.
Do just enough work to get by.
Don't see things through.
Shirk my duties.

Extraversion

Feel comfortable around people.
Make friends easily.
Am skilled in handling social situations.
Am the life of the party.
Know how to captivate people.
Have little to say.
Keep in the background.
Would describe my experiences as somewhat dull.
Don't like to draw attention to myself.
Don't talk a lot.

Neuroticism

Often feel blue.
Dislike myself.
Am often down in the dumps.
Have frequent mood swings.
Panic easily.
Rarely get irritated.
Seldom feel blue.
Feel comfortable with myself.
Am not easily bothered by things.
Am very pleased with myself.

Openness to Experience

Believe in the importance of art.
Have a vivid imagination.
Tend to vote for liberal political candidates.
Carry the conversation to a higher level.
Enjoy hearing new ideas.
Am not interested in abstract ideas.
Do not like art.
Avoid philosophical discussions.
Do not enjoy going to art museums.
Tend to vote for conservative political candidates.

Appendix E: Study 2 Survey

Dear Potential Research Participant,

You are invited to take part in a research study conducted by Elizabeth McCune, a PhD student in Portland State University's Psychology Department. This study is part of Elizabeth's dissertation. The purpose of the study is to understand the thought process that individuals go through when completing a personality inventory. You were selected as a possible participant in this study because you are enrolled in an undergraduate psychology course.

If you decide to participate, you will be asked to respond to a questionnaire that will take 30-60 minutes to complete. You will be offered extra credit for participating in this study. In addition, the results of this study will contribute to a greater understanding of personality tests and their uses.

Any information that is obtained in connection with this study and that can be linked to you or identify you will be kept confidential. It will be stored in a locked and secure location.

Your participation is voluntary and confidential. You do not have to participate in this study, and you may withdraw from the study at any time. If you feel uncomfortable answering any question you may skip it. Likewise, your individual questionnaire answers will be kept confidential. The results of this research will only be reported in aggregate form (everyone's information will be pooled together and summarized). Your participation will not affect your employment standing and will not be provided to your employer. All information you provide will only be used for research purposes.

The proposed study is relatively risk-free and non-threatening. However, there is a chance you may experience some discomfort. However, we would like to remind you that your individual responses are not linked to your name and only the researchers will have access to your responses. In addition, you can stop at any time and skip any question you do not feel comfortable answering.

If you have any concerns or questions about this study, please contact Elizabeth McCune at mccunee@pdx.edu or (503) 320-6023. If you have questions about your rights as a research participant please contact the Chair of the Human Subjects Research Review Committee, Office of Research and Sponsored Projects, 600 Unitus Building, 2121 SW 4th Avenue, Portland, Oregon, 97201, (503) 725-4288.

Sincerely,
Elizabeth McCune

IMPORTANT: If you completed this survey in the Spring term, or if you would prefer to complete the alternative assignment in order to receive your extra credit, please go here: <https://survey.oit.pdx.edu/ss/wsb.dll/s/2bfg790>

1) Please check the box below if you would like to participate.

I agree to participate in this study.

2)

Below are phrases describing people's behaviors. Please use the options provided to describe how accurately each statement 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, in relation to other people you know of the same sex as you are, and roughly your same age.

Remember that your responses will be kept in absolute confidence. Please read each statement carefully, and then select the bubble that corresponds to the response you have chosen.

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Have a good word for everyone.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Believe that others have good intentions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Respect others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Accept people as they are.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Make people feel at ease.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3)

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Have a sharp tongue.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cut others to pieces.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Suspect hidden motives in others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Get back at others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insult people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4)

Below are phrases describing people's behaviors. Please use the options provided to describe how accurately each statement 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, in relation to other people you know of the same sex as you are, and roughly your same age.

Remember that your responses will be kept in absolute confidence. Please read each statement carefully, and then select the bubble that corresponds to the response you have chosen.

