

DISSERTATION

AN INVESTIGATION OF THE UNDERLYING MECHANISM OF DISCRIMINATION: THE  
PROTOTYPE MATCHING MODEL

Submitted by

Adam Vanhove

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2013

Doctoral Committee:

Advisor: Alyssa M. Gibbons

Kevin R. Murphy  
Benjamin A. Clegg  
Troy V. Mumford

## ABSTRACT

### AN INVESTIGATION OF THE UNDERLYING MECHANISM OF DISCRIMINATION: THE PROTOTYPE MATCHING MODEL

Within this study a model explaining the underlying process that results in employment discrimination was proposed and tested. The prototype matching model represents a moderated mediational model that builds on the existing lack-of-fit model (Heilman, 1983) by considering evaluators' prototypes in determining "fit". Each of the propositions of the model was tested using the Goldberg paradigm design among a sample of Amazon Mechanical Turk (MTurk) participants. Participants described their prototypical jobholders for two different positions (sales and customer service) using two factors: warmth and competence. Participants viewed a set of stimulus applicants to each position and indicated their warmth and competence stereotypes of each applicant, then rated each applicant on three evaluative outcomes (coworker desirability, coworker undesirability, and likelihood of success). The presence of discrimination was assessed with regard to two applicant characteristics: applicant sex (male/female) and weight (normal weight/overweight). Evidence of sex-based discrimination was found, as female applicants received more positive evaluations than males across all three outcomes. No differences in evaluations were found due to applicant weight. Females were stereotyped as both warmer and more competent than males. Normal weight applicants were stereotyped as more competent, while overweight applicants were stereotyped as warmer. Both warmth and competence were strong predictors of each of the three evaluative outcomes. Warmth stereotypes were more strongly associated with coworker desirability evaluations, competence stereotypes were more strongly associated with likelihood of success evaluations, and applicants stereotyped as both warm and competent generally received the most positive evaluations. In addition, stereotypes fully mediated the relationship between applicant sex and all three evaluative outcomes. Finally, evidence indicated that participants' evaluations of applicants were influenced by the (mis)match between evaluators' prototypical jobholder and the stereotypes they attributed to applicants, with stereotype-prototype

competence match more strongly predicting evaluations for the sales position and stereotype-prototype warmth match more strongly predicting evaluations for the customer service position. Findings suggest that the prototype matching model may have utility in explaining evaluators' cognitive processes that lead to discrimination. Future directions are discussed.

## TABLE OF CONTENTS

ABSTRACT.....	ii
LIST OF TABLES.....	v
LIST OF FIGURES.....	vi
CHAPTER 1: INTRODUCTION.....	1
Sex and Weight Discrimination.....	3
Stereotyping.....	7
Prototypes.....	12
Stereotype Content Model (SCM).....	13
The Prototype Matching Model.....	15
Context of the Current Study.....	18
CHAPTER 2: METHOD.....	20
Participants.....	20
Procedure.....	20
Materials.....	22
Measures.....	23
Data Cleaning.....	24
Analyses.....	26
CHAPTER 3: RESULTS.....	28
Preliminary Factor Structure Analyses.....	28
Preliminary Analyses.....	29
Multilevel Modeling Analyses.....	30
Step 1 (X → Y): Applicant Sex and Weight Predicting Outcomes.....	31
Step 2 (X → M): Applicant Sex and Weight Predicting Stereotype Content (Warmth and Competence).....	34
Step 3 (M → Y): Stereotype Content (Warmth and Competence) Predicting Outcomes.....	36
Step 4 (X, M → Y): Applicant Sex and Weight Predicting Outcomes (Controlling for Stereotype Content: Warmth and Competence).....	39
Prototype Matching.....	40
CHAPTER 4: DISCUSSION.....	44
Discrimination as a Function of Applicant Sex and Weight (Step 1).....	44
Stereotype Differences as a Function of Applicant Sex and Weight (Step 2).....	46
Discrimination as a function of stereotype differences (Step 3).....	46
Stereotypes Mediating Discrimination Based on Applicant Sex (Step 4).....	47
Prototype Matching.....	48
Limitations.....	49
Future Research Directions.....	50
Conclusion.....	53
REFERENCES.....	81
APPENDIX II.....	97
APPENDIX III.....	100
APPENDIX IV.....	103
APPENDIX V.....	106
APPENDIX VI.....	109

## LIST OF TABLES

Table 1 .....	63
Table 2 .....	64
Table 3 .....	66
Table 4 .....	68
Table 5 .....	69
Table 6 .....	70
Table 7 .....	72
Table 8 .....	74
Table 9 .....	75
Table 10 .....	76
Table 11a.....	77
Table 11b .....	78

## LIST OF FIGURES

Figure 1 .....	55
Figure 2 .....	56
Figure 3 .....	57
Figure 4 .....	58
Figure 5 .....	59
Figure 6 .....	60
Figure 7 .....	61
Figure 8 .....	62

## CHAPTER 1: INTRODUCTION

Discrimination is characterized by situations in which one group has a systematic advantage over another. It is a problem that has received continued attention in organizational settings, particularly with regard to personnel decisions such as hiring and promotion. On the one hand, organizations must discriminate; that is, they must make distinctions among job applicants and among employees. This, in fact, is the premise behind the use of selection procedures and performance evaluations: to systematically advantage those highest on job-related knowledge, skills, and abilities (KSAs) over others. From an organizational standpoint, this type of discrimination is necessary and useful (Cascio & Aguinis, 2011).

On the other hand, personnel decisions are often influenced by characteristics that are job-irrelevant; for example, by the employee or applicant's sex (e.g., Ryan & Haslam, 2008; Turnage & Muchinsky, 1984), race (e.g., Dean, Roth, & Bobko, 2008; Roth, Bobko, McFarland, & Buster, 2008), age (e.g., Gordon & Arvey, 2004), weight (e.g., Roehling, Pichler, & Bruce, 2013; Rudolph, Weller, Wells, & Baltes, 2009), or sexual orientation (e.g., Croteau, 1996). Discrimination based on these characteristics creates a number of negative consequences for individuals, organizations, and society (Cox, 2001; Kane, 2003). For brevity, I will use the term "discrimination" hereafter to refer specifically to discrimination based on job-irrelevant personal characteristics.

Research on employment-related discrimination has typically focused on establishing the extent to which and under what conditions discrimination exists. There are a wide range of discriminable characteristics, some of which were mentioned above, and the majority of this research has concentrated on a single characteristic or, in some cases, on the interaction between two characteristics. Early discrimination research focused largely on sex (e.g., Arvey, 1979; Fidell, 1970; Goldberg, 1968; Rosen & Jerdee, 1974) and race (e.g., deJung & Kaplan, 1962; Farr, O'Leary, & Bartlett, 1971; Fox & Lefkowitz, 1974; Landy & Farr, 1973; Schmidt & Johnson, 1973; see Landy and Farr, 1980, for review). More recently however, interest in a wider range of characteristics (e.g., weight, Puhl & Heuer, 2009; Roehling, 1999; sexual orientation, Croteau, 1996; Kormanik, 2009; Lee Badgett, 1995) has grown.

One challenge to applying employment-based discrimination research is the lack of integration across specific research areas. In addition, comparatively little employment-based discrimination research has focused on the underlying processes by which employment decisions are made. One notable exception is the lack-of-fit model (Heilman, 1983, 2001), which proposes that discrimination occurs when the stereotypes evaluators hold of a person do not match the work requirements of the given job. Unfortunately, the lack-of-fit model has rarely been extended beyond sex discrimination (for exceptions, see Pichler, Varma, & Bruce, 2010, and Polinko & Popovich, 2001). However, research conducted outside of employment settings has provided valuable insight into the underlying processes of discrimination. For example, a considerable amount of cognitive and social psychology research has supported the importance of stereotypes (e.g., Allport, 1954; Devine, 1989; for reviews see Dovidio & Gaertner, 2010 and Macrae & Bodenhausen, 2000) and prototypes (e.g., Smith & Minda, 1998; Nosofsky, 1987) in evaluative judgments, although the types of judgments in these studies are seldom made with regard to discrimination in employment contexts.

The goal of this study is to integrate existing theory by proposing and testing a model of the underlying mechanism of discrimination, which can be applied across multiple discriminable personal characteristics. The model provides a flexible framework for understanding discrimination in personnel decision making, and this framework may provide valuable insight for (a) researchers studying both discrimination processes and specific types of discrimination, and (b) practitioners whose goal is to avoid discrimination in decision making.

In the following sections, I review theory and research on each of the specific components of the proposed model. An initial hurdle in developing a generalizable model of the underlying processes of discrimination is to be able to describe two basic types of discrimination through the same mechanism. Describing two different types of discrimination through the same model would provide initial support for the idea that there may be an underlying mechanism that generalizes across discrimination types. Thus, I begin by reviewing findings of workplace discrimination literature on the two discriminable characteristics examined in the present study: sex and weight. Second, I review relevant stereotype



research and theory (e.g., lack-of-fit model; Heilman, 1983, 2001). Third, I discuss the importance of prototypes to discrimination and review one particular prototype-based theory that has been applied to employment settings: implicit leadership theory (ILT; Lord & Maher, 1991). Finally, I integrate the above theories and present a prototype matching model as a potentially useful framework for describing the underlying mechanism of multiple different types of discrimination (e.g., sex, race, age, weight).

### **Sex and Weight Discrimination**

In an effort to begin to uncover a generalizable underlying mechanism of discrimination, this study focuses on two very different individual characteristics: sex and weight. These particular characteristics were chosen because both are individually well-studied, providing a broad basis of evidence to inform hypotheses. In addition, these two characteristics differ in multiple important ways, allowing for a strong test of a common explanatory mechanism.

First, sex and weight differ in legal considerations. That is, sex discrimination is currently prohibited under US employment law. Federal law prohibits discrimination based on a handful of individual characteristics that are believed to be job-irrelevant (i.e., sex, race, color, national origin, and religion, Title VII of the Civil Rights Act, 1964; disability, Americans with Disabilities Act, 1990; age, Age Discrimination Act, 1967). However, discrimination based on these characteristics remains a serious topic. For example, recent evidence has shown that sex discrimination continues to be evident in employment settings (e.g., Dean, Roth, & Bobko, 2008; Eagly & Karau, 2002; Heilman & Eagly, 2008; Lyness & Heilman, 2006). There is an array of other characteristics that are neither job-related nor protected by law, but have also been shown to factor into personnel decisions. Weight is one prominent example. Individuals' weight has been shown to influence a wide range of work-related evaluative judgments (Roehling, 1999; Roehling et al., 2013), even when controlling for target job qualifications (see Roehling et al., 2013).

Second, the stereotypes – “shared beliefs about a group” (Dovidio & Gaertner, 2010, p. 1084) – attributed to individuals differ based on sex and weight. Researchers have long argued that stereotypes play an important role in evaluations (e.g., Allport, 1954). Stereotypes often suggest a negative

connotation (Landy, 2008), but it is well accepted that stereotypes can be positive (Cuddy, Fiske, & Glick, 2008; see also Eagly & Karau, 2002). For example, females are typically ascribed both positive and negative stereotypes (e.g., Eagly & Karau, 2002; Fiske, Cuddy, Glick, & Xu, 2002). In contrast, those viewed as overweight are typically ascribed negative stereotypes (e.g., Roehling, 1999).

Third, the amount of control individuals are perceived to have with regard to these two characteristics differs. Sex is typically perceived to be out of the individual's control. In other words, most people are not perceived as having a choice to be male or female. This is not the case with perceptions of weight. That is, weight is often perceived to be under the individual's control, and the result of one's behavior (Allon, 1982; Crocker, Cronwell, & Major, 1993). Consequently, weight has been suggested to be the last characteristic on which it is socially acceptable to discriminate (see Puhl & Brownell, 2001). Below, I review the evidence of sex and weight discrimination.

**Sex discrimination.** Sex-based discrimination has been an important topic among social scientists for decades (e.g., Eagly, 1985), and research in organizational settings is no exception (e.g., Fidell, 1970; Eagly & Karau, 2002; Eagly, Makhijani, & Klonsky, 1992). Research has been conducted both in the lab and in the field. Lab-based studies have commonly used the Goldberg research paradigm (Goldberg, 1968). Goldberg research designs ask participants to evaluate "paper people" – that is, resumes or other job-related materials in which the sex of the rating target is manipulated (i.e., as being male or female). The rationale is that by controlling for job-related information, any resulting differences in judgments about targets must be due to the target's sex.

Research following this paradigm is abundant, and work-related outcomes such as performance evaluations, hiring decisions, and others have been examined. This has led to multiple meta-analyses of the lab-based research (Bowen, Swim, & Jacobs, 2000; Davison & Burke, 2000; Olian, Schwab, & Haberfield, 1988; Swim, Borgida, Maruyama, & Myers, 1989), which have consistently shown a small overall advantage for males ( $d < .10$ , with the exception of Olian et al.'s  $d = .41$ ).

A sizeable body of evidence comparing males and females also comes from field samples. Primary studies have examined a range of work-related outcomes including performance appraisals (see

Roth et al., 2012), ratings of incumbent potential (e.g., Turnage & Muckinsky, 1984), performance measures used for selection (Dean, Roth, & Bobko, 2008; Graves & Powell, 1995), and attitude measures (e.g., Graves & Powell, 1995). The evidence from field settings has been mixed with regard to the direction of effects. Some primary study evidence shows an advantage for males with regard to performance evaluations (e.g., Turnage & Muchinsky, 1984; Pulakos, White, Oppler, & Borman, 1989), evaluations of potential (Turnage & Muchinsky, 1984), and the probability of job offers (e.g., Graves & Powell, 1995). However, meta-analyses of field research have also shown an advantage for females in both performance ( $d = -.11$ ; Roth et al., 2012) and selection settings ( $d = -.19$ ; Dean et al., 2008).

Two points become salient from the lab- and field-based sex discrimination research described above. First, the practical significance of sex difference findings has typically been described as small (i.e., Bowen et al., 2000; Davison & Burke, 2000; Dean et al., 2008; Olian et al., 1988; Roth et al., 2008; Swim et al., 1989). However, these seemingly small practical differences have been argued to have greater long term effects (Agars, 2004; Eagly, 1995). Martell, Lane, and Emrich (1996) used a simulation study to model the cumulative long term effects of sex discrimination across several promotion cycles. They sought to describe the effects that can build up over time in the presence of small, but consistent levels of discrimination. They simulated data for a cohort of hypothetical employees across eight promotion periods to represent moving from the bottom of the organizational ladder to the top. With sex set to account for only 1% of the variance in a given round of promotion decisions (meaning 99% of the variance in decisions were due to differences other than employee sex), an initial cohort that was 53% female upon entering the organization led to a top management cohort that was only 35% female. With sex set to account for 5% of the variance in promotion decisions, the initial cohort of 58% females upon entering bottom levels of the organization led to a top management cohort that was only 29% female. These findings mirror labor market statistics showing similar employment ratios for males and females in low-level positions, but sex disparity in leadership positions (U.S. Bureau of Labor Statistics, 2004).

Second, the directionality of findings has been inconsistent. As described above, the evidence is mixed regarding whether females are targets or beneficiaries of sex-related bias. A key moderator of the

relationship between sex and performance and selection criteria is job type (e.g., Davison & Burke, 2000). That is, females are advantaged in “feminine” jobs (e.g., nurse, secretary, librarian), and disadvantaged in “masculine” jobs (e.g., engineer, carpenter). Thus, it appears that the discussion of sex-based discrimination requires consideration of the job. In fact, these findings have led to the consideration of work roles in the formulation of job fit theories (Eagly & Karau, 2002; Heilman, 1983).

**Weight discrimination.** A considerable amount of research also exists on weight discrimination. Research has been conducted in both the lab (e.g., Bellizzi & Hasty, 2001; Brink, 1988; Ding & Stillman, 2005; Larkin & Pines, 1979; Polinko & Popovich, 2001; Rothblum, Miller, & Garbutt, 1988) and in the field (e.g., Averett & Korenman, 1996; Baum & Ford, 2004; Collins & Zebrowitz, 1995; Judge & Cable, 2011; Leigh & Berger, 1989; Morris, 2006; Zhang & Wang, 2004). This research has examined a wide range of criteria, including personality and behavioral dimensions (e.g., Larkin & Pines, 1979), attitudes (e.g., Kennedy & Homant, 1984), hiring decisions (e.g., Larwood, 1995), coworker desirability (e.g., Jasper & Klassen, 1990), income (e.g., Register & Williams, 1990), and employment status (e.g., Tunceli, Li, & Williams, 2006). Finally, multiple meta-analyses have recently summarized these findings both in the lab ( $d = .36$  and  $.52$ , Roehling et al., 2013; Rudolph et al., 2009, respectively) and in real world settings ( $d = .02$ , Vanhove & Gordon, in press). These meta-analyses have also found a number of moderators that affect the magnitude of weight discrimination. One important moderator is target sex. That is, overweight females have been shown to experience greater weight discrimination than do overweight males (Roehling et al., 2013).

**Summary.** Evidence suggests that discrimination as a function of both of these characteristics does exist and is relatively small at the aggregate level, but the magnitude of discrimination due to sex and weight is exacerbated under certain conditions. As described above, one particularly important moderator of sex-based discrimination is job type (i.e., masculine/feminine; Davison & Burke, 2000), and one particularly important moderator of weight-based discrimination is target sex (i.e., male/female; Roehling et al., 2013). I consider the role of job type in the design of this study, a topic I discuss in more

depth in later sections, as well as examine the interaction between target sex and weight in predicting discrimination.

It is important to pursue establishing a framework that can be applied to any type of discrimination as it will have utility for both research and practice. In the following sections I review literature relevant to such a process model. I begin by reviewing stereotype research, as stereotypes have been shown to play a central role in instances of discrimination.

### **Stereotyping**

In this section I review basic theory behind stereotyping, beginning with the concept of categorization – the starting point in stereotypical thinking. Then I describe the specific stereotypes attributed to females and overweight individuals. Finally, I describe contemporary stereotype theory relevant to the current study.

**Categorization.** The onset of stereotypical thinking begins with categorization. That is, in order to associate attributes with individuals based on salient characteristics such as sex or weight, we must first associate the individuals with those categories (i.e., male/female, normal weight/overweight). Only after an individual is categorized into a stereotyped group are the stereotypes attributed to that group generalized to the individual (e.g., Macrae, Stangor, & Milne, 1994; Macrae & Bodenhausen, 2000).

The functional role of social categorization was realized early on (Lippmann, 1922). Early theorists understood that (a) humans function within complex environments and social situations, and (b) humans have limited cognitive resources to process social stimuli (e.g., Allport, 1954; Bruner, 1957). Efficiently and effectively processing the vast amount of social information is a great challenge, if not altogether impossible, and it is unlikely that perceivers can, or are willing to, engage in the tedious process of evaluating all available information (see Dovidio & Gaertner, 2010; see also Macrae & Bodenhausen, 2000). Thus, simplifying the process of perceiving others through social categorization plays an important role in our ability to manage the world around us and efficiently function within it (Allport, 1954; Brewer, 1988; Fiske & Neuberg, 1990; also see Dovidio & Gaertner, 2010, and Macrae & Bodenhausen, 2000) by allowing us to perceive others through abstract, yet simplified categories. In turn,

categorizing the people we encounter based on salient characteristics is believed to free up valuable cognitive resources (Sherman et al., 1998).

Social categorization is a process necessary for efficient social functioning, and one that can be very effective at times. Yet, its efficiency can also lead to ineffective judgments, especially in the case of personnel evaluation where the goal is to accurately evaluate job applicants and incumbents. Ideally, evaluators should focus exclusively on differentiating individuals based on the observable evidence of KSAs related to job performance (e.g., Campbell, McCoy, Oppler, & Sager, 1993). In reality, social perception tasks are complex and challenging (Macrae & Bodenhausen, 2000). In addition, some personnel evaluation settings, such as selection, provide evaluators with limited job-relevant, observable evidence, which may lead evaluators to rely, at least partially, on salient but job-irrelevant characteristics such as an applicant's sex, weight, or both. Relying on salient physical characteristics, as opposed to job-related information, provides evaluators with an efficient way to compare and differentiate individuals, but it also may lead to inaccurate personnel decisions (e.g., Heilman & Haynes, 2008) such as hiring less qualified applicants or promoting the wrong employees.

**Ascribing stereotypical attributes.** Categorization, alone, does not necessarily lead to discrimination. For example, an evaluator may categorize a target as “female” or “overweight” yet their evaluative judgments may not be biased. Problems stem from instances when evaluators focus on job-irrelevant characteristics (e.g., weight) and infer stereotypical job-related characteristics (e.g., lazy) that are not actually observed. Thus, discrimination is more directly related to the stereotypes evaluators attribute to the female or overweight categories (Brewer, 1988; Fiske & Neuberg, 1990), as opposed to the categories themselves.

Early researchers (e.g., Allport, 1954) argued that the more negative the stereotypical attributes associated with the category, the greater the discrimination. We now know that the relationship between stereotypes and evaluative outcomes is more complex. I discuss this issue in more detail later in this section. For now, let it suffice to say that there is broad consensus that stereotypes are an important starting point in the process of discrimination with regard to both sex (Allport, 1954; Heilman, 1983;

Eagly & Karau, 2002) and weight (Crocker, Cornwell, & Major, 1993; Puhl & Heuer, 2009; Roehling, 1999).

**Stereotypes of females and overweight individuals.** As mentioned above, stereotypes are not necessarily negative. Female stereotypes include both positive and negative attributes, for example: “communal”, “warm”, “helpful”, “kind”, and “nice” (Heilman & Eagly, 2008; Eagly & Karau, 2002; Fiske et al., 2002), but also “dependent” and “incompetent” (Fiske et al., 2002). Social perceivers, including personnel evaluators, often label target females with these characteristics, regardless of whether or not they actually observe these attributes in the target’s behavior (Eagly & Karau, 2002). Despite the fact that some female stereotypes are positive, they still often lead to negatively biased evaluations.

Although females are labeled with both positive and negative stereotypes, this is not the case for overweight individuals. Instead, the stereotypical attributes ascribed to these individuals are largely negative (Allon, 1982). A number of studies have captured stereotypical attributes associated with being overweight, and a narrative review by Roehling (1999) summarizes these findings in the employment context. Stereotypes of overweight individuals that have been found include: being lazy; low in competence, ability, and skill; sloppy; more likely to have negative personality traits; lacking self-discipline and control; and having poor work habits. Consistent findings of negative stereotypes toward overweight individuals have led a number of researchers to make a connection between weight stereotypes and discrimination in employment settings (e.g., Crocker et al., 1993; Finkelstein, Frautschy Demuth, & Sweeney, 2007; Roehling, 1999).

**“Fit” Theories.** Early theory on stereotyping suggested that negative evaluative outcomes for individuals resulted from the negative stereotypes associated with the groups to which the individuals belonged (e.g., Allport, 1954). In the case of the consistent, negative stereotypes of overweight individuals, this logic appears to hold for describing findings of weight discrimination. However, stereotypes and evidence of discrimination towards some groups do not appear to be congruent with this model. As described above, females are ascribed both positive and negative attributes. Thus, the experience of discrimination by females may not be purely the result of negative stereotypes. More

recent theory has suggested that the relationship between stereotypes and evaluative outcomes depends on the congruence (i.e., fit) between the stereotypical attributes ascribed to the evaluative target and the work role requirements of the job (Heilman, 1983, 2001; Eagly & Karau, 2002).

The lack-of-fit model (Heilman, 1983, 2001) directly posits the idea of “fit” as a way of explaining the complex relationship between stereotypes and discrimination. Heilman (1983, 2001) proposes that the less congruent the characteristics attributed to the target are with work role requirements the lower the evaluation the target will receive. A common example of the application of this model is in the leadership context (i.e., role incongruence theory; Eagly & Karau, 2002). Characteristics often associated with successful leaders include frankness, aggressiveness, and assertiveness (Heilman, Block, Martell, & Simon, 1989). Although common female stereotypes such as warm, helpful, kind, and nice are generally viewed as positive, they are somewhat incongruent with the characteristics typically associated with leaders (Heilman et al., 1989; Schein, 1973). Thus, the lack-of-fit model (and specifically, role incongruence theory) provides the rationale for why females, with both negative and positive stereotypical attributes, have been shown to experience discrimination. In addition, the lack-of-fit model provides an explanation for the findings of Davison and Burke (2000) and others (see Martinko and Gardner, 1983) regarding the role of sex-typed jobs (where female targets were only disadvantaged when evaluated for masculine jobs).

The “fit” theories represent advances over early stereotype theory in explaining the mechanisms that underlie sex-based discrimination, and this may generalize across types of discrimination. However, many gaps remain. First, the proposition that discrimination results from incongruence between target stereotypes and work requirements suggests a moderated mediation model. That is, “fit” theories suggest that evaluators attend to the evaluative target’s specific, job-irrelevant individual characteristics, which leads to ascribing stereotypical attributes (e.g., warm, incompetent) based on those characteristics (e.g., sex, weight). These stereotypical attributes are then suggested to interact with work role requirements (e.g., coldness/warmth, competence/incompetence) to predict evaluative outcomes. However, existing research has yet to explicitly test this process model. Instead, the theory has been based on the integration



of research findings using group membership (e.g., male/female, normal weight/overweight) as a predictor of (a) stereotypical attributes (Crocker et al., 1993; Harris, Harris, & Bochner, 1982; Koenig, Eagly, Mitchell, & Ristikari, 2011; Shapiro, King, & Quinones, 2007), and/or (b) evaluative outcomes for specific work roles (Bowen et al., 2000; Davison & Burke, 2000; Dean et al., 2008; Finkelstein et al., 2007; Roehling et al., 2013; Swim et al., 1989). In other words, research has not explicitly tested the mediational hypothesis that possessing certain characteristics (e.g., being female or overweight) leads evaluators to ascribe stereotypical attributes to those individuals, which in turn leads to discrimination.

Second, “fit” was initially developed as an explanation of sex-based discrimination. Although “fit” has potential for explaining discrimination across multiple characteristics, few studies have applied “fit” outside of the sex discrimination literature (e.g., Pichler, Varma, & Bruce, 2010; Polinko & Popovich, 2001). Thus, an important step in establishing the generalizability of the idea of “fit” to explain discrimination is to extend this theory to applicant characteristics other than sex.

Third, central to “fit” theory is the idea that work requirements, such as those that would be defined through a skill-based job analysis – that is, a systematic approach to uncovering the most important KSAs for a particular job (see Brannick, Levine, & Morgeson, 2007) – are consistent across evaluators. This may be an oversimplification and somewhat imprecise for understanding the underlying process of discrimination, as it ignores the variation that exists in individual evaluators’ mental models of a successful employee. Indeed, evaluators often have different ideas of what is necessary to be successful in a given job. This is a problem that has driven half a century of research on increasing interrater agreement in performance appraisals and subjective performance measures (e.g., Gordon & Medland, 1965; Woehr & Huffcutt, 1994; see Landy & Farr, 1980). Thus, it may not be the work role requirements themselves that evaluators use in determining how well targets’ stereotypes “fit”. Instead, it is more likely that evaluators rely on their own individual perceptions of what attributes are necessary for success in a particular job. Evaluators’ specific prototypes, as opposed to standardized work role requirements, may provide valuable insight into the processes proposed by “fit” theory. Research and theory focusing

on prototypes in the existing IO literature provides valuable insight into the potential role of prototypes when applied to “fit” theory.

### **Prototypes**

A prototype is a mental representation of a category member (Homa, 1984). Humans have prototypes for all types of constructs, including females and overweight individuals, and we use these prototypes to evaluate how well the individuals we perceive fit into those categories (Nosofsky & Zaki, 2002). We also have prototypes for jobholders of specific jobs such as “salesperson”, or more specifically “car dealer”, and possibly even more specifically “car dealer at Acme Automotive”. These prototypes are particularly important to evaluators tasked with selecting applicants best suited for a “salesperson” or “leadership” work role. Prototypes are related to, but distinct from, stereotypes. Whereas we ascribe stereotypes to individuals based on the groups under which we categorize them (e.g., that person is a female, therefore she must be warm), we compare individuals to prototypes to determine how well they fit into those categories (e.g., that female is cold; she doesn’t fit my prototype of females).

**Implicit leadership (and followership) theories.** Implicit leadership theory (ILT; Lord & Maher, 1991) proposes that individuals create complex prototypes of leaders, and argues that subordinates rely on these prototypes when evaluating their own leaders. A number of studies have provided evidence that evaluators use prototypes in making judgments regarding the effectiveness of specific leaders (e.g., Cronshaw & Lord, 1987; Foti, Fraser, & Lord, 1982; Fraser & Lord, 1988; Lord, De Vader, & Alliger, 1986; Lord, Foti, & Phillips, 1982; Nye & Forsyth, 1991). Research has extended the idea of ILT to followers (implicit follower theories; IFTs; e.g., Engle & Lord, 1997; van Gils, van Quaquebeke, & van Knippenberg, 2010; Sy, 2010), suggesting that prototypes are used to evaluate the full range of individuals we encounter in the workplace. I will extend this to suggest that individuals have prototypes for all types of specific work roles, including prototypes for each of the positions for which hiring or promotion decisions are made.

As described earlier, “fit” theory has assumed a static mental model for a given work role to compare to evaluative target stereotypes in determining person-job “fit”. However, prototype theories,

such as ILT and IFT, suggest that this is an oversimplification. Alternatively, prototype theories would suggest that a determination of “fit” would be more precisely captured by comparing evaluators’ prototypes of a specific work role to the stereotypes they attribute to evaluative targets. Thus, I borrow the assumptions put forth by ILT (and IFT) as a way of building on “fit” theory.

### **Stereotype Content Model (SCM)**

Much research over recent decades has focused on the process of stereotyping (e.g., see Macrae & Bodenhausen, 2000). The goal has been to uncover systematic principles behind stereotyping that generalize across time, context, and group (Eckes, 2002; Fiske et al., 2002). However, Zebrowitz (1996) pointed out that a comprehensive understanding of stereotyping requires more than just an understanding of the process and highlighted the importance of understanding stereotype content – that is, the stereotypical attributes associated with different groups. Building on Zebrowitz’s proposition, I suggest that the value of understanding stereotype content is even further reaching in that stereotype content plays an important role in understanding the process of discrimination. That is, to understand the underlying mechanisms of discrimination, it is necessary to understand the stereotypes that are ascribed to different groups. In addition, both stereotypes *and* prototypes play central roles in the model being tested in the present study, and a content model for which the two can be directly compared is also necessary. That is, it is important to be able to examine the (mis)match between content in evaluators’ stereotypes and prototypes. An efficient and parsimonious model of stereotype content (to be compared to prototype content) remains to be somewhat of an issue for researchers. There are as many stereotype labels as there are adjectives. Personality research has provided some guidance through various useful models that narrow down the mass of descriptive terms into personality factors (e.g., Cattell, 1943; Costa & McCrae, 1995). More directly relevant, however, is research on social judgment. Multiple models of social judgment have proposed that humans’ judgments of others can be described through two factors (e.g., Abele & Wojciszke, 2007; Bakan, 1966; Rosenberg, Nelson, & Vivekananthan, 1968). Recently, Fiske et al. (2002) have applied this two factor model to stereotype content. The stereotype content model (SCM) was developed in an aim to capture the content dimensions people use in stereotyping others, and it

provides possibly the most useful framework for comparing stereotype and prototype content in explaining the underlying process of discrimination.

The SCM suggests that stereotypes can be categorized through two orthogonal dimensions: warmth and competence. There appears to be strong convergence between the SCM and other models aiming to capture the content of personality, social perceptions, and human judgments through two overarching factors (Judd, Hawkins, Yzerbyt, & Kashima, 2005). For example, the description of Bakan's (1966) general personality dimensions of agency (self-orientation) and communion (other-orientation) largely aligns with competence and warmth, respectively; Rosenberg, et al. (1968) created a two-factor structure from 64 personality characteristics that they labeled as intellectual and social; Wojciszke (2005) labeled factors of self- and other-perceptions as competence and morality; within the organizational research realm, the Ohio State Leadership Studies (e.g., Fleishman, 1953) used initiating structure (i.e., task-orientation) and consideration (i.e., interpersonal-orientation) to describe perceptions of leadership behaviors. Cuddy et al. (2008) propose that these models often "boil down" to warmth and competence dimensions, despite having different labels.

The consensus among researchers regarding the similarity among two factor structures provides confidence for using this framework in the model tested in the present study. Evidence has also accumulated specifically supporting the SCM. For example, various groups' warmth and competence profiles have been consistently replicated in US samples (e.g., Asians are consistently perceived as being highly competent, but not warm; Cuddy et al., 2007; Eckes, 2002; Fiske et al., 2002). In addition, stereotypes have been shown to differ as a function of warmth and competence perceptions across seven European and three Asian nationalities (Cuddy, Fiske, Kwan, et al., 2009). Both correlational (e.g., Cuddy et al., 2007; Fiske et al., 2002) and experimental (Caprariello, Cuddy, & Fiske, 2007) research has also shown that perceived group status and competitiveness are predictors of group competence and warmth perceptions, respectively, and these relationships appear to be generally consistent across international samples (see Cuddy et al., 2009).

The SCM proposes that groups can fall under any combination of the two dimensions. Groups perceived as low in competence, but high in warmth are referred to as “pitied” groups. Groups perceived as high in competence, but low in warmth are referred to as “envied” groups. Those perceived as low in competence, and low in warmth are referred to as “contempt” groups. The SCM also suggests that different combinations elicit different emotions and prejudices (Cuddy, Fiske, & Glick, 2008), and that discrimination and bias result from these emotional and perceptual responses. Support has been found for the responses different combinations elicit (Cuddy et al., 2007; Fiske et al., 2002), and evidence also suggests stereotype content to predict behavioral tendencies. That is, warmth stereotypes have been shown to predict both active facilitation (e.g., helping) and harm (e.g., harassing) ( $r = .73$  and  $-.55$ , respectively), and competence stereotypes to predict both passive facilitation (e.g., admiring) and harm (e.g., holding contempt) ( $r = .77$  and  $-.68$ , respectively). What has not yet been examined is the effect of warmth and competence stereotypes on work-related outcomes. This study takes an important step in this regard by extending these findings to outcomes relevant to workplace discrimination.

The SCM provides a parsimonious and practical framework to directly compare the (in)congruence between evaluators’ stereotypes and prototypes as a way to explain discrimination. That is, by comparing evaluators’ perceptions of warmth and competence dimensions with regard to both their stereotypes of applicants and their prototypes of qualified job holders for a given job I can directly examine how stereotypes interact with prototypes in explaining discrimination.

### **The Prototype Matching Model**

The research and theory described above suggest a prototype matching model to explain the underlying process by which discrimination occurs in personnel evaluation. In the realm of stereotype research, the development of the lack-of-fit model (Heilman, 1983, 2001) amended the more basic position that more negative stereotypes lead to more negative evaluations (Allport, 1954) by taking into consideration (in)congruence between target stereotypes and work roles. However, as described above, “fit” theory’s conceptualization of work role attributes may be an oversimplification, as it ignores individual differences among evaluators’ perceptions regarding what is necessary to be successful in a

given job. Thus, the prototype matching model builds on “fit” theory and may more precisely represent the process of discrimination by capturing individual evaluators’ unique prototypes.

In sum, both stereotype- and prototype-based theories offer rich bodies of evidence in their respective areas. However, these literatures have remained largely separate. In this study, I integrate stereotype-based “fit” theories and prototype-based ILT (and IFT) in an aim to describe the interaction between stereotypes and prototypes as the underlying mechanism of discrimination. The prototype matching model represents a moderated mediation model. The model is presented in Figure 1 below.

Although evaluators attend to job-related KSAs when evaluating targets, this model begins with the proposition that evaluators often also attend to at least some job-irrelevant characteristics. Thus, for this model to have value in explaining discrimination, the presence of discrimination must be established. I tested the prototype matching model (Figure 1) using two characteristics: sex and weight. Based on evidence regarding discrimination I reviewed above, I expect:

*Hypothesis 1a:* Male targets will receive more positive evaluations than females.

*Hypothesis 1b:* Normal weight targets will receive more positive evaluations than overweight targets.

Evidence also suggests job-type as a potential moderator of the effects of target sex on evaluative outcomes. For this reason, I included two different positions (a sales position and a customer service position) in the study design. The sales and customer service positions were chosen due to (a) the masculine and feminine features, respectively, associated with these positions, and (b) the familiarity that most people have with these two types of positions. Specific hypotheses are as follows:

*Hypothesis 1c:* Females will receive more positive evaluations than males for the service position, and males will receive more positive evaluations than females for the sales position.

In addition, previous evidence has suggested greater discrimination against overweight females than overweight males (Roehling et al., 2013; Vanhove & Gordon, in press). Thus, hypothesis 1d is as follows:

*Hypothesis 1d:* Overweight males will receive more positive evaluations than overweight females.

The remainder of the hypotheses relate to the processes by which discrimination occurs. Path 2 in Figure 1 proposes that evaluators ascribe stereotypical attributes to targets based on characteristics such as sex and weight. The SCM proposes that stereotypes can be broken down into warmth and competence dimensions (e.g., Cuddy et al., 2002), and research using that model suggests different warmth and competence stereotype profiles based on sex and weight. Using the SCM framework to capture stereotypes, I expect:

*Hypothesis 2a:* Females to be rated as warmer than males, and males to be rated as more competent than females.

*Hypothesis 2b:* Normal weight targets to be rated as both warmer and more competent than overweight targets.

In addition to males and females (normal weight and overweight individuals) being associated with different stereotypes, I expect these stereotypes to predict evaluative outcomes (Path 3 in Figure 1). I included both coworker desirability and likelihood of success outcomes in this study. It is plausible to expect that those viewed as warm, friendly, and good-natured (i.e., warmth stereotypes) may also be evaluated as more desirable coworkers, due to the interpersonal-oriented nature of these stereotypes. It is also plausible to expect that those viewed as competent, intelligent, and confident (i.e., competence stereotypes) may also be evaluated as more likely to succeed, due to the achievement-oriented nature of these stereotypes. The formal hypotheses regarding these relationships are as follows:

*Hypothesis 3a:* Warmth stereotype ratings will be more strongly related to coworker desirability evaluations than will competence stereotype ratings.

*Hypothesis 3b:* Competence stereotype ratings will be more strongly related to likelihood of success evaluations than will warmth stereotype ratings.

As a final step in the mediation process proposed by Baron and Kenny (1986), target characteristics (i.e., sex and weight) should not significantly predict evaluative outcomes when examined along with stereotype ratings, if full mediation is present.

*Hypothesis 4a:* Stereotype ratings will mediate the relationship between target sex and evaluative outcomes.

*Hypothesis 4b:* Stereotype ratings will mediate the relationship between target weight and evaluative outcomes.

Thus far, each of the propositions proposed above is in line with “fit” theory’s approach to describing the process of discrimination. The final path (Figure 1, Path 5) of the prototype matching model, however, is what differentiates the model from “fit” theory. Using the framework of ILT (and IFT), the final piece of the prototype matching model suggests that the relationship between stereotypes and evaluators’ evaluations are moderated by the evaluators’ perceptions of the prototypical jobholder. That is, the less congruent the evaluator’s stereotypes of the target are with the prototype the evaluator holds for the position (i.e., the worse the stereotype-prototype match) the greater the negative bias, as reflected through lower evaluations. In contrast, the more congruent the evaluator’s stereotypes of the target are with his or her prototype, the more positive the evaluation. To test this hypothesis, participants made “warmth” and “competence” ratings for their prototypical jobholder for both the sales and the customer service positions, and these prototypical warmth and competence ratings were compared to evaluators’ warmth and competence stereotype ratings of each target.