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Am always prepared.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pay attention to details.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Get chores done right away.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Carry out my plans.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Make plans and stick to them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5)

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Waste my time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Find it difficult to get down to work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do just enough work to get by.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Don't see things through.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Try to avoid my duties.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6)

Below are phrases describing people's behaviors. Please use the options provided to describe how accurately each statement 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, in relation to other people you know of the same sex as you are, and roughly your same age.

Remember that your responses will be kept in absolute confidence. Please read each statement carefully, and then select the bubble that corresponds to the response you have chosen.

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Feel comfortable around people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Make friends easily.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Am skilled in handling social situations.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Am the life of the party.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Know how to captivate people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
-------------------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

7)

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Have little to say.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Keep in the background.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Would describe my experiences as somewhat dull.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Don't like to draw attention to myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Don't talk a lot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8)

Below are phrases describing people's behaviors. Please use the options provided to describe how accurately each statement 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, in relation to other people you know of the same sex as you are, and roughly your same age.

Remember that your responses will be kept in absolute confidence. Please read each statement carefully, and then select the bubble that corresponds to the response you have chosen.

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Often feel blue.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dislike myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Am often down in the dumps.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Have frequent mood swings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Panic easily.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9)

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Rarely get irritated.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Seldom feel blue.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Feel comfortable with myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Am not easily bothered by things.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Am very pleased with myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10)

Below are phrases describing people's behaviors. Please use the options provided to describe how accurately each statement 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, in relation to other people you know of the same sex as you are, and roughly your same age.

Remember that your responses will be kept in absolute confidence. Please read each statement carefully, and then select the bubble that corresponds to the response you have chosen.

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Believe in the importance of art.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Have a vivid imagination.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tend to vote for liberal political candidates.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Carry the conversation to a higher level.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Enjoy hearing new ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11)

	Very Inaccurate	Inaccurate	Neither Accurate nor Inaccurate	Accurate	Very Accurate
Am not interested in abstract ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not like art.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Avoid philosophical discussions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do not enjoy going to art museums.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tend to vote for conservative political candidates.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

12)

Instructions: In the next section of the survey, you will be asked to reflect back on the responses you provided during the first section. Specifically, you will be asked to indicate which situations you were thinking of when responding to the questions.

You will be provided with seven options. **PLEASE CHECK ALL THAT APPLY.** Here are your response options and their definitions:

At Work: Choose this response if you were thinking of how you are when you are at work or when you interact with your boss and/or coworkers.

At Home: Choose this response if you were thinking of how you are when you are at home, either alone or with whomever you live with.

At School: Choose this response if you were thinking of how you are when you are on campus, in class, or when you interact with instructors and classmates.

With Friends: Choose this response if you were thinking of how you are when you are with your friends.

With Family: Choose this response if you were thinking of how you are when you are with your family. This includes your immediate family (parents, siblings), extended family (aunts, uncles, grandparents, cousins) and your significant other.

In General: Choose this response if you were thinking of how you are in general.

Other: Choose this response if you were thinking of how you are when in some other situation that is not included in the other response options. *Please use the "Additional comments" box to describe these situation(s).*

Click Next Page to Proceed.

13)

Please indicate which situations you were thinking of when responding to each question in the first section of the survey.

PLEASE CHECK ALL THAT APPLY

At Work: Choose this response if you were thinking of how you are when you are at work or when you interact with your boss and/or coworkers.

At Home: Choose this response if you were thinking of how you are when you are at home, either alone or with whomever you live with.

At School: Choose this response if you were thinking of how you are when you are on campus, in class, or when you interact with instructors and classmates.

With Friends: Choose this response if you were thinking of how you are when you are with your friends.

With Family: Choose this response if you were thinking of how you are when you are with your family. This includes your immediate family (parents, siblings), extended family (aunts, uncles, grandparents, cousins) and your significant other.