*Hypothesis 5:* The relationship between evaluators’ stereotypes of targets and their evaluations of those targets will be moderated by the prototypes they hold for the given job.

### **Context of the Current Study**

The hypotheses presented above test each of the propositions of the prototype matching model (Figure 1). This study is the first in a program of research aimed to better understand the complex processes involved in discrimination in personnel decision making. As such, I employ a Goldberg paradigm design (i.e., fictitious job applicants as evaluative stimuli) with a sample of adult participants, who likely have little real world personnel evaluation experience. I do so to establish whether the model presented above *can* explain discrimination (Mook, 1983). If the processes described in the prototype matching model do reflect evaluators’ cognitive processes, the effects will be more easily identified in a



lab setting, which typically offers greater experimental control than field settings using real world decision makers. Thus, my focus in this study is on establishing internal validity, which is suggested to have greater importance during the early stages of theory development than external validity – that is, without establishing the internal validity of a model there is little value in attempts to generalize it to the target population (Anderson, Lindsay, & Bushman, 1999).

Discrimination can exist as the result of any type of personnel decision, the two most common of which are performance evaluation and selection. Discrimination may be especially prevalent in selection settings, as opposed to job performance evaluation settings. This is because there is likely less job-relevant information on applicants available in selection situations, as compared to when actual performance can be observed. For this study, I focus on the early stages of the selection process where the only job-relevant information available is applicant resumes, increasing the likelihood that evaluators will focus on characteristics such as target sex and weight in categorizing targets (Martinko & Gardner, 1983; Puhl & Heuer, 2009).

## CHAPTER 2: METHOD

### **Participants**

Data were collected from 610 participants (56.9% Female) through a recruitment announcement posted on Amazon's Mechanical Turk (MTurk). MTurk ([www.MTurk.com](http://www.MTurk.com)) is an online workplace where "workers" complete tasks posted by "requesters". MTurk has become a popular medium for social scientists to collect data, due to the availability of its large pool of participants. Research has shown that MTurk samples may provide greater demographic diversity than either college student or other internet samples (Buhrmester, Kwang, & Gosling, 2011). However, data quality issues regarding MTurk samples have also been raised (see Goodman, Cryder, & Cheema, 2012). To address these concerns I conducted a series of data cleaning procedures (described in greater detail below). As a result of the data cleaning procedures, the final sample including for analysis was 499 (61.7% Female). Participants who completed the study were compensated \$0.25-\$0.75 for their participation.

### **Procedure**

Participants completed the study online. A link to the study was provided on the MTurk recruitment posting. Those who followed this link were met with the study consent form, and were prohibited from advancing to the study materials unless they provided consent to participate, which was indicated by clicking the "consent to participate" button at the end of the consent form. Those who consented were then provided a brief overview (Appendix I), which set up the context of the study and task. First, participants were informed that they would be playing the role of decision maker in charge of conducting an initial screening of applicants applying to a fictitious organization. The organization was described as a healthcare organization named StayActive, Inc. Participants were provided with a general description of the company then informed that StayActive was hiring for two positions – a sales position and a customer service position – and that their task was to conduct an initial screening of applicants for each of these jobs, and in doing so respond to a number of questions regarding their perceptions of each applicant.

Next, participants were presented with a description of the first position, which was randomly determined to be the sales or customer service position, for which they would evaluate target applicants (Appendix II). After viewing the first job description participants were then given the opportunity to describe their ideal applicant for the first job in an open-ended response, then they were presented with “warmth” and “competence” items (Appendix III), and asked to “indicate how well each word applies to their ideal applicant”. These ratings served as participants’ prototype for the first job.

Participants were then presented with four LinkedIn profiles, representing the four applicants for the first position. Each LinkedIn profile included a resume indicating previous experience and was accompanied by a photo of either a(n) (a) overweight male, (b) overweight female, (c) normal weight male, or (d) normal weight female. Photos and LinkedIn profiles were randomly paired across participants, and the order in which these applicant photo-resume pairs were presented was also randomized. Participants viewed each LinkedIn profile separately. After viewing each applicant, participants’ rated the applicant on the same warmth and competence measures used to assess participants’ prototypes. However, these warmth and competence ratings would represent participants’ stereotypes of each applicant. After completing the warmth and competence ratings for a given applicant and before advancing to the next applicant LinkedIn profile, participants were given the opportunity to add any additional thoughts regarding the applicant using an open-ended response, then asked to evaluate the applicant with regard to the two outcome measures (coworker desirability and likelihood of success; Appendix III). Participants completed this process separately for each of the four applicants.

Once participants completed evaluating each of the four applicants for the first position, they were presented with the second job description and repeated the process described above, with a new set of applicants (i.e., resumes and photos) to the second position. After completing the process for the second position participants completed a conscientiousness measure, indicated their sex, and responded to four dichotomous items regarding previous work experience. Finally, participants were provided a study completion code which they entered on the MTurk recruitment page to receive compensation for their participation.

## **Materials**

**Applicant materials.** Eight separate LinkedIn profile resumes were used. Four of these were specific to the sales position and four were specific to the customer service position. The applicant names were androgynous to allow for random pairings of the LinkedIn profiles with either male or female stimulus photos. Eight stimulus photos were included (two photos for each of the four sex x weight combinations). Randomization was used to determine which of the, for example, overweight male photos would be presented as an applicant for the sales position and which would be presented as an applicant for the customer service position. This was done in order to (a) minimize the potential of artifactual error that could result from presenting the same applicant photo for the same position across participants, and (b) to ensure that the same photo was not presented to a participant for both positions.

**Pilot study of photos and resume materials.** A series of photos of overweight and normal weight males and females who appear to be in their early 20s (the age at which many college seniors are applying for entry level jobs, such as those used in this study) were gathered from stock photo websites. These photos were then pilot tested among 26 undergraduate students. Ratings of each photo were made on a 5-point scale with regard to the photo model's weight and attractiveness. Across the 16 photos that were pilot tested, the two from each of the four sex x weight conditions that showed (a) the greatest mean difference in weight ratings, and (b) the greatest similarity in attractiveness ratings were chosen as the stimulus applicant photos for inclusion in the study. Attractiveness was controlled for in order to isolate the effects associated with applicant weight, as applicant weight and attractiveness likely covaried. Descriptives of pilot test ratings of all photos are presented in Table 1.

Resume materials were also pilot tested (see Table 2 for descriptives). This was done to control the effects of the "strength" of resumes on participant ratings of applicants and help ensure that participant ratings were based on applicant sex and weight. One set of credentials (e.g., education, previous experience, extracurricular activities) was created for the sales position applicant resumes, and a second set was created for the customer service applicant resumes. Credentials were pilot tested among the 32 undergraduates used for pilot testing applicant photos. Pilot test participants rated each credential

on the nine warmth and competence items. These data were then aggregated and the four resumes for each job were created so that each resume was associated with similar levels of perceived warmth and competence. Stimulus materials (photos and an example resume) are presented in Appendix IV.

## **Measures**

**Warmth and Competence.** A set of nine adjectives used in previous stereotype content model (SCM) research by Cuddy et al. (2002) was used to measure warmth and competence. Five of these items reflect competence and four reflect warmth. These items were used to collect both prototype and stereotype ratings, which resulted in slightly different directions being presented to participants based on the context in which items were presented. The directions associated with prototype and stereotype ratings are presented along with the items in Appendix III. For both prototype and stereotype ratings responses were made using a 7-point Likert-type scale (1=not at all, 7=extremely).

**Control Measures.** Nine items from the Big Five Inventory (BFI; John, Donahue, & Kentle, 1991) were used to assess participants' conscientiousness on a 7-point scale (1=strongly disagree, 7=strongly agree). Participants' sex and four questions regarding participants' previous experience with formal evaluation, and working in sales, customer service, and the healthcare or pharmaceutical industry were included as demographic controls. Each of these items was dichotomous. Control measures are presented in Appendix V.

**Criterion measures.** Two criterion measures, created for use in this study, were included to reflect different aspects of evaluation and employee perceptions. The first set of items aimed to capture participants' desirability to have each applicant as a coworker (coworker desirability). The coworker desirability measure included four items (two reverse-coded), and ratings were made using a 7-point Likert scale (1=not at all, 7=extremely). The second set of items aimed to capture participants' perceptions of whether each applicant would be successful in the job (likelihood of success). The likelihood of success measures also included four items (one reverse-coded), and used the same response scale and options as the coworker desirability measure. Ratings were made using the same response options as the coworker outcome items. Criterion measures are presented in Appendix VI.

## **Data Cleaning**

The model I am testing inherently relies on evaluators' use of stereotypes as a heuristic to make judgments, and participants completed this study online in an unproctored setting for monetary compensation. This context requires a tender balance to be struck. On the one hand, evaluators' use of heuristics, such as stereotypes, as a cognitive shortcut in effortful responding is an important piece of the prototype matching model. On the other hand, responding without putting forth effort, such as to receive compensation, introduces measurement error. Thus, I used multiple data cleaning procedures based on the recommendations of Huang, Curran, Keeney, Poposki, and DeShon (2012) in order to minimize potential sources of error, while preserving data reported from those who simply used stereotypes but were engaged in the task.<sup>1</sup> Specifically, data were screened based on response (in)frequency, completion time, and invariant responding.

**Response frequency.** Participants who failed to respond to a high percentage of the total study questions may have lacked the motivation to respond accurately throughout their participation in the study. I conducted the response frequency screen using a 95% cutoff. There were a total of 169 Likert-scale questions included as part of this study, and participants were required to respond to at least 161 (95%) of these questions in order to avoid having their data excluded. A total of 89 (14.6%) participants failed to respond to 95% of the Likert-scale questions and their data were removed prior to my conducting the hypothesis testing analyses.

**Completion time.** There was likely a minimum amount of time that was necessary in order for participants to receive the manipulation in this study – that is, to be able to attend to job descriptions and resumes and read, comprehend, and respond to each of the study questions in an effortful fashion. Those who completed the study in less than the estimated minimum completion time may have been unlikely to provide valid responses. Given that this study included a cover story, as well as multiple job descriptions and applicant resumes in which participants were asked to review, estimating the minimal study completion time was less straightforward than survey-based research, in which minimum study

completion time can be estimated simply by multiplying the number of questions in the study by the completion time for a single question.

In order to determine the minimum completion time for this study I directed two pilot testers to complete the study as fast as possible while fully reading and comprehending all directions and study questions. Each pilot tester completed this task three times. Each time pilot testers' completion time fell between eight and 10 minutes. In addition, I manually examined the completion times of the 610 participants in the original dataset, comparing completion time to whether or not participants had failed to pass the 95% response frequency screen. The assumption is that there should be a strong relationship between completion time and the number of questions completed. Of the 27 participants who completed the study in less than eight minutes, 16 (63.0%) had already been screened out due to response infrequency. Of the 34 participants who took between eight and 10 minutes to complete the study, 10 (29.4%) had already been screened out due to response infrequency. Of the 27 participants who took between 10 and 12 minutes to complete the study, 4 (14.0%) had already been screened out due to response infrequency. Taking the initial pilot test data, supported by descriptive data among the actual study data, I chose to use eight minutes as the minimum cutoff for completion time. This resulted in the data from 10 additional participants (1.9%) being removed from the dataset.

**Response invariance.** As a final data screening tool I examined cases for a lack of variance in responses. It is somewhat unlikely that participants who responded honestly and accurately would have provided the same response to each item within a given measure. Moreover, it is very unlikely for this to be the case within measures including both positively- and negatively-worded items. I employed a number of measures that included both forward- and reverse-coded items. First, the 9-item conscientiousness measure (used as a control measure) included four reverse-coded items. In addition, participants completed two 4-item outcome measures (coworker desirability and likelihood of success) with regard to each of the eight applicants ( $2 \times 8 = 16$  outcome observations). The coworker desirability measure included two reverse-coded items, and the likelihood of success measure included a single reverse-coded item. In order to examine data for invariant responses I first created a set of 17 new

variables (one conscientiousness measure and 16 outcome observations) which indicated the variance within each set of responses (e.g., coworker desirability for the first applicant for the first position, likelihood of success for the final applicant for the second position, self-reported conscientiousness). These variances were then aggregated to reflect the mean variance across the 17 measures including both positively- and negatively-worded items. There are a number of subjective cutoffs used to screen out participants based on invariant response patterns. I chose a very conservative cutoff – a mean response variance equal to or less than 0.25. This resulted in the exclusion of 12 additional participants (2.3%). In total, 111 (18.8%) of the initial 610 participants' data were removed through the three data screens that were conducted. Thus, the final sample for analyses included 499 participants (61.7% Female).

### **Analyses**

A multilevel approach was appropriate for this study design due to the multiple measurements (i.e., applicant stereotypes and evaluative ratings) nested within individual participants (Hox, 1998). The rationale behind the use of MLM, as opposed to a single level approach (e.g., regression), is that it allows the variance within participants (i.e., differences among one participant's evaluations of the four target applicants) to be separated from the variance due to between-participant predictors (e.g., prototype ratings, conscientiousness; Goldstein, 2011; Singer & Willett, 2003).

I used a model building process that reflects the steps for testing mediation presented by Baron and Kenny (1986). The first set of models (Step 1) assessed the direct effects of applicant sex and weight and the position to which applicants were applying (i.e., position) on evaluative outcomes (Hypotheses 1a-1d). These models tested the first important proposition of the broader prototype matching modeling; that is, that target applicants differed in evaluative ratings as a function of sex and/or weight (i.e., discrimination). The second set of models (Step 2) assessed whether participants ascribed target applicants different warmth/competence stereotypes based on target sex/weight (and their interaction) (Hypotheses 2a and 2b). That is, stereotype ratings served as the dependent variable in this analysis. The third set of models (Step 3) assessed whether the stereotypes participants ascribed to target applicants, predicted evaluative outcomes (Hypotheses 3a-b). The fourth set of models (Step 4) included the direct



effects of both target sex and weight (and their interaction) and stereotypes to predict evaluative outcomes (Hypothesis 4a-b). The fifth set of models builds on the previous models by including participants' ratings of their prototypical job holder for the given job (i.e., sales or customer service position). Specifically, this tested whether stereotypes and prototypes interacted to predict evaluative outcomes (Hypothesis 5).

## CHAPTER 3: RESULTS

### **Preliminary Factor Structure Analyses**

I used confirmatory factor analysis (CFA) in order to assess the factor structure of the stereotype measure and outcome measures, separately. The final dataset included 499 unique participants who made eight sets of warmth and competence (i.e., stereotype) and evaluative outcome ratings. Because the eight sets of ratings within participants were not independent, I conducted separate CFAs on each set for both the stereotype and outcome measures. Reported are the mean (and range) of fit statistic values for each model.

**Warmth and competence.** The scale I used to measure warmth and competence has been used in previous SCM research (Cuddy et al., 2002). However, there is an absence of evidence regarding its internal structure. Thus, I compared the intended two factor model (four-item warmth factor and five-item competence factor) to a model in which all nine items served as indicators of a single factor (one factor model). The two factor model, CFI=.95 (.91-.97), RMSEA=.12 (.10-.15), SRMR=.04 (.03-.06), fit the data substantially better than the single factor model, CFI=.81 (.75-.90), RMSEA=.22 (.16-.25), SRMR=.08 (.06-.10), as indicated by higher CFI values and lower RMSEA and SRMR values. These findings suggest separate warmth and competence scores to be more appropriate than a single stereotype score for testing the hypotheses of this study.

**Coworker desirability and success outcomes.** The coworker desirability and likelihood of success items were created specifically for this study. As with the SCM measure, it was important to test the internal structure of these measures prior to use in hypothesis testing analyses. I followed the same process as for the warmth and competence items, this time examining whether the intended two-factor structure (a four-item coworker desirability factor and a four-item likelihood of success factor) fit the observed data better than a single-factor structure including all eight items. The two-factor model, CFI=.84 (.82-.88), RMSEA=.23 (.21-.25), SRMR=.11 (.09-.11), indicated slightly better fit than the single factor model, CFI=.82 (.77-.86), RMSEA=.24 (.22-.26), SRMR=.11 (.09-.12). However, the mean values associated with each of the reported fit indices for both the single-factor and two-factor models

indicated less than acceptable fit (see Hu & Bentler, 1999). Moreover, the factor loadings for items in the two-factor models indicated that the three reverse-coded items had substantially lower factor loadings than the positively-worded items. Consequently, I tested a three-factor model in which the two positively-worded coworker desirability items made up a “coworker desirability” factor, the two reverse-coded coworker desirability items made up a “coworker undesirability” factor, and the single reverse-coded likelihood of success item was removed, creating a 3-item likelihood of success factor. The three-factor model fit the observed data well, CFI=.99 (.99-1.00), RMSEA=.05 (.01-.08), SRMR=.01 (.01-.02). As a result of these findings, I chose to use three outcome measures: a 2-item coworker desirability measure, a 2-item coworker undesirability measure (with higher scores indicating less undesirability), and a 3-item success outcome measure.

### **Preliminary Analyses**

The means, standard deviations, and correlations among study variables are presented in Table 3. Means, standard deviations, and effect sizes (*ds*) associated with differences on warmth and competence scores are presented separately based on applicant weight and sex (see Table 4). These results show that participants stereotyped normal weight applicants as more competent ( $d=0.12$ ), but overweight applicants as warmer ( $d= -0.22$ ). Participants also stereotyped female applicants as both more competent ( $d= -0.20$ ) and warmer ( $d= -0.25$ ). Table 5 presents means, standard deviations, and *ds* for participants’ evaluations of applicants on the coworker desirability, coworker undesirability, and likelihood of success outcome measures, separately for the sales and service positions. There are a number of important findings from Table 5 worth noting. First, among evaluations of applicants to the sales position, overweight applicants were evaluated as both more desirable coworkers and less undesirable coworkers, while normal weight applicants were evaluated as more likely to succeed. For the customer service position, the effect sizes associated with each of the mean differences in evaluations for both the sales and customer service positions may be trivial from a practical standpoint, with  $ds \leq .06$  and confidence intervals that include zero. That said, overweight applicants were again evaluated as more desirable coworkers, while there was no difference in coworker undesirability ratings, and overweight applicants were evaluated as more

likely to succeed. Second, females were evaluated more positively than males across all three outcomes for both the sales and service position. Findings regarding the effects of applicant sex on evaluative outcomes suggest non-trivial differences, as the associated confidence intervals do not include zero. However, given existing theory regarding job “fit”, what stands out about these findings is the lack of variation in effect sizes as a function of the position (sales/customer service). That is, “fit” theory would suggest that males better “fit” the sales position, and thus, should receive more positive evaluations. However, these preliminary findings suggest a consistent pro-female bias across evaluative outcomes for both positions.

### **Multilevel Modeling Analyses**

At each step described below I began by examining the unconditional model (Model 1), in which no predictors were included. An advantage of using a multilevel framework is that it separates the variance due to differences in multiple measurements of an outcome taken from each participant (i.e., within-participant variance) from variance due to between participant differences (i.e., between-participant variance). The purpose for the unconditional model was to differentiate the proportion of variance due to differences between participants’ ratings of applicants and that due to differences within participants’ ratings of each of the different applicants. This was established through the intraclass correlation (ICC) for each unconditional model. Next, I added the main effects of the Level 2 (L2) control variables: participant self-reported conscientiousness, sex, and the four previous experience items (Model 2). Subsequent models added the substantive Level 1 (L1) predictors. At Step 1 and Step 2, this involved adding the main effects (Model 3) and, interactions (Model 4) for (1) applicant sex and (2) weight and (3) position to predict evaluative outcomes (Hypothesis 1a-d) and stereotypes (Hypothesis 2a-b), respectively. At Step 3, the main effects (Model 3) and, as appropriate, interactions (Model 4) for participants’ (1) warmth and (2) competence stereotypes were used to predict evaluative outcomes (Hypothesis 3a-b). At Step 4, applicant characteristics and participants’ stereotypes of applicants were included in the same model to predict evaluative outcomes (Hypothesis 4a-b). In the final set of models, participants’ prototypes, as well as the stereotype x prototype interactions were included as predictors of

evaluative outcomes (Hypothesis 5). Model fit statistics for each model at each step are presented in Table 6. Bonferroni corrections were made for interpreting the statistical significance of each individual predictor to account for experiment-wise error due to multiple outcomes examined at each step (critical  $p < .0063$ ).<sup>2</sup> Cohen's  $d$  effect size values (Cohen, 1988) are presented in-text for significant substantive predictors. These values were calculated from  $t$  values associated with individual predictors within the MLM results and represent the effect size, controlling for other predictors in the MLM model.

### **Step 1 (X → Y): Applicant Sex and Weight Predicting Outcomes**

A preliminary step in testing the prototype matching model is to establish that applicant characteristics significantly predict the examined outcomes. That is, without demonstrating significant relationships between applicant sex and weight (predictors) and ratings on outcome variables tests of subsequent assumptions and prototype matching would not be warranted. Therefore, I first examined whether applicant sex and weight predicted coworker desirability, coworker undesirability, and likelihood of success outcome ratings.

**Coworker desirability outcome.** The results for each model included in the model building process for each of the three outcomes examined at this step are presented in Table 7. The first outcome I examined was coworker desirability. Model 1a ( $\chi^2_{498}=11,802$ ) represents the unconditional model and suggests that 38.7% of the variance in the reported coworker desirability was due to between participant differences, indicating that the remaining 61.3% of the variance was due to differences within participants' evaluations of the applicants they rated. Model 2a ( $\chi^2_{487}=11,648$ ) added the main effects of the L2 control variables (conscientiousness, participant sex, and responses to the four previous experience questions). The addition of the L2 control variables in Model 2a resulted in significantly better fit than Model 1a, based on the deviance test ( $\Delta D$ ) of -2 log-likelihood values and reported as  $\chi^2$  values,  $\Delta D(11)=154, p < .05$ . In Model 2a, the main effects for three of the L2 control variables were significant predictors of coworker desirability (see Table 7). Participants who self-reported higher conscientiousness also rated applicants as more desirable coworkers, while both male participants and those with previous

experience in formal evaluation rated applicants as less desirable coworkers. The main effects of the L1 applicant characteristics (weight and sex) and position were added in Model 3a ( $\chi^2_{484}=11,630$ ), which significantly improved model fit,  $\Delta D(3)=18, p<.05$ . Participant conscientiousness, sex, and previous experience remained significant, and the relationships between these variables and coworker desirability remained in the same direction as in Model 2a (see Table 7). Of the three L1 predictors, only applicant sex significantly predicted evaluators' coworker desirability ratings, suggesting that female applicants were rated as more desirable coworkers ( $d = -0.17$ ). The non-significant coefficients associated with the other two L1 predictors suggest that evaluators did not discriminate against applicants based on weight, and did not differ in their ratings based on the position to which applicants were applying. Model 4a ( $\chi^2_{481}=11,638$ ) resulted in worse fit than Model 3a,  $\Delta D(3)=-8$ . In addition, none of the interactions included in this model were statistically significant (see Table 7).

**Coworker undesirability outcome.** The unconditional model (Model 1b,  $\chi^2_{498}=11,542$ ) indicated that 41.6% of the variance in participants' coworker undesirability ratings was due to between-participant differences, while the remaining 58.4% of the variance was due to differences within-participants. Model 2b ( $\chi^2_{487}=11,372$ ) added the main effects of the L2 control variables, and significantly improved model fit,  $\Delta D(11)=170, p<.05$ . Those self-reporting greater conscientiousness rated applicants as less undesirable coworkers, while male participants and participants with previous experience in formal evaluation rated applicants as more undesirable coworkers (see Table 7). Model 3b ( $\chi^2_{484}=11,366$ ) introduced the main effects of the three L1 predictors, but did not significantly improve model fit,  $\Delta D(3)=6, p>.05$ . In Model 3b, participant conscientiousness, sex, and previous evaluation experience remained significant, and only the main effect of participant sex was significant among the newly added L1 predictors (see Table 7). The significant effect of applicant sex indicates that females were rated as less undesirable coworkers than were males ( $d = -0.15$ ). Model 4b ( $\chi^2_{481}=11,374$ ) add the L1 interaction terms, but resulted in worse fit than Model 2b,  $\Delta D(9)=-2$ . In addition, none of interactions were statistically significant.

**Likelihood of success outcome.** The unconditional model (Model 1c,  $\chi^2_{498}=12,514$ ) suggests that 31.3% of the variance in likelihood of success ratings was associated with between-participant differences, while the remaining 68.7% of the variance in the data was due to differences within-participants. Model 2c ( $\chi^2_{487}=12,382$ ) included the main effects of the L2 control variables and resulted in a significantly better fit to the data than Model 1c,  $\Delta D(11)=132, p<.01$ . Conscientiousness scores were found to be positively related to likelihood of success ratings, and male participants rated applicants as less likely to succeed. Model 3c ( $\chi^2_{484}=12,354$ ) added the main effects of the three L2 predictors, significantly improving model fit,  $\Delta D(3)=28, p<.05$ . Both participant conscientiousness and sex remained statistically significant (see Table 7). With regard to L1 predictors, female applicants were rated as more likely to succeed ( $d = -0.17$ ), as were applicants applying for the sales position ( $d = -0.12$ ), while applicant weight did not significantly affect participants' ratings of likelihood of success. Model 4c ( $\chi^2_{481}=12,364$ ) added the L1 interactions, but resulted in a slight decrease in fit,  $\Delta D(3)= -10$ . In addition, only participant conscientiousness and applicant sex remained significant predictors of likelihood of success ratings in Model 4c (see Table 7).

**Step 1 summary.** With regard to the main effects of L2 control variables a consistent positive relationship was found between participants' conscientiousness and each of the three outcomes. That is, participants self-reporting higher conscientiousness also viewed applicants as more desirable coworkers, less undesirable coworkers, and more likely to succeed. Across all three outcomes there was also a consistent trend for male participants to give applicants lower ratings. Finally, participants with previous formal experience in evaluation rated participants lower on both coworker desirability and undesirability.

The inclusion of the three L1 predictors (applicant sex and weight and position) significantly improved model fit, over models including only the main effects for the six control variables, in predicting coworker desirability and likelihood of success ratings. In addition, applicant sex significantly predicted each of the three outcomes. This provides partial support for Hypothesis 1a. That is, these results suggest that sex-based discrimination was present based on evaluators' ratings. However, the effect was in the opposite direction than was predicted, in that female applicants received higher

evaluations than did male applicants. Partial support was also found for Hypothesis 1c, in that females received higher ratings than with regard to the customer service position, but females also received higher ratings than males with regard to the sales position.

I failed to find support for Hypothesis 1b. That is, applicant weight did not have a significant effect on evaluators' ratings with regard to any of the three outcomes, indicating that evaluators did not discriminate due to this applicant characteristic. In addition, the interaction between applicant weight and applicant sex was not a significant predictor of any of the three outcomes, and I failed to support Hypothesis 1d. In general, these findings reflect the effect sizes based on mean differences that were presented above in Table 5.

Despite my failure to find support for the initial hypotheses, a number of relationships warrant further consideration. First, although the effects of applicant sex were not in the hypothesized direction, the effects for applicant sex suggest that presence of sex-based discrimination. This suggests that further testing is warranted to examine whether stereotypes mediate the main effect of applicant sex on each of the evaluative outcomes. Second, applicant weight did not significantly affect evaluators' ratings on any of the three outcomes, which suggests that further mediational testing regarding weight-based discrimination is not warranted. However, it may be informative to examine whether stereotypes differ as a function of applicant weight. Thus, applicant weight was included in Step 2 analyses but not in analyses at subsequent steps.

### **Step 2 (X → M): Applicant Sex and Weight Predicting Stereotype Content (Warmth and Competence)**

As a second step in Baron and Kenny's (1986) approach to testing mediation, I tested whether applicant characteristics (i.e., sex and weight) significantly predicted the proposed mediating variables in the prototype matching model (i.e., warmth and competence stereotype ratings). A failure to find support for such relationships would suggest that the proposed mediators do not mediate the relationship between applicant characteristics and the examined outcomes, and further subsequent analyses regarding the prototype matching model would become unwarranted.



**Warmth.** Next, I examined the extent to which applicant sex, weight and position predicted participants' stereotypes of applicants (warmth and competence). The unconditional model (Model 1d,  $\chi^2_{498}=10,486$ ) for warmth indicated that 43.4% of the variance in participants' ratings of warmth was due to between-participant differences, and 56.6% was due to within-participant differences. Model 2d ( $\chi^2_{487}=10,374$ ) added the L2 controls, resulting in a significant improvement in model fit,  $\Delta D(11)=112$ ,  $p<.05$ . Participant conscientiousness and sex significantly predicted warmth ratings, with those higher on conscientiousness rating applicants as warmer, and males rating applicants as less warm (see Table 8). Model 3d added the main effects of the L1 predictors of applicant sex and weight and position ( $\chi^2_{484}=10,156$ ). Model 3d fit the data significantly better than Model 2d,  $\Delta D(3)=218$ ,  $p<.05$ . Both participant conscientiousness and sex remained significant, and all three L1 predictors significantly predicted warmth (see Table 8). Overweight applicants were rated as warmer ( $d=0.23$ ), as were female applicants ( $d=-0.25$ ) and applicants applying to the customer service position ( $d=0.36$ ). Model 4d added the interactions between the three L1 predictors, including their three-way interaction ( $\chi^2_{480}=10,168$ ). The addition of the interaction terms resulted in a decrease in model fit,  $\Delta D(4)=-10$ . In addition, each of the main effects remained significant, but none of the interactions significantly predicted evaluators' warmth ratings.

**Competence.** Model 1e, the unconditional model ( $\chi^2_{498}=9,920$ ), indicated that 46.1% of the variance in competence ratings was associated with between-participant differences, and the remaining 53.9% of the variance was due to within-participant differences. Model 2e ( $\chi^2_{487}=9,802$ ) included the L2 control variables and provided a significant improvement in fit compared to the unconditional model,  $\Delta D(11)=118$ ,  $p<.05$ . Participant sex and conscientiousness were significant predictors of competence stereotype ratings (see Table 8), as a positive relationship was found between participants' conscientiousness and ratings of applicant competence, and male participants' ratings of applicants' competence were significantly lower female participants' ratings. Model 3e added the L1 main effects ( $\chi^2_{484}=9,428$ ), which significantly improved model fit,  $\Delta D(3)=374$ ,  $p<.05$ . Significant relationships remained for participant sex and conscientiousness. Each of the newly added L1 variables also

significantly predicted ratings of applicant competence (see Table 8). Overweight and male applicants were viewed as less competent ( $d = 0.14$  and  $-0.22$ , respectively), and lower ratings of competence were also reported across applicants when evaluated for the customer service position ( $d = -0.59$ ). Model 4e added the L1 interaction terms ( $\chi^2_{480} = 9,440$ ), but resulted in a decrease in model fit,  $\Delta D(4) = -12$ . In Model 4e, none of the L1 interactions significantly predicted competence ratings, and both applicant weight and sex became non-significant.

**Step 2 summary.** As was the case at Step 1, participant conscientiousness and participant sex significantly predicted outcomes at Step 2. Participants who viewed themselves as more conscientious also viewed applicants as both warmer and more competent, while male participants viewed applicants as less warm and less competent.

In comparison to predicting evaluative outcomes in Step 1, the L1 predictors of applicant sex, applicant weight and position were much stronger predictors of evaluators' warmth and competence stereotype ratings, and each was significantly related to both warmth and competence ratings. In general, these findings support research suggesting different stereotypes are associated with different groups. Specifically, females were rated as both warmer and more competent than males. This provides partial support for Hypothesis 2a, which predicted that females would be rated as warmer, but males as more competent. Partial support was also found for Hypothesis 2b, which predicted that normal weight applicants would be rated as both warmer and more competent than overweight applicants. Normal weight applicants were rated as more competent, but not as warm as overweight applicants. Applicants to the customer service position were rated as warmer, and applicants to the sales position were rated as more competent. Although no specific hypotheses were made with regard to these relationships, this does provide evidence that the manipulation of resume content was effective, as sales resumes were intended to reflect greater competence, and customer service resumes greater warmth.

### **Step 3 (M → Y): Stereotype Content (Warmth and Competence) Predicting Outcomes**

The third assumption of Baron and Kenny's (1986) approach to mediation is that proposed mediating variables significantly predict the outcomes examined. As was the case for Step 1 and 2

analyses, a failure to find support for proposed mediators (i.e., warmth and competence stereotype ratings) as significant predictors of outcomes would render further mediational analyses unwarranted. Thus, as a second step I tested the assumption that warmth and competence stereotype ratings would significantly predict coworker desirability, coworker undesirability, and likelihood of success outcome ratings.

In evaluating the effects of stereotype content on the three work-related outcomes, the unconditional models (Model 1) and models which added the L2 control variables (Model 2) in the Step 3 are identical to Models 1a-c and Models 2a-c, respectively, in Step 1. The difference in the model building process between Step 1 and Step 3 is that Model 3a-c (Step 1) introduced the main effects of applicant characteristics (sex and weight), whereas Model 3f-h (Step 3) introduced the main effects of stereotype content (warmth and competence). As a result, I describe the model building process for each outcome beginning with the results of Model 3f-h.

**Coworker desirability outcome.** I examined the main effects for L1 stereotypes (warmth and competence) in predicting coworker desirability in Model 3f ( $\chi^2_{485}=10,046$ ), which marked a significant improvement in fit to the data in comparison to Model 2a,  $\Delta D(2)=1,602$ ,  $p<.05$ . Among the six control variables, only previous experience in formal evaluation was significant, in that those with previous experience rated applicants as significantly less desirable coworkers. Among the L1 predictor main effects, both warmth ( $d =0.93$ ) and competence significantly ( $d =0.55$ ) predicted coworker desirability (see Table 9). Model 4f ( $\chi^2_{484}=9992$ ) added the interaction between warmth and competence, which further improved model fit,  $\Delta D(1)=54$ ,  $p<.05$ . The main effects for previous experience in formal evaluation, warmth ( $d =0.97$ ), and competence ( $d =0.60$ ) remained significant, and the interaction between warmth and competence was also significant ( $d =0.25$ ; see Table 9). The significant interaction is depicted in Figure 2, and shows that applicants stereotyped as warmer received higher ratings on evaluative outcomes when they were also stereotyped as more competent.

**Coworker undesirability outcome.** The addition of warmth and competence in Model 3g ( $\chi^2_{485}=9942$ ) to the control variables included in Model 2b resulted in significantly better fit,  $\Delta D(2)=1430$ ,

$p < .05$ . Among control variables, more conscientious participants rated applicants as less undesirable coworkers, while males and those with previous experience in formal evaluation and the healthcare industry rated applicants as less undesirable coworkers. Among L1 predictors, applicants rated as warmer ( $d = 0.72$ ) and more competent ( $d = 0.69$ ) were also rated as less undesirable coworkers (see Table 9). Model 4g added the interaction between warmth and competence ( $\chi^2_{484} = 9924$ ), which significantly improved model fit,  $\Delta D(1) = 18$ ,  $p < .05$ . There were no changes from Model 3g to Model 4g in the significance or direction of main effects (see Table 9). Specifically, both warmth ( $d = 0.73$ ) and competence ( $d = 0.71$ ) remained significant predictors, and a significant effect was found for the interaction between warmth and competence ( $d = 0.16$ ). This interaction is shown in Figure 3, and suggests that applicants rated as warmer were rated as less undesirable coworkers when they were also rated as more competent, in comparison to those rated as less competent.

**Likelihood of success outcome.** Model 3h ( $\chi^2_{485} = 10,114$ ) added the main effects of warmth and competence to the L2 control variables included in Model 2c, resulting in significantly better model fit,  $\Delta D(2) = 2268$ ,  $p < .05$ . Participant conscientiousness negatively related to ratings of applicants' likelihood of success, while the main effects for both warmth ( $d = 0.60$ ) and competence ( $d = 1.18$ ) were positively related to ratings likelihood of success (see Table 9). The interaction between warmth and competence was added in Model 4h ( $\chi^2_{484} = 10,120$ ), which represented a decrease in model fit,  $\Delta D(1) = -4$ . Each of the three main effects remained significant (see Table 9), and a non-significant effect was found for the warmth x competence interaction.

**Step 3 summary.** Warmth and competence stereotypes were strong predictors of all three outcomes. Additionally, the coefficients reported in Table 9 show that warmth stereotypes were stronger predictors of coworker desirability ratings, and competence stereotypes were stronger predictors of likelihood of success ratings. These findings support Hypotheses 3a and 3b. Finally, in the cases of the coworker desirability and undesirability outcomes the interaction between warmth and competence was also significant, indicating that applicants rated as high on both warm and competent were rated as the

most desirable and least undesirable coworkers. There were a number of significant main effects for control variables. However, the significance of these effects was less consistent than at Step 1 and Step 2.

**Step 4 (X, M → Y): Applicant Sex and Weight Predicting Outcomes (Controlling for Stereotype Content: Warmth and Competence)**

As a final step in Baron and Kenny's (1986) approach to mediation, a single model was examined with regard to each of the three evaluative outcomes. The model included the predictors present in the final model at Step 3 and the L1 predictors found to be significant at Step 1. Applicant sex was a significant predictor of each of the outcomes at Step 1 and included in each of the Step 4 final mediation models. Position was a significant predictor of likelihood of success at Step 1 and included in the Step 4 model predicting that outcome. Applicant weight was not a significant predictor of any of the three outcomes at Step 1, and not included in the Step 4 models. Mediation is indicated in instances where predictors retained from Step 1 become non-significant when they are included with the proposed mediating variables (L1 predictors from Step 3) in the Step 4 models. This would indicate full mediation. Results of the models predicting each of the evaluative outcomes are presented in Table 10.

**Coworker desirability.** In the Step 4 model predicting coworker desirability ( $\chi^2_{483}=9998$ ), the main effects for previous experience in formal evaluation, warmth ( $d=0.96$ ), competence ( $d=0.60$ ), and the interaction between warmth and competence ( $d=0.25$ ) significantly predicted coworker desirability (see Table 10). Applicant sex, which had been a significant predictor of coworker desirability at Step 1, became non-significant when warmth, competence, and their interaction were included at Step 4 ( $d = -0.02$ ). The addition of applicant sex in the Step 4 model resulted in worse model fit,  $\Delta D(1) = -6$ .

**Coworker undesirability.** There were a number of significant predictors of coworker undesirability in the Step 4 model ( $\chi^2_{483}=9930$ ), including the main effects for warmth ( $d=0.73$ ) and competence ( $d=0.71$ ) and their interaction ( $d=0.16$ ; see Table 10). Applicant sex, which was a significant predictor of coworker undesirability at Step 1, was not a significant predictor at Step 4 ( $d = 0.00$ ). Similar to the Step 4 model predicting coworker desirability, the addition of applicant sex in predicting coworker undesirability resulted in worse fit,  $\Delta D(1) = -6$ .

**Likelihood of success.** The Step 4 model predicting likelihood of success ( $\chi^2_{483}=10,118$ ) resulted in significant main effects for warmth ( $d=0.54$ ) and competence ( $d=1.12$ ), and a non-significant effect for applicant sex ( $d=-0.01$ ), indicating full mediation. Once again, the addition of applicant sex in predicting likelihood of success ratings resulted in a decrease in model fit,  $\Delta D(2)=-4$ .