In General: Choose this response if you were thinking of how you are in general.
Other: Choose this response if you were thinking of how you are when in some other situation that is not included in the other response options. *Please use the additional comments box to describe the situation(s) you were thinking of.*

	At Work	At Home	At School	With Friends	With Family	In General	Other
Have a good word for everyone.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Believe that others have good intentions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Respect others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Accept people as they are.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Make people feel at ease.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

14)
PLEASE CHECK ALL THAT APPLY

	At Work	At Home	At School	With Friends	With Family	In General	Other
Have a sharp tongue.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cut others to pieces.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Suspect hidden motives in others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Get back at others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Insult people.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

15)
Please indicate which situations you were thinking of when responding to each question in the first section of the survey.
PLEASE CHECK ALL THAT APPLY
At Work: Choose this response if you were thinking of how you are when you are at work or when you interact with your boss and/or coworkers.
At Home: Choose this response if you were thinking of how you are when you are at home, either alone or with whomever you live with.
At School: Choose this response if you were thinking of how you are when you are on campus, in class, or when you interact with instructors and classmates.
With Friends: Choose this response if you were thinking of how you are when you are with your friends.
With Family: Choose this response if you were thinking of how you are when you are with your family. This includes your immediate family (parents, siblings),

extended family (aunts, uncles, grandparents, cousins) and your significant other. In General: Choose this response if you were thinking of how you are in general. Other: Choose this response if you were thinking of how you are when in some other situation that is not included in the other response options. *Please use the additional comments box to describe the situation(s) you were thinking of.*

	At Work	At Home	At School	With Friends	With Family	In General	Other
Am always prepared.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pay attention to details.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Get chores done right away.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Carry out my plans.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Make plans and stick to them.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**16)
PLEASE CHECK ALL THAT APPLY**

	At Work	At Home	At School	With Friends	With Family	In General	Other
Waste my time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Find it difficult to get down to work.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do just enough work to get by.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Don't see things through.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Try to avoid my duties.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**17)
Please indicate which situations you were thinking of when responding to each question in the first section of the survey.
PLEASE CHECK ALL THAT APPLY**

At Work: Choose this response if you were thinking of how you are when you are at work or when you interact with your boss and/or coworkers.

At Home: Choose this response if you were thinking of how you are when you are at home, either alone or with whomever you live with.

At School: Choose this response if you were thinking of how you are when you are on campus, in class, or when you interact with instructors and classmates.

With Friends: Choose this response if you were thinking of how you are when you are with your friends.

With Family: Choose this response if you were thinking of how you are when you are with your family. This includes your immediate family (parents, siblings), extended family (aunts, uncles, grandparents, cousins) and your significant other.

In General: Choose this response if you were thinking of how you are in general.

Other: Choose this response if you were thinking of how you are when in some other situation that is not included in the other response options. *Please use the additional comments box to describe the situation(s) you were thinking of.*

	At Work	At Home	At School	With Friends	With Family	In General	Other
Feel comfortable around people.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Make friends easily.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Am skilled in handling social situations.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Am the life of the party.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Know how to captivate people.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

18)
PLEASE CHECK ALL THAT APPLY

	At Work	At Home	At School	With Friends	With Family	In General	Other
Have little to say.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Keep in the background.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Would describe my experiences as somewhat dull.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Don't like to draw attention to myself.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Don't talk a lot.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

19)
Please indicate which situations you were thinking of when responding to each question in the first section of the survey.
PLEASE CHECK ALL THAT APPLY
At Work: Choose this response if you were thinking of how you are when you are at work or when you interact with your boss and/or coworkers.
At Home: Choose this response if you were thinking of how you are when you are at home, either alone or with whomever you live with.
At School: Choose this response if you were thinking of how you are when you are on campus, in class, or when you interact with instructors and classmates.
With Friends: Choose this response if you were thinking of how you are when you are with your friends.
With Family: Choose this response if you were thinking of how you are when you are with your family. This includes your immediate family (parents, siblings),

extended family (aunts, uncles, grandparents, cousins) and your significant other.
In General: Choose this response if you were thinking of how you are in general.
Other: Choose this response if you were thinking of how you are when in some other situation that is not included in the other response options. *Please use the additional comments box to describe the situation(s) you were thinking of.*