**Step 4 summary.** Applicant sex was a significant predictor of each of the three outcomes at Step 1. The inclusion of participants' warmth and competence stereotype ratings of applicants along with applicant sex at Step 4 resulted in a non-significant effect for applicant sex. According to Baron and Kenny's (1986) approach to mediation these results suggest that participants' stereotypes of applicants fully mediated the relationship between applicant sex and participants' evaluations of applicants on all three of the evaluative outcomes. These findings support Hypothesis 4a. Because applicant weight was not found to be a significant predictor of any of the three outcomes at Step 1 it was not included in the Step 4 models, and a test of mediation was not conducted for weight (Hypothesis 4b).

### **Prototype Matching**

What the above series of analyses has not addressed is whether or not participants' prototype ratings moderate the relationship between their stereotypes and outcome ratings. To examine this relationship, I conducted an additional set of multi-level analyses. With the exception of position, the main effects included in the Step 4 model predicting each outcome were retained for prototype matching analyses. Participants indicated separate warmth and competence prototype ratings for the sales and customer service position. Thus, instead of including position as a predictor, separate analyses were conducted using each of the two sets of prototypes with regard to each of the three outcomes. In addition to the predictors retained from the Step 4 models, the main effects for prototype warmth and competence ratings and the stereotype-prototype warmth and stereotype-prototype competence interactions were included, in order to directly test whether stereotype-prototype match affected participants' ratings on evaluative outcomes.

**Sales position (coworker desirability).** The model testing stereotype-prototype match predicting coworker desirability in the sales position ( $\chi^2_{480}=5070$ ) resulted in significant effects for both

warmth and competence stereotype ratings (see Table 11a). Effect sizes associated with these effects were  $d=0.81$  and  $0.51$ . The only other significant predictor in this model was the interaction between stereotype and prototype competence ratings ( $d=0.19$ ; Figure 4a). The Figure suggests that participants with higher prototypical competence standards for the sales position better differentiated between the low and high stereotype competence applicants on coworker desirability ratings, with those high on stereotype competence (i.e., better matching competence prototypes) receiving the most positive coworker desirability ratings. The stereotype-prototype warmth interaction ( $p=.02$ ;  $d=0.10$ ) was not significant as a result of the Bonferroni correction used throughout these analyses. Nonetheless, this interaction is depicted in Figure 4b, and suggests a similar trend for stereotype-prototype matching.

**Sales position (coworker undesirability).** Similar to the coworker desirability model, both stereotype main effects and the interaction between stereotype and prototype competence ratings were the only significant predictors in the model predicting coworker undesirability ( $\chi^2_{480}=5042$ ). The coefficients associated with each predictor included in the model predicting coworker undesirability ratings are presented in Table 11a. Stereotype warmth ( $d=0.48$ ) and competence ( $d=0.75$ ) were again strong predictors of coworker undesirability ratings. Figure 5a depicts the stereotype-prototype competence interaction ( $d=0.17$ ), and suggests that those with higher prototypical competence standards better differentiated between applicants stereotyped as low and high on competence. The stereotype-prototype warmth interaction ( $p=.02$ ;  $d=0.11$ ) was again not significant based on the corrected  $p$  value. However, this interaction is also depicted (Figure 5b), and although the effect was not significant, it does suggest that some degree of stereotype-prototype matching may be present.

**Sales position (likelihood of success).** Neither of the stereotype-prototype interactions were statistically significant in the model predicting likelihood of success in the sales position ( $\chi^2_{480}=4888$ ). The main effects for warmth ( $d=0.34$ ) and competence stereotypes ( $d=1.40$ ) were positively, and competence prototypes ( $d= -0.16$ ) negatively, related to likelihood of success ratings in the sales position model (see Table 11a). This suggests that prototype matching did not affect likelihood of success ratings.

**Customer service position (coworker desirability).** The same predictors included in the models predicting each of the three outcomes for the sales position were included in the models predicting outcomes for the customer service position. The coworker desirability ( $\chi^2_{480}=5164$ ) resulted in a number of significant predictors (see Table 11b). Applicants stereotyped as warmer ( $d=0.83$ ) and more competent ( $d=0.56$ ) were viewed as more desirable coworkers, while participants who held higher competence prototypes rated applicants as less desirable coworkers ( $d= -0.23$ ). Also significant were both stereotype-prototype interactions (Figure 6a and 6b). Figure 6a depicts the stereotype-prototype competence interaction, which represents a moderating effect opposite to that which was expected – that is, those with lower prototypical competence standards actually better differentiated applicants based on stereotype competence ratings ( $d= -0.14$ ). Figure 6b depicts the stereotype-prototype warmth interaction, and suggests that participants with high prototypical ratings better differentiated applicants based on stereotype competence ratings ( $d=0.30$ ).

**Customer service position (coworker undesirability).** The coworker undesirability model ( $\chi^2_{480}=5062$ ) resulted in the main effects for both stereotype and prototype warmth and competence significantly predicting coworker desirability ratings. Higher stereotype warmth ( $d=0.71$ ) and competence ( $d=0.62$ ) and prototype warmth ratings ( $d=0.19$ ) were associated with ratings of less coworker undesirability, while higher prototypical competence ratings ( $d= -0.24$ ) were associated with ratings of greater coworker undesirability (Table 11b). In addition, a significant effect was found for the stereotype-prototype warmth interaction (Figure 7), which suggests that participants with higher prototypical warmth standards better differentiated between low and high warmth stereotyped applicants ( $d= -0.08$ ).

**Customer service position (likelihood of success).** The main effects for stereotype warmth ( $d=0.71$ ) and competence ( $d=0.93$ ) and prototype competence ( $d= -0.20$ ) were significant predictors in the model predicting likelihood of success ( $\chi^2_{480}=5176$ ). The coefficients for each predictor included in the model are presented in Table 11b. In addition, there was a significant stereotype-prototype warmth



interaction ( $d=0.12$ ). As shown in Figure 8, participants with higher prototypical warmth standards better differentiated between applicants who were stereotyped as low and high on warmth.

**Prototype matching summary.** Results of the models including prototype ratings showed a number of significant stereotype-prototype interactions, providing evidence in support of prototype matching in predicting evaluative outcomes (Hypothesis 5). More specifically, competence prototypes tended to play a stronger moderating role between stereotypes and evaluative outcomes for the sales position than did warmth prototypes, and vice versa for the customer service position. This is an important finding given that pilot testing was conducted to ensure the job description and resume materials used with the sales position conveyed competence, and the customer service position warmth. The implications of these findings are discussed in greater detail below.

## CHAPTER 4: DISCUSSION

This study is the first to test the propositions of the prototype matching model. This model builds on existing “fit” theory (e.g., Heilman, 2001) by considering the role of evaluators’ unique prototypes in determining person-job fit. I proposed the prototype matching model as an explanation of the underlying mechanism of discrimination, one that could generalize across discrimination type. To test the propositions of this model I examined the role of two applicant characteristics (sex and weight) and two stereotype/prototype factors (warmth and competence) as predictors of three evaluative outcomes (coworker desirability, coworker undesirability, and likelihood of success) for two separate jobs (sales and service). The key findings regarding each proposition are discussed below.

### **Discrimination as a Function of Applicant Sex and Weight (Step 1)**

As a model proposed to explain the underlying mechanism of discrimination, the most basic proposition of the prototype matching model is the presence of discrimination. I examined two applicant characteristics that have been shown to be a source of discrimination, sex and weight. A significant effect was found for applicant sex, with females being evaluated more positively than male applicants across each of the three outcomes under examination. Although the effect of sex was in the opposite direction than hypothesized (Hypothesis 1a), this finding demonstrated the presence of discrimination (the basic proposition of the prototype matching model). The directionality of the effect of applicant sex does have precedent in the research literature, which has often found inconsistent effects for sex-based discrimination and even an overall pro-female bias in some selection-based employment contexts (e.g., assessment center simulation exercises; Dean et al., 2008).

A non-significant effect was found for applicant weight with regard to each of the three outcomes included in this study (Hypothesis 1b). The practical significance of the effects of applicant weight was also small, much smaller than what has been found in meta-analyses of lab-based research on employment-based weight discrimination ( $d=.36$ , Roehling et al., 2013;  $d=.52$ , Rudolph et al., 2009). However, a number of lab-based studies of weight discrimination have shown small effects for weight

(Bellizzi & Hasty, 2001; Bevins, 2003; Rothblum, Miller, & Garbutt, 1988), or even effects favoring overweight stimuli (Hebl, 1997; Polinko & Popovich, 2001).

One possible explanation for my failure to find an effect for applicant weight may be due to the way I operationalized the normal weight and overweight conditions. The operationalization of some discriminable characteristics is a fairly straightforward task. Using the example of applicant sex, it is often quite easy to distinguish between male and female stimuli. Operationalizing weight differences, however, is not as straightforward, and the way in which weight differences have been operationalized in the weight discrimination literature has varied considerably. For example, Brink (1988) contrasted written descriptions of a 165-pound (normal weight) and 425-pound (overweight) job candidate to assess the effects of weight discrimination. Conversely, Rothblum et al. (1988) contrasted photographs of females perceived by pilot testers to weight approximately 125 pounds (normal weight stimulus) and 155 pounds (overweight stimulus). Perhaps not coincidentally, the magnitude of the effect Brink (1988) found was approximately one-half a standard deviation greater than that found by Rothblum et al. (1988). The operationalization of applicant weight in the present study is more similar to the approach taken by Rothblum et al. (1988). This may have led to similarly weaker effects for weight discrimination. Another possible explanation for my failure to find weight discrimination is that I controlled for applicant attractiveness. That is, I pilot tested photos on stimulus attractiveness in order to minimize the effects due to this potential confound. However, weight and attractiveness have been shown to be negatively related (e.g., Singh & Young, 1995) and controlling for the effects of applicant attractiveness may have further reduced the differences in ratings that were observed between normal weight and overweight applicants.

A number of my initial hypotheses were not supported – that is, regarding the main effect for applicant weight (Hypothesis 1b), and the interaction between applicant sex and position (Hypothesis 1c) and applicant weight (Hypothesis 1d). However, the data did show the presence of discrimination due to applicant sex (Hypothesis 1a), indicating that further analyses were warranted with regard to this characteristic. Because I failed to find a significant main effect for applicant weight at Step 1, I did not include it as a predictor in the final step of the mediational analyses, nor in the subsequent analyses

regarding prototype matching. However, I did explore the effects of applicant weight, along with applicant sex, on participants' stereotypes of applicants at Step 2.

### **Stereotype Differences as a Function of Applicant Sex and Weight (Step 2)**

The second proposition of the prototype matching model is that the stereotypes evaluators attribute to applicants differ as a function of applicant characteristics, such as sex and weight. I found that both applicant sex and weight significantly affected evaluators' warmth and competence stereotypes of applicants'. However, the direction of some of these effects was unexpected. With regard to applicant sex, I hypothesized that females would be rated as warmer than males, and males as more competent than females (Hypothesis 2a). However, female applicants were rated as both warmer and more competent than male applicants. With regard to applicant weight, I expected normal weight applicants to be rated as both warmer and more competent than overweight applicants (Hypothesis 2b). Normal weight applicants were rated as more competent than overweight applicants. However, overweight applicants were rated as warmer than normal weight applicants. In terms of the directionality of results, these findings provide only mixed support for hypotheses regarding the specific stereotypes evaluators attribute to individuals based on sex and weight. From a broader standpoint, however, these findings do demonstrate significantly different stereotypes due to applicant sex and weight, which provides support for the second proposition of the prototype-matching model – that stereotypes differ as a function of applicant characteristics.

### **Discrimination as a function of stereotype differences (Step 3)**

The third proposition of the prototype matching model is that stereotypes predict evaluative outcomes. Stated differently, this proposition suggests that evaluators discriminate against applicants based on the stereotypes they attribute to them. Stereotypes strongly predicted each of the three outcomes used in this study. Specifically, a positive relationship was found between both warmth and competence stereotypes and each of the three outcomes. In addition, warmth stereotypes were a stronger predictor of coworker desirability ratings than were competence stereotypes (Hypothesis 3a), and competence stereotypes were a stronger predictor of likelihood of success than were warmth stereotypes (Hypothesis

3b). These findings suggest that evaluators gave greater weight to their warmth stereotypes of each applicant when determining the extent to which they would like to work alongside him or her, and gave greater weight to their competence stereotypes of each applicant when determining the extent to which they perceived the likelihood of him or her succeeding. This demonstrates that these two stereotype factors have somewhat different correlates. From a broader standpoint, these findings also support the third proposition of the prototype matching model.

#### **Stereotypes Mediating Discrimination Based on Applicant Sex (Step 4)**

The fourth proposition of the prototype matching model is that stereotypes mediate the relationship between applicant characteristics (e.g., applicant sex) and evaluative outcomes resulting in discrimination. Because no effect was found for applicant weight at Step 1, this analysis was restricted to examining whether stereotypes mediated the relationship between applicant sex and each of the three outcomes. Full mediation was demonstrated with regard to each of the three outcomes. For coworker desirability and coworker undesirability, mediation was demonstrated through a non-significant effect for applicant sex and significant main effects for warmth and competence stereotype ratings and their interaction. The interaction suggests that warmth ratings had a stronger effect on coworker desirability and coworker undesirability ratings when applicants were also rated as high on competence, with applicants stereotyped as both warm and competent being rated as the most desirable and least undesirable coworkers. For likelihood of success, mediation was demonstrated through a non-significant effect for applicant sex and significant main effects for both stereotype warmth and competence ratings. In addition, warmth stereotypes remained the stronger predictor of coworker desirability ratings, and competence stereotypes remained the stronger predictor of likelihood of success. Taken together, the results of Steps 1 through 4 demonstrate that stereotypes mediate the relationship between applicant sex and evaluative outcomes (Hypothesis 4a), and provide the rationale for testing the final proposition of the prototype matching model – the role of evaluators’ prototypes as a moderator of the relationship between applicant stereotypes and evaluative outcomes.

## **Prototype Matching**

Participants provided separate prototypical warmth and competence ratings for the sales and customer service positions. Thus, I conducted separate analyses on these data in order to compare evaluators' sales (customer service) prototypes to their stereotypes of each applicant applying to the sales (customer service) position in predicting evaluative outcomes. With regard to the sales position, I found evidence that evaluators used prototype matching in evaluating applicants through significant competence stereotype-prototype interactions for the coworker desirability and coworker undesirability outcomes. Specifically, evaluators who had high prototypical competence standards for the sales position better differentiated between applicants who were stereotyped as low and high on competence, with the most positive coworker ratings being given by evaluators with high prototypical competence standards to applicants who were stereotyped as being highly competent. Marginally significant effects were also found for warmth stereotype-prototype interactions.

I found additional support for evaluators' use of prototype matching when I examined the customer service position. Here, I found significant warmth stereotype-prototype interactions for all three of the outcomes. These interactions indicated that evaluators with high prototypical warmth standards better differentiated between applicants stereotyped as more and less warm. In addition, these interactions showed the most pronounced effects in predicting coworker desirability and likelihood of success ratings. That is, applicants stereotyped as less warm were given lower ratings by evaluators who had high prototypical warmth standards than by evaluators who had low prototypical warmth standards, but applicants stereotyped as highly warm were given higher ratings by evaluators with high warmth prototypes than by evaluators with low warmth prototypes. I also found a significant effect for the competence stereotype-prototype interaction in predicting coworker desirability. However, it was evaluators with low prototypical competence standards who better differentiated between applicants stereotyped as low and high on competence in their coworker desirability ratings. Despite this one unexpected finding, each of the other interactions supports the expected effects of prototype matching (Hypothesis 5). I did not present specific hypotheses regarding which stereotype-prototype interaction

(warmth or competence matching) would have the strongest effect for each position. However, given that pilot testing was done to ensure that sales job descriptions and applicant resumes more strongly reflected competence, and customer service job descriptions and applicant resumes more strongly reflected warmth, this is a particularly important finding. That is, it appears that participants attended to the competence-relevant (warmth-relevant) information in the sales (customer service): (a) job description in reporting their prototypes for each job, and (b) resume information when reporting their stereotypes for applicants. These findings also suggest that participants weighed the relevance of information (competence or warmth stereotype-prototype match) when evaluating applicants with respect to each position.

Finally, it is important to note the effect sizes associated with these interactions. Statistically significant interactions were typically associated with what Cohen (1988) describes as small effects. However, this is not to say these effects are not important. At the outset of the present study I proposed the prototype matching model to build upon existing theory to more precisely capture the process of discrimination in work-related evaluations. Thus, the goal of this study was to present a model that accounted for additional individual-level variance that has not been accounted for by existing models. Consequently, the variance associated with prototype matching was largely expected to be less than that associated with the well-supported effects of stereotypes and other broadly studied variables.

### **Limitations**

There are a number of limitations associated with this study. First, the MTurk population may not be representative of the broader workforce population, and is likely not representative of the population of organizational decision makers. As mentioned above, research has begun to explore the population characteristics of MTurk participants, but it is still unclear how well this data generalizes to actual organizational decision making situations. In addition, the lab-based design and study materials (i.e., job descriptions and resumes) likely do not reflect the real world contexts and quantity of information from which actual employment decisions are made. On the one hand, the low-consequence decisions associated with the context of this study may have resulted in participants putting less effort into the decision making process than what would be expected from real world decision makers. On the

other hand, the quantity of information, in terms of (a) the number of job candidates evaluators had to evaluate and (b) the amount resume and other job-relevant information on each job candidate, study participants were faced with processing was also likely less than evaluators typically must process within in real selection contexts. Taken together, both the quality and quantity of cognitive processing completed by study participants may have differed from that of real world decision makers.

Nonetheless, this study aimed to establish the internal validity of this model prior to assessing external validity (Mook, 1983). Thus, the greatest potential limitations of this study were related to factors threatening internal validity. The fact that participation was completed online and not proctored represents possible threats to the internal validity of these findings. Participants' level of motivation to complete this study and the extent to which participants' external motivation (small compensation) may have affected the accuracy of their responses are unclear. Considerable efforts were made, through data cleaning procedures, to minimize these threats. However, these procedures may not have fully accounted for all sources of measurement error due to this unique population.

Evidence presented here provides initial support for the prototype matching model, specifically with regard to the characteristic of applicant sex. My failure to find the presence of weight discrimination in this data may raise questions regarding the generalizability of the prototype matching model. However, it is important to point out that this failure resulted in my inability to *test* the prototype matching model with regard to weight; this is opposed to failing to explain weight discrimination through the prototype matching model. As such, the flexibility in the prototype matching model to explain discrimination other than that due to applicant sex remains untested, and future research should aim to extend these findings to other discriminable characteristics.

### **Future Research Directions**

In addition to extending the prototype matching model to other discriminable characteristics, there are a number of factors that may also affect discrimination. For example, future research may explore whether evaluator-applicant similarity on discriminable characteristics affects the stereotypes and prototypes evaluators use when evaluating applicants. Research suggests that rater x ratee interactions



can affect the magnitude of discrimination (Bowen et al., 2000), and these interactions may also be relevant to predicting evaluators' stereotypes and prototypes. In applying the prototype matching model, it is also important that future research explore the role of contextual factors such as how prototypes differ between jobs other than sales and customer service, and how the prototype matching model generalizes to work-related settings beyond that of selection.

Importantly, this study showed that evaluators' prototypes of successful jobholders differ. That is, a number of significant main effects were found for participant prototypes, suggesting that evaluators' prototypes, at least to some extent, represent individual differences. An important step for future research is to better understand the factors and lead to these differences. This includes exploring whether evaluators have well-established prototypes for job holders of specific positions or whether these are created once evaluators are faced with a specific evaluative task. I would expect that real world evaluators have better-established prototypes for specific positions under their charge than did the MTurk participants for the sales and customer service positions in this study. However, the quantity and quality of evaluators' relevant experiences likely affect the precision and stability of real world evaluators' prototypes.

In conceptualizing the prototype matching model, I chose to use the term prototype in referring the mental schemas evaluators possess and compare to stereotypes. However, cognitive psychology continues to debate the importance of prototypes v. exemplars in creating these mental schemas (e.g., Nosofsky & Zaki, 2002; Smith & Minda, 1998). Multiple exemplar-based models have been developed (e.g., Medin & Schaffer, 1978; Brooks, 1978), but each generally, when applied to the present context, suggests that evaluators rely on one or a set of specific examples of "successful jobholders" to which applicants are compared (Nosofsky & Zaki, 2002). These frames-of-reference differ from prototypical frames-of-reference, which are believed to be developed based on the aggregated characteristics of those that evaluators have categorized as "successful jobholders". By using the term prototype I am not suggesting that prototypes hold greater importance than exemplars to the mental schemas evaluators use to compare to applicant stereotypes. Instead, I use the term prototype due to organizational

psychologists' greater familiarity with the term (e.g., Lord, Brown, Harvey, & Hall, 2001). It was not possible to identify the extent to which participants used prototypes or exemplars in developing what I refer to as "prototype ratings" in the present study. It may be informative for future research to further establish the extent to which evaluators rely on exemplars, as opposed to prototypes, in evaluating applicants. In addition, the present study did not measure the strength of participants' stereotypes or prototypes (or exemplars) with regard to warmth and competence. Although extant evidence from various fields suggest a bifactor structure of constructs reflecting warmth and competence (see Cuddy et al., 2008), it remains unclear how important participants viewed these factors to be with regard to stereotype and prototype perceptions in this evaluative context. Further, individual differences likely exist with regard to stereotype and prototype strength on warmth and competence factors, and the strength of individual stereotypes and/or prototypes likely affect the presence of prototype matching. That is, prototype matching is likely more pronounced among those with stronger warmth and competence stereotypes and prototypes, and less pronounced among those with weaker warmth and competence stereotypes and prototypes.

My failure to find weight discrimination also highlights important conceptual and methodological issues. First, I pilot tested applicant photos to obtain some control over applicant attractiveness, and I found no evidence of weight discrimination in this study. I mentioned above that my failure to find weight discrimination may have been due to a relatively subtle weight manipulation (i.e., difference between the normal weight and overweight condition), the relationship between weight and attractiveness, or possibly a combination of the two. These findings highlight a number of potentially important issues conducting and interpreting research on weight discrimination in the future. First, there is little consistency between studies in the way in which normal weight and overweight conditions are operationalized. As described above, the difference in weight conditions has ranged from 30 pounds to over 200 pounds. In order to make strong conclusions regarding the magnitude of weight discrimination, greater emphasis should be placed on understanding how the magnitude of weight discrimination differs as a function of the difference between normal- and overweight conditions. As described above, studies

employing more extreme weight differences between conditions (e.g., Brink, 1988) also typically find stronger magnitudes of weight discrimination than studies employing more subtle weight differences between conditions (e.g., Rothblum et al., 1988). However, relatively little attention has been given to the effects of weight differences on the observed magnitude of weight discrimination. Second, it is possible that some evidence of weight discrimination has actually been a function of discrimination due to attractiveness. Although lab-based weight discrimination research using stimulus photos has controlled for attractiveness (Shapiro, King, & Quinones, 2007), it is still largely unknown how attractiveness interacts with weight to predict work-related outcomes at different weight difference intervals (e.g., 25, 50, 75, 100 pounds). Greater consideration of these two potential moderators may ultimately lead to a more accurate understanding of the effects of weight in personnel evaluation.

## **Conclusion**

This study presented the first attempt to comprehensively test the relationships that make up the prototype matching model. These findings indicate that the prototype matching model can explain the process from which at least some discrimination results. Specifically, evidence presented in this study suggests (a) that pro-female discrimination occurred on all three evaluative outcomes, (b) that stereotypes fully mediated the relationship between applicant sex and each evaluative outcome, and (c) that evaluators' prototypes may be important to better understanding the process which leads to discrimination – as indicated by prototypes moderating the relationship between evaluators' stereotypes of applicants and ratings of applicants on evaluative outcomes. These findings suggest that the prototype matching model may have utility in explaining discrimination, and I have outlined a number of possible future directions for further assessing the potential utility of this model.

## TABLES AND FIGURES

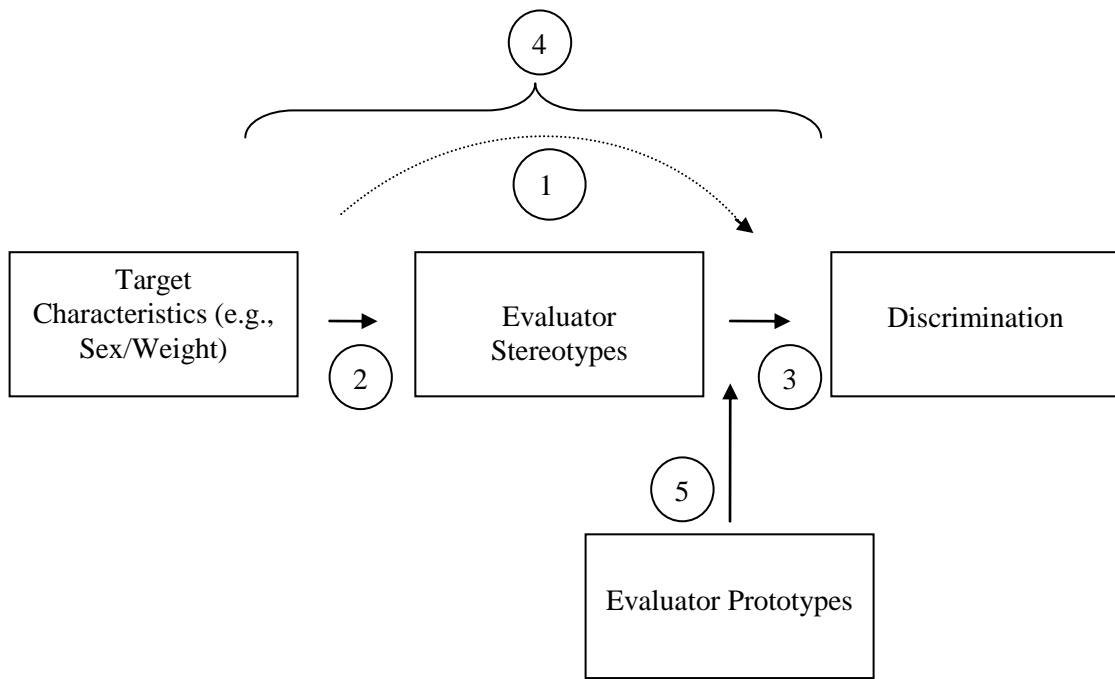


Figure 1

*Moderated Mediation Model for Biased Evaluation*

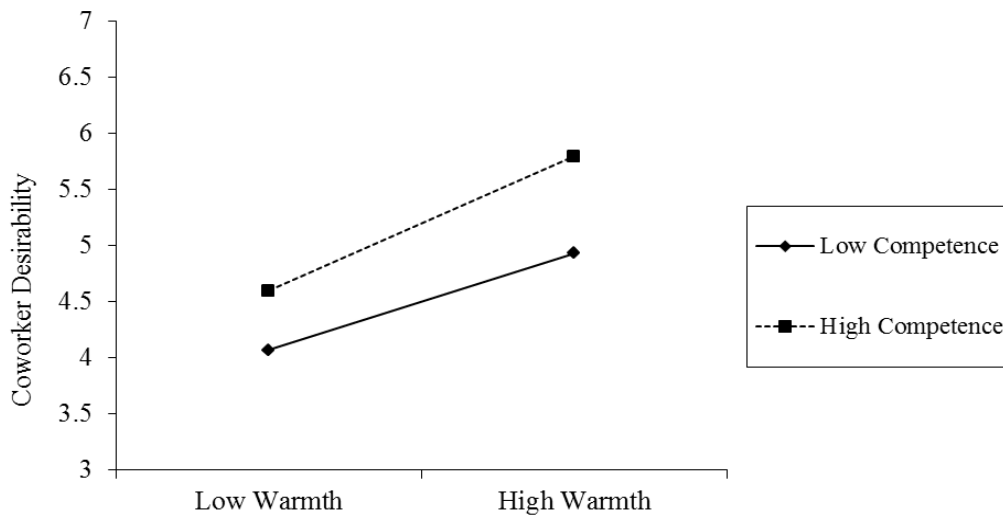


Figure 2

*Stereotype Warmth x Competence Predicting Coworker Desirability*

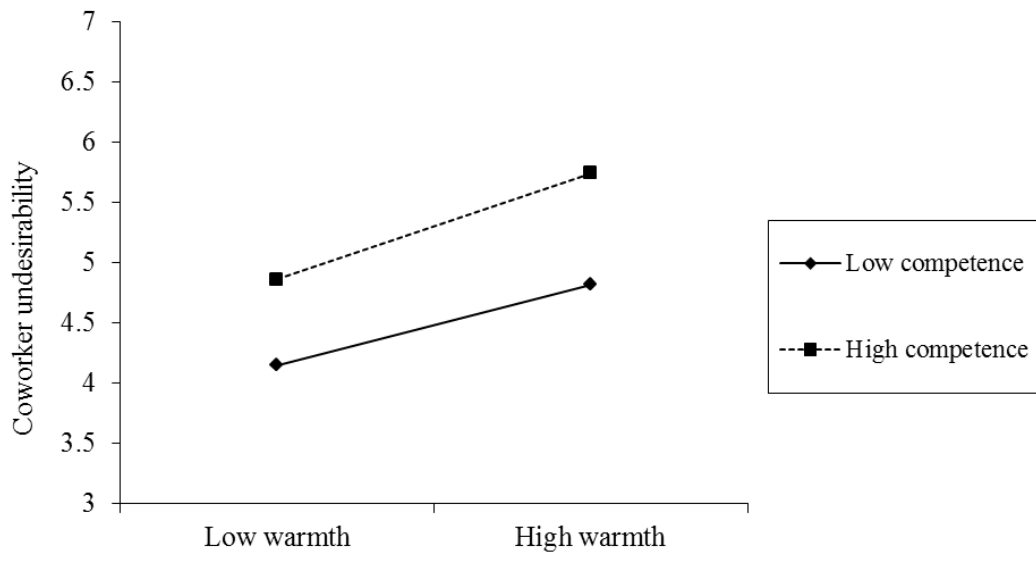


Figure 3

*Stereotype Warmth x Competence Predicting Coworker Undesirability*

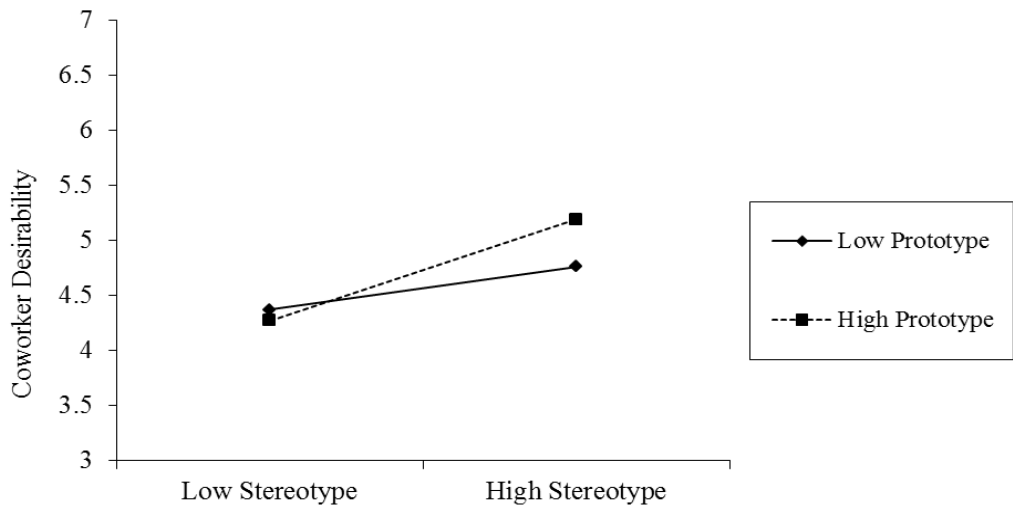


Figure 4a

*Stereotype Competence x Prototype Competence Predicting Coworker Desirability in Sales Position*

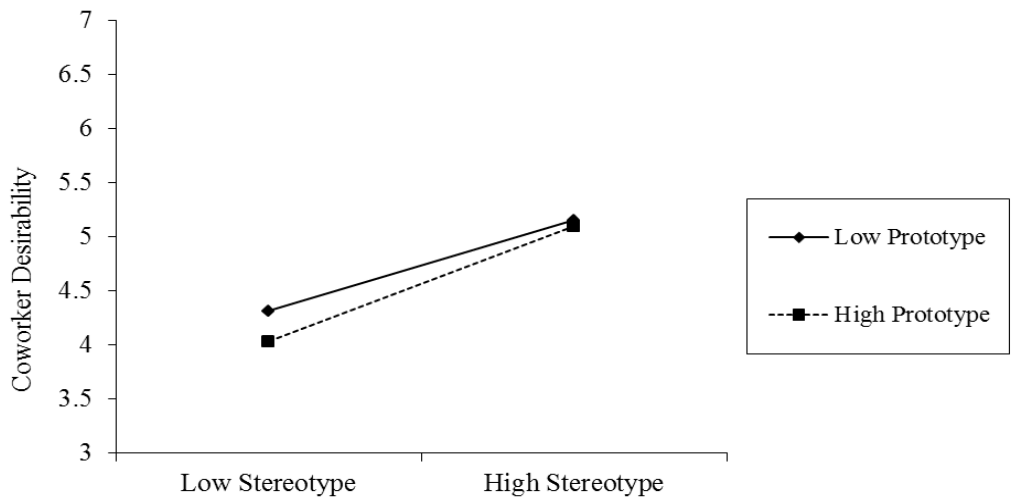


Figure 4b

*Stereotype Warmth x Prototype Warmth Predicting Coworker Desirability in Sales Position*



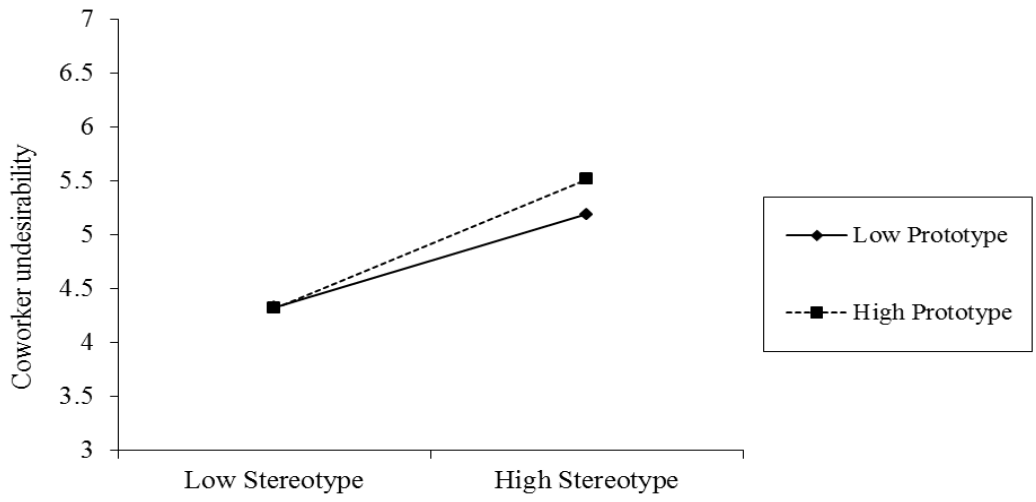


Figure 5a

*Stereotype Competence x Prototype Competence Predicting Coworker Undesirability in Sales Position*

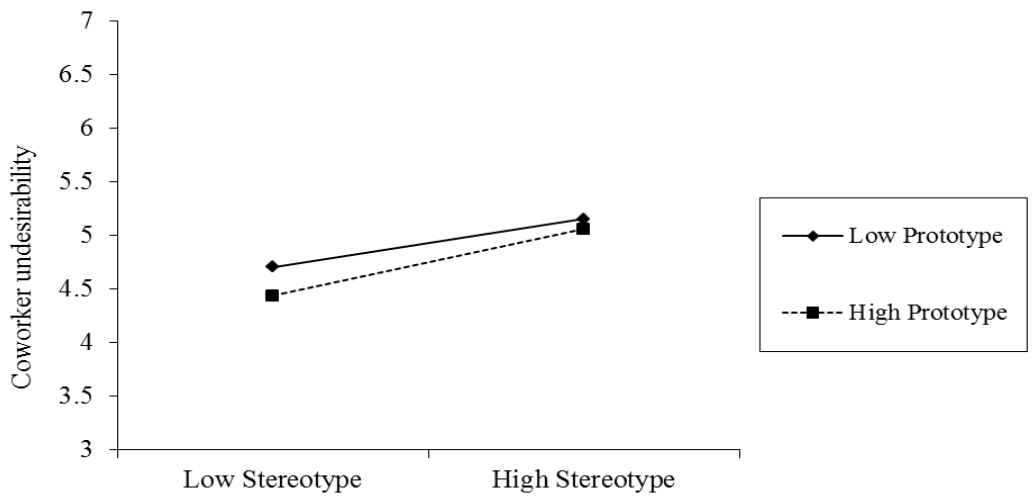


Figure 5b

*Stereotype Warmth x Prototype Warmth Predicting Coworker Undesirability in Sales Position*

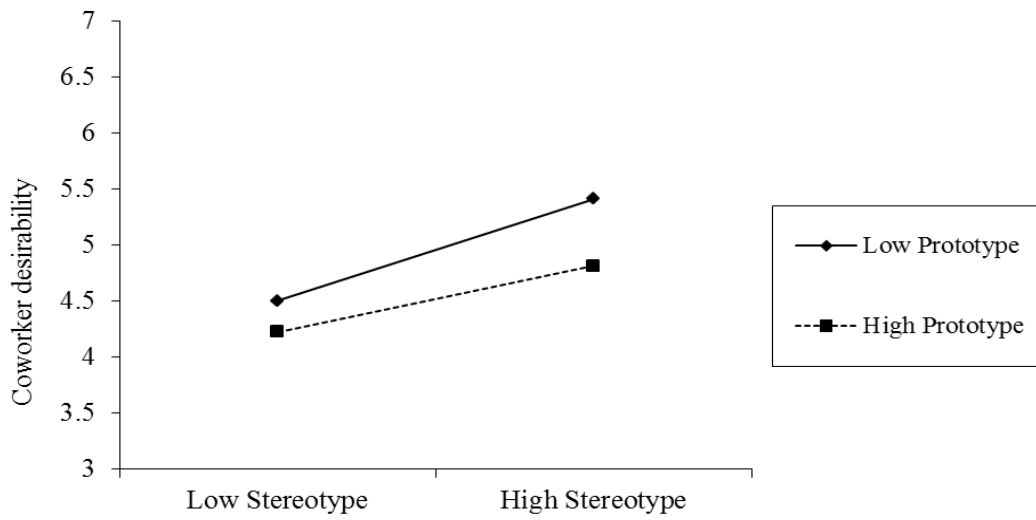


Figure 6a

*Stereotype Competence x Prototype Competence Predicting Coworker Desirability in Customer Service Position*