	At Work	At Home	At School	With Friends	With Family	In General	Other
Often feel blue.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dislike myself.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Am often down in the dumps.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have frequent mood swings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Panic easily.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

20)
PLEASE CHECK ALL THAT APPLY

	At Work	At Home	At School	With Friends	With Family	In General	Other
Rarely get irritated.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Seldom feel blue.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Feel comfortable with myself.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Am not easily bothered by things.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Am very pleased with myself.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

21)
Please indicate which situations you were thinking of when responding to each question in the first section of the survey.
PLEASE CHECK ALL THAT APPLY
At Work: Choose this response if you were thinking of how you are when you are at work or when you interact with your boss and/or coworkers.
At Home: Choose this response if you were thinking of how you are when you are at home, either alone or with whomever you live with.
At School: Choose this response if you were thinking of how you are when you are on campus, in class, or when you interact with instructors and classmates.
With Friends: Choose this response if you were thinking of how you are when you are with your friends.
With Family: Choose this response if you were thinking of how you are when you are with your family. This includes your immediate family (parents, siblings), extended family (aunts, uncles, grandparents, cousins) and your significant other.

In General: Choose this response if you were thinking of how you are in general.
Other: Choose this response if you were thinking of how you are when in some other situation that is not included in the other response options. *Please use the additional comments box to describe the situation(s) you were thinking of.*

	At Work	At Home	At School	With Friends	With Family	In General	Other
Believe in the importance of art.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Have a vivid imagination.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tend to vote for liberal political candidates.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Carry the conversation to a higher level.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enjoy hearing new ideas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

22)
PLEASE CHECK ALL THAT APPLY

	At Work	At Home	At School	With Friends	With Family	In General	Other
Am not interested in abstract ideas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do not like art.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Avoid philosophical discussions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do not enjoy going to art museums.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tend to vote for conservative political candidates.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

23) Please indicate which of the following options describe the situation you are CURRENTLY in while completing this survey.

- I am at work.
- I am at home.
- I am at school.
- I am with my friends.
- I am with my family.
- Other (please specify)

If you selected other, please specify

24) You are more than halfway done! Thank you for your diligence up to this point. If you're feeling tired, now would be a great time to take a short break and maybe stretch your legs or get a drink of water. The more focused you are when answering the questions, the better the quality of the data you provide will be.

Thanks again for taking the time to participate in this important study! Click Next Page to keep going!

25) Please consider the extent to which each statement describes you, and then select the appropriate response.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I can only argue for ideas in which I already believe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would probably make a good actor.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find it hard to imitate the behavior of other people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In a group of people I am rarely the center of attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I may deceive people by being friendly when I really dislike them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

26) Please consider the extent to which each statement describes you, and then select the appropriate response.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
In different situations and with different people, I often act like very different persons.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can make impromptu speeches even on topics about which I have almost no information.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
At parties and social gatherings, I do not attempt to do or say things that others will like.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have trouble changing my behavior to suit different people and different situations.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
At a party I let others keep the jokes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

and stories going.					
--------------------	--	--	--	--	--

27) Please consider the extent to which each statement describes you, and then select the appropriate response.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I have considered being an entertainer.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I guess I put on a show to impress or entertain others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel a bit awkward in public and do not show up quite as well as I should.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can look anyone in the eye and tell a lie with a straight face (if for a right end).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

28) Please consider the extent to which each statement describes you, and then select the appropriate response.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I am not particularly good at making other people like me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I'm not always the person I appear to be.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have never been good at games like charades or improvisational acting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would not change my opinions (or the way I do things) in order to please someone or win their favor.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

29) The following questions ask you to describe your behavior at work.
***If you have more than one job, please choose the one for which you work the most hours.**
****If you are not currently employed, please skip to the next page.**

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I engage in activities that positively affect my performance evaluation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I meet formal performance requirements of the job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I adequately complete assigned duties.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I fulfill responsibilities specified in the job description.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I perform tasks that are expected of me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I fail to perform essential job duties.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

30) Please consider both your work and home life when answering these questions.