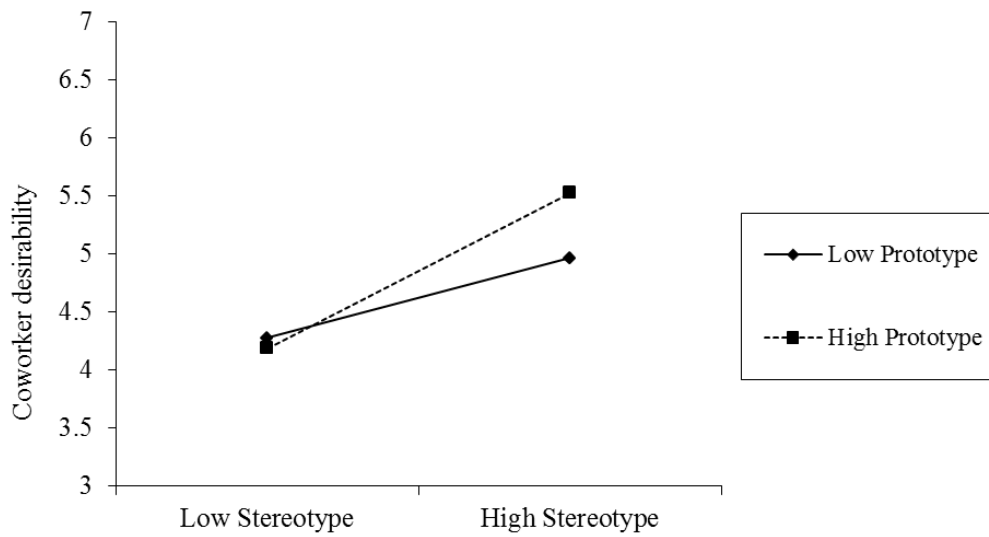


Figure 6b

*Stereotype Warmth x Prototype Warmth Predicting Coworker Desirability in Customer Service Position*

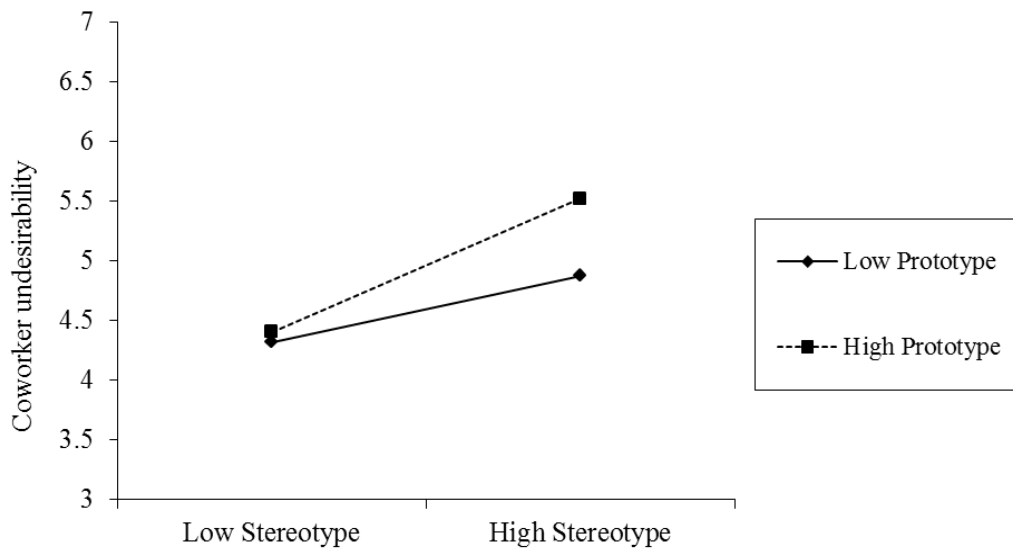


Figure 7

*Stereotype Warmth x Prototype Warmth Predicting Coworker Undesirability in Customer Service Position*

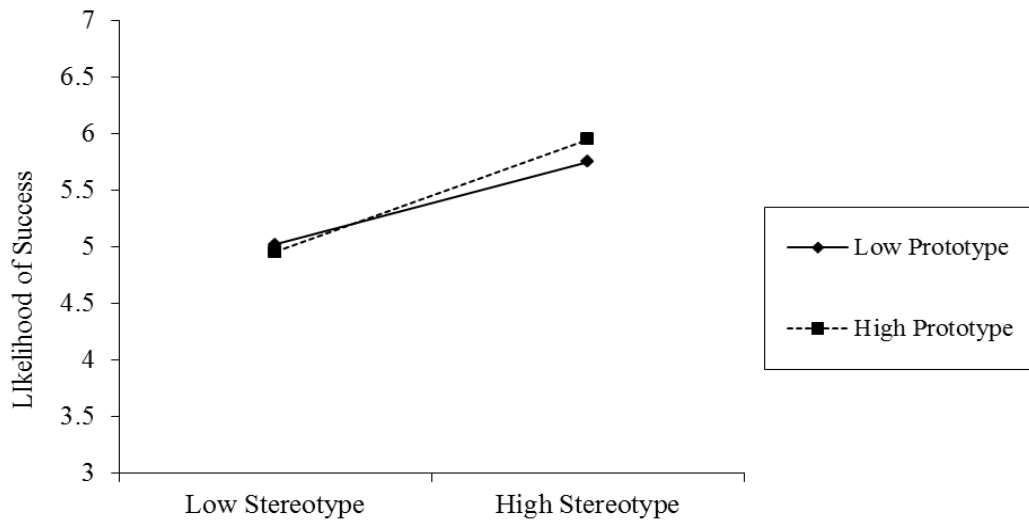


Figure 8

*Stereotype Warmth x Prototype Warmth Predicting Likelihood of Success in Customer Service Position*

Table 1

*Means and Standard Deviations for Pilot Study Weight and Attractiveness Ratings of Applicant Photos*

	Weight		Attractiveness	
	M	SD	M	SD
Overweight male #1	3.47	0.92	1.53	0.67
*Overweight male #2	3.22	0.97	2.34	1.00
*Overweight male #3	3.94	0.89	1.65	0.66
Overweight male #4	2.66	1.07	2.03	0.66
Normal weight male #1	1.22	0.42	3.22	0.75
*Normal weight male #2	1.63	0.75	1.94	0.80
Normal weight male #3	1.06	0.25	3.66	0.83
*Normal weight male #4	1.03	0.18	2.34	0.87
Overweight female #1	3.19	0.90	2.06	0.84
*Overweight female #2	3.00	0.76	2.41	0.84
Overweight female #3	2.50	0.95	2.22	0.97
*Overweight female #4	3.44	0.91	2.00	0.98
*Normal weight female #1	1.53	0.67	3.09	0.82
Normal weight female #2	1.09	0.39	3.38	0.66
*Normal weight female #3	1.00	0.00	1.91	0.89
Normal weight female #4	1.13	0.42	3.66	0.83

*Notes:*  $N = 26$ ; \* indicates that photo was used in main study.

Table 2

*Means and Standard Deviations for Pilot Study Warmth and Competence Ratings of Resume Content*

	Competence		Warmth	
	M	SD	M	SD
<b>Sales resume #1</b>				
Bachelor's degree, Business Management.	3.86	1.14	5.33	0.93
College GPA of 3.0.	3.80	1.02	4.38	0.72
Business owner, Amway.	4.08	1.03	6.09	0.72
Part-time customer service, Johnson's Auto Sales (car dealership).	4.23	1.28	3.86	1.02
Student President, Future Business Leaders of America (FBLA), Wofford University chapter.	4.88	0.91	5.96	0.90
Total	4.17	1.08	5.13	0.86
<b>Sales resume #2</b>				
Bachelor's degree, Business Management.	3.86	1.14	5.33	0.93
College GPA of 3.0.	3.80	1.02	4.38	0.72
Internship, IBM Sales Division.	4.06	1.15	5.48	0.89
Part-time Sales representative GNC.	4.35	0.83	4.37	0.94
Tutor, St. Albert's School of Business Administration.	5.39	0.96	5.64	0.81
Total	4.29	1.02	5.04	0.86
<b>Sales resume #3</b>				
Bachelor's degree, Marketing.	4.35	0.92	5.19	0.81
College GPA of 3.2.	4.17	1.27	4.74	0.84
Owner of summer painting business, Student Painting Services, LLC.	4.54	1.01	5.44	1.02
Part-time commissioned sales, Garrett's Footwear.	4.28	1.04	4.43	1.25
Member, Evansville St. University Entrepreneur Club.	4.28	1.30	5.42	1.17
Total	4.32	1.11	5.05	1.02
<b>Sales resume #4</b>				
Bachelor's degree, Marketing.	4.35	0.92	5.19	0.81
College GPA of 3.2.	4.17	1.27	4.74	0.84
Internship, Coca-Cola Marketing Division.	4.41	1.02	5.28	0.88
Self-employed, EBAY purchasing and sales, Antique Findings, Inc., LLC.	3.31	1.11	4.98	1.15
Peer advisor, Fairview Business School.	5.57	0.83	5.01	0.83
Total	4.36	1.03	5.04	0.90

*(table continues)*

Competence

Warmth

Service resume #1	M	SD	M	SD
Bachelor's degree, Psychology.	5.22	1.09	5.14	0.91
College GPA of 2.8.	3.73	0.97	3.65	0.72
Office administrator, Hilltop Health Center.	4.62	1.23	4.72	1.21
Part-time university tour guide.	5.60	0.82	4.51	0.97
Volunteer counselor, Keystone Juvenile Rehabilitation Center.	6.92	3.29	5.18	0.78
Total	5.22	1.48	4.64	0.92

Service resume #2	Competence		Warmth	
	M	SD	M	SD
Bachelor's degree, Social Work.	5.48	1.04	4.87	0.78
College GPA of 3.0.	3.80	1.02	4.38	0.72
Part-time mental health care aid, St. Michael's.	5.67	0.79	4.92	0.86
Part-time sales/customer service, Jared - Galleria of Jewelry.	4.65	0.82	4.40	0.87
Volunteer, St. Mary's Homeless Shelter.	6.27	0.69	3.74	1.15
Total	5.18	0.87	4.46	0.88

Service resume #3	Competence		Warmth	
	M	SD	M	SD
Bachelor's degree, Nursing.	5.57	0.88	5.27	0.85
College GPA of 3.2.	4.17	1.27	4.74	0.84
Licensed Practical Nurse (LPN), Malone Senior Living Center.	5.90	0.84	5.28	0.84
Part-time server, Juan's Mexican Bistro.	4.38	1.09	3.72	0.95
Volunteer, PeaceCorp.	6.07	0.63	5.36	0.74
Total	5.22	0.94	4.88	0.84

Service resume #4	Competence		Warmth	
	M	SD	M	SD
Bachelor's degree, Physical Therapy.	4.88	1.19	5.06	0.86
College GPA of 2.8.	3.73	0.97	3.65	0.72
Peer counselor, Daniels College.	6.14	0.76	5.18	0.85
Certified Nursing Assistant (CNA), Malone Senior Living Center.	5.91	0.79	5.03	0.80
Volunteer, Habitat for Humanity.	5.64	1.30	4.52	1.16
Total	5.26	1.00	4.69	0.88

Table 3

*Means, Standard Deviations, and Correlations Among Study Variables*

	M	SD	1	2	3	4	5	6	7
1. Conscientiousness (participant)	4.09	0.66							
2. Sex (participant)	0.38	0.49	-.16						
3. Experience: evaluation (participant)	0.44	0.50	.12	-.01					
4. Experience: sales (participant)	0.47	0.50	.15	.00	.20				
5. Experience: customer service (participant)	0.75	0.43	.04	-.09	.23	.38			
6. Experience: healthcare (participant)	0.24	0.43	.02	-.13	.16	.07	.11		
7. Sex (applicant)	0.50	0.50	-	-	-	-	-	-	
8. Weight (applicant)	0.50	0.50	-	-	-	-	-	-	-
9. Applying position (applicant)	0.50	0.50	-	-	-	-	-	-	-
10. Warmth stereotype	5.37	1.00	.17	-.12	-.03	.05	.04	.00	-.08
11. Competence stereotype	5.39	1.06	.16	-.12	-.03	.06	.01	-.01	-.09
12. Warmth prototype (sales)	5.86	0.90	.28	-.23	.01	.04	-.01	.04	.00
13. Competence prototype (sales)	6.10	0.73	.30	-.19	-.01	.00	.02	.00	.00
14. Warmth prototype (service)	6.36	0.84	.22	-.16	-.08	-.02	-.02	.00	.00
15. Competence prototype (service)	5.48	0.85	.25	-.13	-.02	.09	.01	.02	.00
16. Coworker desirability	5.10	1.21	.16	-.13	-.09	.05	.01	-.06	-.06
17. Coworker undesirability	5.26	1.19	.18	-.14	-.09	.05	-.01	-.07	-.06
18. Likelihood of success	5.22	1.27	.09	-.10	-.04	.02	-.01	-.01	-.07

*Notes:* correlations of .03 and greater are statistically significant.  $N=3976-3992$ .

*(table continues)*



	8	9	10	11	12	13	14	15	16	17
1. Conscientiousness (participant)										
2. Sex (participant)										
3. Experience: evaluation (participant)										
4. Experience: sales (participant)										
5. Experience: customer service (participant)										
6. Experience: healthcare (participant)										
7. Sex (applicant)										
8. Weight (applicant)										
9. Applying position (applicant)	-									
10. Warmth stereotype	-.05	-.21								
11. Competence stereotype	.09	.13	.62							
12. Warmth prototype (sales)	.00	.00	.32	.34						
13. Competence prototype (sales)	.00	.00	.44	.37	.49					
14. Warmth prototype (service)	.00	.00	.27	.27	.39	.49				
15. Competence prototype (service)	.00	.00	.35	.32	.45	.43	.49			
16. Coworker desirability	.02	.02	.52	.58	.22	.30	.22	.18		
17. Coworker undesirability	.01	.01	.55	.55	.21	.33	.26	.18	.82	
18. Likelihood of success	.00	-.05	.68	.57	.21	.29	.19	.18	.67	.73

Notes: correlations of .03 and greater are statistically significant.  $N=3976-3992$ .

Table 4

*Means, Standard Deviations, and Effect Sizes (ds) for Applicant Warmth and Competence Ratings based on Sex and Weight*

	Competence			Warmth		
	M	SD	<i>d</i> [95% CIs]	M	SD	<i>d</i> [95% CIs]
Overweight	5.33	0.78	0.12 [.07, .17]	5.48	0.78	-0.22 [-.27, -.17]
Normal weight	5.42	0.76		5.30	0.83	
Male	5.30	0.79	-0.20 [-.24, -.15]	5.29	0.81	-0.25 [-.30, -.20]
Female	5.45	0.75		5.49	0.77	

*Notes:* positive effect sizes represent biases towards normal weight and male applicants. *N*=499.

Table 5

*Means, Standard Deviations, and Effect Sizes (ds) for Applicant Evaluations based on Sex and Weight*

Sales Position	<u>Coworker Desirability</u>			<u>Coworker Undesirability</u>			<u>Success</u>		
	M	SD	<i>d</i> [95% CIs]	M	SD	<i>d</i> [95% CIs]	M	SD	<i>d</i> [95% CIs]
Overweight	5.10	1.02	-0.06 [-.12, .00]	5.26	1.01	-0.03 [-.09, .03]	5.28	1.01	0.01 [-.05, .07]
Normal weight	5.04	0.97		5.23	0.98		5.29	1.02	
Male	4.99	1.01	-0.16 [-.22, -.10]	5.19	1.00	-0.12 [-.18, -.06]	5.21	1.03	-0.15 [-.21, -.09]
Female	5.15	0.99		5.31	0.97		5.36	0.99	

Service Position	<u>Coworker</u>			<u>Coworker Undesirability</u>			<u>Success</u>		
	M	SD	<i>d</i> [95% CIs]	M	SD	<i>d</i> [95% CIs]	M	SD	<i>d</i> [95% CIs]
Overweight	5.15	1.03	-0.05 [-.11, .02]	5.27	1.03	0.00 [-.07, .07]	5.17	1.08	-0.02 [-.09, .05]
Normal weight	5.10	1.06		5.27	1.06		5.15	1.09	
Male	5.05	1.02	-0.14 [-.20, -.08]	5.19	1.03	-0.15 [-.21, -.09]	5.07	1.06	-0.17 [-.24, -.11]
Female	5.19	1.00		5.34	1.00		5.25	1.01	

*Notes:* positive effect sizes represent biases towards normal weight and male applicants.  $N=499$ .

Table 6

*Model Fit Statistics for each Model at each Step.*

	AIC	BIC	-2logLik	$\chi^2$	$\Delta D$	(df)
Step 1	Coworker desirability					
Model 1	11808	11827	-5901	11802		
Model 2	11667	11723	-5824	11648	154	(11)*
Model 3	11655	11730	-5815	11630	18	(3)*
Model 4	11669	11763	-5819	11638	-8	(3)*
	Coworker undesirability					
Model 1	11548	11567	-5771	11542		
Model 2	11390	11447	-5686	11372	170	(11)*
Model 3	11390	11465	-5683	11366	6	(3)
Model 4	11404	11498	-5687	11374	-2	(9)
	Likelihood of success					
Model 1	12520	12538	-6257	12514		
Model 2	12400	12456	-6191	12382	132	(11)*
Model 3	12379	12454	-6177	12354	28	(3)*
Model 4	12394	12489	-6182	12364	-10	(3)
Step 2	Warmth					
Model 1	10492	10510	-5243	10486		
Model 2	10392	10449	-5187	10374	112	(11)*
Model 3	10180	10256	-5078	10156	218	(3)*
Model 4	10201	10301	-5084	10168	-12	(4)
	Competence					
Model 1	9926	9944	-4960	9920		
Model 2	9821	9877	-4901	9802	118	(11)*
Model 3	9453	9528	-4714	9428	374	(3)*
Model 4	9472	9573	-4720	9440	-12	(4)

*(table continues)*

		AIC	BIC	-2logLik	$\chi^2$	$\Delta D$	(df)
Step 3	Coworker desirability						
	Model 3	10068	10137	-5023	10046	1602	(2)*
	Model 4	10017	10092	-4996	9992	54	(1)*
	Coworker undesirability						
	Model 3	9964	10033	-4971	9942	1430	(2)*
	Model 4	9949	10024	-4962	9924	18	(1)*
	Likelihood of success						
	Model 3	10136	10205	-5057	10114	2268	(2)*
	Model 4	10144	10219	-5060	10120	-6	(1)
Step 4	Coworker desirability						
		10024	10106	-4999	9998	-6	(1)
	Coworker undesirability						
		9956	10038	-4965	9930	-6	(1)
Likelihood of success							
		10143	10225	-5059	10118	-4	(2)
Prototype matching							
Sales	Coworker desirability						
		5103	5192	-2535	5070		
	Coworker undesirability						
		5073	5163	-2521	5042		
Likelihood of success							
		4920	5010	-2444	4888		
Customer service	Coworker desirability						
		5195	5284	-2582	5164		
	Coworker undesirability						
		5094	5183	-2531	5062		
Likelihood of success							
		5208	5297	-2588	5176		

Notes: \*indicates model fits significantly better than previous model ( $p < .05$ ); the exception is Model 4 (Step 1) predicting coworker undesirability, in which Model 4 was compared to Model 2, as Model 3 did not result in a significant improvement over Model 2. Model 3 (for each outcome) at Step 3 were compared to the relevant Model 2 at Step 1, as the unconditional model (Model 1) and control variable model (Model 2) did not differ between Steps 1 and 3. Step 4 models are compared to the preferred model at Step 3.