***If you have more than one job, consider ALL OF YOUR JOBS when responding to these questions.**

****If you are not currently employed, please go to the next page.**

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
The demands of my work interfere with my home and family life.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The amount of time my job takes up makes it difficult to fulfill my family responsibilities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Things I want to do at home do not get done because of the demands my job puts on me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My job produces strain that makes it difficult to fulfill family duties.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

31) What is your cumulative college GPA as of Winter 2010? Please round to two decimal places (e.g., 3.67)

32) The following questions ask you to describe your behavior at school. Please answer honestly. Remember that responses are only reported in groups, and your individual responses cannot be identified.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I engage in activities that positively affect my school performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I meet school performance requirements.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I adequately complete assigned school assignments.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I fulfill specified school responsibilities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I perform school assignments that	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

are expected of me.					
I fail to perform essential school activities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

33) Please indicate how frequently each of these things has happened to you during this academic year (Fall 2009-Spring 2010).

	Never	About once each quarter	About once each month	About once each week	Several times each week
Skipped several classes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Had trouble completing assignments	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fallen far behind in reading	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Come to class unprepared	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Taken a test I was not fully prepared for	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Turned in an assignment late	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sought tutoring for help with a course	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Had lower grades than expected	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

34) The next section of the survey asks basic questions about you and your life.

35) Age

36) Gender

- Male
- Female

37) Ethnicity (please select all that apply)

- White/Caucasian
- Black/African American
- American Indian or Alaska Native
- Asian
- Hawaiian or Pacific Islander

- Hispanic or Latino
- Other (please specify)

If you selected other, please specify

38) Highest level of education

- Grade school or less
- Some high school
- High school graduate
- Some college
- College graduate
- Some graduate study
- Graduate degree

39) Were you born in the United States?

- Yes
- No

40) What is your current status in school?

- Freshman
- Sophomore
- Junior
- Senior
- Post-baccalaureate/graduate student
- Other (please specify)

If you selected other, please specify

41) Are you currently employed?

- Yes
- No

42) If you are currently employed, how many paying jobs do you have?

43) How would you classify your position? If you have more than one job, please check all that apply.

- I am not currently employed.
- Officials & Managerial
- Professional
- Technician
- Sales
- Office/Clerical
- Craft Worker
- Machine Operator
- Laborer
- Service Worker
- Other (please specify)

If you selected other, please specify

44) What is your current or most recent job title? If you have more than one job, please list them all.

45) About how many hours per week do you work across all of your paying jobs?

46) Please indicate how you are paid. If the method varies across the jobs you hold, please check all that apply.

- Salaried
- Paid by the hour
- Paid by commission/tips
- Paid by the hour and commission/tips
- Other (please specify)

If you selected other, please specify

47) How long have you worked for your current organization? Please enter "0" if you are not currently employed or have never had a job. If you have more than one job, please refer to your primary occupation when answering the question.

Years
Months

48) How long have you been working in your field? Please enter "0" if you have never had a job. If you have more than one job, please refer to your primary occupation when answering the question.

Years _____
Months _____

49) Please estimate the years and months total of work experience that you have across all jobs that you have held.

Years _____
Months _____

50) Please estimate the number of jobs you have applied for in your lifetime.

51) Please describe your marital status.

- Single
- Not married, but living with partner
- Married
- Other (please specify)

If you selected other, please specify

52) Have you been married before?

- Yes
- No

53) Please indicate the number of individuals under the age of 18 that live with you at least 3 days per week, and for whom you are primarily responsible.

54) Please indicate the number of individuals under the age of 18 who do not live with you at least 3 days a week, but for whom you are at least partially responsible.

55) How many living parents, including step-parents, do you have?

56) If you are married or living with your partner please indicate how many living parents, including step-parents, your spouse or partner has.

57) How many hours per week do you spend caring for your parents/step-parents or your spouse or partner's parents/step-parents? This could involve shopping, home maintenance, transportation, emotional support, financial management, making meals, bathing, etc.