Table 7

*Step 1: Applicant Sex and Weight Predicting Outcomes*

	Coworker Desirability						Coworker Undesirability					
	Model 2a		Model 3a		Model 4a		Model 2b		Model 3b		Model 4b	
	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE
Intercept	4.15	0.24	4.18	0.24	4.21	0.24	4.27	0.24	4.32	0.24	4.33	0.24
<i>Level 2 Control Variables</i>												
Participant conscientiousness	0.28*	0.05	0.28*	0.05	0.28*	0.05	0.30*	0.05	0.30*	0.05	0.30*	0.05
Participant sex	-0.28*	0.07	-0.28*	0.07	-0.28*	0.07	-	0.07	-0.32*	0.07	-0.32*	0.07
Participant exp. 1 (evaluation)	-0.27*	0.07	-0.27*	0.07	-0.27*	0.07	-	0.07	-0.27*	0.07	-0.27*	0.07
Participant exp. 2 (sales)	0.15	0.08	0.15	0.08	0.15	0.08	0.13	0.08	0.13	0.08	0.13	0.08
Participant exp. 3 (service)	-0.01	0.09	-0.01	0.09	-0.01	0.09	-0.04	0.09	-0.04	0.09	-0.04	0.09
Participant exp. 4 (healthcare)	-0.18	0.08	-0.18	0.08	-0.18	0.08	-0.21	0.08	-0.21	0.08	-0.21	0.08
<i>Level 1 Predictors</i>												
Applicant sex			-0.16*	0.03	-0.23*	0.05			-0.13*	0.03	-0.17*	0.05
Applicant weight			0.05	0.03	0.00	0.05			0.01	0.03	-0.02	0.05
Position for which applying			0.05	0.03	0.04	0.05			0.02	0.03	0.05	0.05
Applicant sex x weight					0.11	0.06					0.10	0.06
Applicant sex x position					0.03	0.06					-0.02	0.06
Applicant weight x position					-0.01	0.06					-0.03	0.06

Likelihood of Success						
	Model 2c		Model 3c		Model 4c	
	$\beta$	SE	$\beta$	SE	$\beta$	SE
Intercept	4.79	0.24	4.95	0.24	4.96	0.24
<i>Level 2 Control Variables</i>						
Participant conscientiousness	0.14*	0.06	0.14*	0.06	0.14	0.06
Participant sex	-0.23*	0.08	-0.23*	0.08	-0.23*	0.08
Participant exp. 1 (evaluation)	-0.13	0.08	-0.13	0.08	-0.13	0.08
Participant exp. 2 (sales)	0.09	0.08	0.09	0.08	0.09	0.08
Participant exp. 3 (service)	-0.07	0.09	-0.07	0.09	-0.07	0.09
Participant exp. 4 (healthcare)	-0.04	0.08	-0.04	0.08	-0.04	0.08
<i>Level 1 Predictors</i>						
Applicant sex			-0.17*	0.03	-0.20*	0.06
Applicant weight			0.00	0.03	-0.04	0.06
Position for which applying			-0.13*	0.03	-0.13	0.06
Applicant sex x weight					0.06	0.07
Applicant sex x position					-0.01	0.07
Applicant weight x position					0.02	0.07

Notes: \* $p < .00625$ . Model 1 predicting each outcome was the unconditional model (i.e., no predictors), thus not presented in this Table.

Table 8

*Step 2: Applicant Sex and Weight Predicting Stereotype Content (Warmth and Competence)*

	Warmth						Competence					
	Model 2d		Model 3d		Model 4d		Model 2e		Model 3e		Model 4e	
	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE
Intercept	4.54	0.22	4.41	0.22	4.41	0.22	4.48	0.22	4.81	0.22	4.78	0.22
<i>Level 2 Control Variables</i>												
Participant												
conscientiousness	0.23*	0.05	0.23*	0.05	0.23*	0.05	0.23*	0.05	0.23*	0.05	0.23*	0.05
Participant sex	-0.21*	0.07	-0.21*	0.07	-0.21*	0.07	-0.20*	0.07	-0.20*	0.07	-0.20*	0.07
Participant exp. 1 (evaluation)	-0.11	0.07	-0.11	0.07	-0.11	0.07	-0.13	0.07	-0.13	0.07	-0.13	0.07
Participant exp. 2 (sales)	0.12	0.07	0.12	0.07	0.12	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Participant exp. 3 (service)	-0.03	0.08	-0.03	0.08	-0.03	0.08	0.06	0.08	0.06	0.08	0.06	0.08
Participant exp. 4 (healthcare)	-0.05	0.08	-0.05	0.08	-0.05	0.08	-0.04	0.08	-0.04	0.08	-0.04	0.08
<i>Level 1 Predictors</i>												
Applicant sex			-0.20*	0.02	-0.22*	0.05			-0.15*	0.02	-0.08	0.04
Applicant weight			0.18*	0.02	0.17*	0.05			-0.10*	0.02	-0.07	0.04
Position for which applying			0.28*	0.02	0.30*	0.05			-0.41*	0.02	-0.37*	0.04
Applicant sex x weight					0.06	0.07					-0.08	0.06
Applicant weight x position					-0.02	0.07					-0.12	0.06
Applicant sex x position					-0.06	0.07					-0.03	0.06
Sex x weight x position					0.03	0.10					0.11	0.09

Notes: \* $p < .00625$ . Model 1 predicting each outcome was the unconditional model (i.e., no predictors), thus not presented in this Table.



Table 9

*Step 3: Stereotype Content (Warmth and Competence) Predicting Outcomes*

	Customer Desirability				Customer Undesirability				Likelihood of Success			
	Model 3f		Model 4f		Model 3g		Model 4g		Model 3h		Model 4h	
	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE
Intercept	4.86	0.19	4.85	0.19	4.94	0.18	4.93	0.18	5.68	0.15	5.68	0.15
<i>Level 2 Control Variables</i>												
Participant conscientiousness	0.09	0.04	0.08	0.04	0.12*	0.04	0.12*	0.04	-0.09*	0.03	-0.09*	0.03
Participant sex	-0.12	0.06	-0.11	0.06	-0.16*	0.06	-0.16*	0.06	-0.03	0.05	-0.03	0.05
Participant exp. 1 (evaluation)	-0.18*	0.06	-0.18*	0.06	-0.17*	0.06	-0.17*	0.06	-0.01	0.05	-0.01	0.05
Participant exp. 2 (sales)	0.07	0.06	0.05	0.06	0.06	0.06	0.05	0.06	0.00	0.05	0.00	0.05
Participant exp. 3 (service)	-0.01	0.07	-0.01	0.07	-0.05	0.07	-0.05	0.07	-0.10	0.06	-0.10	0.06
Participant exp. 4 (healthcare)	-0.14	0.07	-0.14	0.07	-0.18*	0.06	-0.18*	0.06	0.01	0.05	0.00	0.05
<i>Level 1 Predictors</i>												
Warmth	0.50*	0.02	0.51*	0.02	0.38*	0.02	0.39*	0.02	0.33*	0.02	0.33*	0.02
Competence	0.31*	0.02	0.35*	0.02	0.39*	0.02	0.41*	0.02	0.68*	0.02	0.69*	0.02
Warmth x competence			0.08*	0.01			0.05*	0.01			0.01	0.01

Notes: \* $p < .00625$ . Model 1 predicting each outcome was the unconditional model (i.e., no predictors), thus not presented in this Table.

Table 10

*Step 4: Applicant Sex Predicting Outcomes Controlling for Stereotypes*

	Customer Desirability		Customer Undesirability		Likelihood of Success	
	Model 2a		Model 2b		Model 2c	
	$\beta$	SE	$\beta$	SE	$\beta$	SE
Intercept	4.85	0.19	4.93	0.18	5.65	0.15
<i>Level 2 Control Variables</i>						
Participant conscientiousness	0.08	0.04	0.12*	0.04	-0.09	0.03
Participant sex	-0.11	0.06	-0.16*	0.06	-0.03	0.05
Participant exp. 1 (evaluation)	-0.18*	0.06	-0.17*	0.06	-0.01	0.05
Participant exp. 2 (sales)	0.05	0.06	0.05	0.06	0.00	0.05
Participant exp. 3 (service)	-0.01	0.07	-0.05	0.07	-0.10	0.06
Participant exp. 4 (healthcare)	-0.14	0.07	-0.18*	0.06	0.00	0.05
<i>Level 1 Predictors</i>						
Applicant sex	-0.01	0.02	0.00	0.02	-0.01	0.03
Position for which applying					0.08	0.03
<i>Level 1 Mediators</i>						
Warmth	0.51*	0.02	0.39*	0.02	0.31*	0.02
Competence	0.35*	0.02	0.41*	0.02	0.70*	0.02
Warmth x competence	0.08*	0.01	0.05*	0.01		

Notes: \* $p < .00625$ . Model 1 predicting each outcome was the unconditional model (i.e., no predictors), thus not presented in this Table.

Table 11a

*Step 5: Effects of Prototypes and Stereotype-Prototype Interactions on Outcomes for Sales Positions*

	Sales Position					
	Coworker Desirability		Coworker Undesirability		Likelihood of Success	
	Model 2a		Model 2b		Model 2c	
	$\beta$	SE	$\beta$	SE	$\beta$	SE
Intercept	4.69	0.21	4.86	0.21	5.45	0.17
<i>Level 2 Control Variables</i>						
Participant conscientiousness	0.11	0.05	0.11	0.05	-0.06	0.04
Participant sex	-0.10	0.07	-0.16	0.06	-0.01	0.05
Participant exp. 1 (evaluation)	-0.15	0.06	-0.15	0.06	0.01	0.05
Participant exp. 2 (sales)	0.11	0.07	0.10	0.06	0.01	0.05
Participant exp. 3 (customer serv.)	-0.05	0.08	-0.10	0.07	-0.13	0.06
Participant exp. 4 (healthcare)	-0.14	0.07	-0.17	0.07	0.03	0.06
<i>Substantive Predictors</i>						
Applicant sex	-0.04	0.03	-0.01	0.03	-0.03	0.03
Warmth (stereotype)	0.47*	0.03	0.28*	0.03	0.19*	0.03
Competence (stereotype)	0.33*	0.03	0.48*	0.03	0.88*	0.03
Warmth (prototype)	-0.09	0.04	-0.08	0.04	-0.04	0.03
Competence (prototype)	0.08	0.05	0.09	0.05	-0.15*	0.04
Warmth (s) x warmth (p)	0.06	0.02	0.06	0.02	0.06	0.02
Competence (s) x competence (p)	0.13*	0.03	0.12*	0.03	0.01	0.03

Notes: \* $p < .00625$ . Model 1 predicting each outcome was the unconditional model (i.e., no predictors), thus not presented in this Table.

Table 11b

*Step 5: Effects of Prototypes and Stereotype-Prototype Interactions on Outcomes for Customer Service Positions*

	Coworker Desirability		Service Position Coworker Undesirability		Likelihood of Success	
	Model 2a		Model 2b		Model 2c	
	$\beta$	SE	$\beta$	SE	$\beta$	SE
Intercept	4.74	0.21	4.78	0.21	5.42	0.19
<i>Level 2 Control Variables</i>						
Participant conscientiousness	0.13	0.05	0.17	0.05	-0.03	0.04
Participant sex	-0.16	0.06	-0.18	0.06	-0.10	0.06
Participant exp. 1 (evaluation)	-0.19	0.06	-0.16	0.06	-0.02	0.06
Participant exp. 2 (sales)	0.04	0.07	0.04	0.07	0.00	0.06
Participant exp. 3 (customer serv.)	-0.02	0.08	-0.05	0.08	-0.11	0.07
Participant exp. 4 (healthcare)	-0.11	0.07	-0.16	0.07	0.00	0.06
<i>Substantive Predictors</i>						
Applicant sex	0.02	0.03	0.01	0.03	0.02	0.04
Warmth (stereotype)	0.51*	0.03	0.42*	0.03	0.43*	0.03
Competence (stereotype)	0.38*	0.03	0.40*	0.03	0.63*	0.03
Warmth (prototype)	0.12	0.04	0.18*	0.04	0.03	0.04
Competence (prototype)	-0.22*	0.04	-0.22*	0.04	-0.18*	0.04
Warmth (s) x warmth (p)	0.16*	0.02	0.14*	0.02	0.07*	0.02
Competence (s) x competence (p)	-0.08*	0.03	-0.04	0.03	0.04	0.03

*Notes:* \* $p < .00625$ . Model 1 predicting each outcome was the unconditional model (i.e., no predictors), thus not presented in this Table.

## ENDNOTES

<sup>1</sup> Analyses were also conducted across all 610 cases – that is, without implementing data screening procedures. This resulted in some differences. At Step 1 predicting both coworker desirability and undesirability, applicant sex became non-significant when Level 1 interactions were included in Model 4 ( $p = .03$  and  $.02$ , respectively), using the Bonferroni correction threshold of  $p = .0063$ . At Step 4, position remained significant in predicting likelihood of success. There were more substantive differences at Step 5. First, for sales position analyses, stereotype competence was non-significant ( $p = .02$ ) in predicting coworker desirability; in predicting coworker undesirability, stereotype warmth ( $p = .18$ ) and competence ( $p = .02$ ) were non-significant, as was the prototype-stereotype competence interaction ( $p = .50$ ), while prototype competence was significant; in predicting likelihood of success, main effects for stereotype and prototype competence were non-significant ( $p = .85$  and  $.21$ , respectively), but their interaction was significant. Second, for service position analyses, the main effects for both stereotype warmth and competence were non-significant ( $p = .02$  and  $.04$ , respectively), as were both warmth ( $p = .13$ ) and competence ( $p = .02$ ) prototype-stereotype interactions, in predicting coworker desirability; in predicting coworker undesirability, prototype warmth was the only significant substantive predictor; in predicting likelihood of success, the main effect of stereotype warmth was non-significant ( $p = .46$ ). Although results using unscreened data were consistent with results using screened data had Bonferroni corrections not been used (see endnote 2), different conclusions can be drawn when unscreened data are used along with the more stringent  $.0063$  significance threshold. Moreover, where differences exist in results of analyses using cleaned and uncleaned data, findings using the uncleaned data do not align with theorized relationships. Taken together, these findings demonstrate the importance of conducting data cleaning prior to conducting substantive analyses, especially when data are collected in non-proctored settings or when the motivation of participants is either unknown or in question (e.g., when collecting data through MTurk). Not conducting data cleaning procedures in these instances may affect the conclusions that are drawn.

<sup>2</sup> The Bonferroni correction threshold of  $p = .0063$  was used as in represents a more conservative estimate. However, had a threshold of  $p = .05$  been used additional substantive predictors would have been interpreted as significant. In the final model at Step 1 and Step 4 predicting likelihood of success, the  $p$  values associated with position were .02 and .007, respectively. At Step 5 (among sales position data), the  $p$  values associated with prototype-stereotype warmth interaction was .02 for all three outcomes, and prototype warmth ratings were associated with coworker desirability ratings at  $p = .04$ . At Step 5 (among service position data), the  $p$  value associated with prototype warmth and coworker desirability ratings was .008, and with the interaction between prototype-stereotype warmth ratings and likelihood of success was .007. These less conservative significance thresholds result in stronger support for prototype matching.

## REFERENCES

- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*, 751-763.
- Agars, M. D. (2004). Reconsidering the impact of gender stereotypes on the advancement of women in organizations. *Psychology of Women Quarterly, 28*, 103-111.
- Age Discrimination in Employment Act, 29 U.S.C. § 621 (1967).
- Allon, N. (1982). The stigma of overweight in everyday life. In B. B. Woldman (Ed.), *Psychological aspects of obesity* (pp. 130-1174). New York: Random House.
- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Americans with Disabilities Act of 1990, 42 U.S.C. § 12101-12213 (2000).
- Anderson, C. A., Lindsay, J. J., & Bushman, B. J. (1999). Research in the psychological laboratory: truth or triviality? *Current Directions in Psychological Science, 8*, 3-9.
- Arvey, R. D. (1979). Unfair discrimination in the employment interview: legal and psychological aspects. *Psychological Bulletin, 86*, 736-765.
- Averett, S., & Korenman, S. (1996). The economic reality of the beauty myth. *The Journal of Human Resources, 31*, 304-329.
- Bakan, D. (1966), *The duality of human existence*. Reading, PA: Addison-Wesley.
- Baum, C. L., & Ford, W. F. (2004). The wage effects of obesity: a longitudinal study. *Health Economics, 13*, 885-899.
- Bellizzi, J. A., & Hasty, R. W. (2001). The effects of a stated organizational policy on inconsistent disciplinary action based on salesperson gender and weight. *The Journal of Personal Selling and Sales Management, 3*, 189-198.
- Bellizzi, J. A., Klassen, M. L., & Belonax, J. J. (1989). Stereotypical beliefs about overweight and smoking and decision making in assignments to sales territories. *Perceptual and Motor Skills, 69*, 419-429.

- Bevins, B. C. (2003). Employability of individuals with varying disabilities and costs of needed workplace accommodations. *Masters Abstracts International*, *41*, 1524-1555.
- Bowen, C., Swim, J. K., Jacobs, R. R. (2000). Evaluating gender biases on actual job performance of real people: a meta-analysis. *Journal of Applied Social Psychology*, *30*, 2194-2215.
- Brannick, M. T., Levine, E. L., & Morgeson, F. P. (2007). *Job and work analysis: methods, research, and applications for human resource management*. Thousand Oaks, CA: Sage.
- Brewer, M. B. (1988). A dual process model of impression formation. In R. S. Wyer, & T. K. Srull (Eds.), *A dual-process model of impression formation: Advances in social cognition* (pp.1-36). Hillsdale, NJ: Lawrence Erlbaum.
- Brink, T. L. (1988). Obesity and job discrimination: mediation via personality stereotypes? *Perceptual and Motor Skills*, *66*, 494.
- Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, N.J.: Erlbaum.
- Bruner, J. (1957). On perceptual readiness. *Psychology Review*, *64*, 123-152.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*, 3-5.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt, & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.
- Caprariello, P. A., Cuddy, A. J. C., Fiske, S. T. (2009). Social structure shapes cultural stereotypes and emotions: a causal test of the stereotype content model. *Group Processes and Intergroup Behavior*, *12*, 147-155.
- Cascio, W. F., & Aguinis, H. (2011). *Applied psychology in human resource management* (7<sup>th</sup> Edition). 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.



- Civil Rights Act of 1964, Pub. L. No. 88-352, 78 Stat. 241 (codified as amended in scattered sections of 2 U.S.C., 28 U.S.C. and 42 U.S.C.).
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Collins, M. A., & Zebrowitz, L. A. (1995). The contributions of appearance to occupational outcomes in civilian and military settings. *Journal of Applied Social Psychology, 25*, 129-163.
- 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.
- Crocker, J., Cornwell, B., & Major, B. (1993). The stigma of overweight: affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology, 64*, 60-70.
- Cronshaw, S. F., & Lord, R. G. (1987). Effects of categorization, attribution, and encoding processes on leadership perceptions. *Journal of Applied Psychology, 72*, 97-106.
- Croteau, J. M. (1996). Research on the work experiences of lesbian, gay, and bisexual people: an integrative review of methodology and findings. *Journal of Vocational Behavior, 48*, 195-209.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology, 92*, 631-648.
- Cuddy, A. J. C., Fiske, S. T., Kwan, S. Y., Glick, P., Demoulin, S., Leyens, J. P., ...Ziegler, R. (2009). Stereotype content model across cultures: towards universal similarities and some differences. *British Journal of Social Psychology, 48*, 1-33.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: the stereotype content model and the BIAS map. *Advances in Experimental Social Psychology, 42*, 61-149.
- Davison, H. K., & Burke, M. J. (2000). Sex discrimination in simulated employment contexts: a meta-analytic investigation. *Journal of Vocational Behavior, 56*, 225-248.
- Dean, M. A., Roth, P. L., & Bobko, P. (2008). Ethnic and gender subgroup differences in assessment center ratings: a meta-analysis. *Journal of Applied Psychology, 93*, 685-691.
- deJong, J. E., &

- Kaplan, H. (1962), Some differential effects of race or rater and rate on early peer ratings of combat aptitude, *Journal of Applied Psychology*, 46, 370-374.
- Devine, P. G. (1988). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5-18.
- Ding, V. J., & Stillman, J. A. (2005). An empirical investigation of discrimination against overweight female job applicants in New Zealand. *New Zealand Journal of Psychology*, 34, 139-148.
- Dovidio, J. F., & Gaertner, S. L. (2010). Intergroup bias. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology*, Vol. 2 (pp. 1084-1121). New York: Wiley.
- Eagly, A. H. (1985). *Sex differences in social behavior: a social-role interpretation*. Hillsdale, NJ: Erlbaum.
- Eagly, A. H. (1995). The science and politics of comparing women and men. *American Psychologist*, 50, 145-158.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109, 573-598.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: a meta-analysis. *Psychological Bulletin*, 111, 3-22.
- Eckes, T. (2002). Paternalistic and envious gender stereotypes: testing predictions from the stereotype content model. *Sex Roles*, 47, 99-114.
- Engle, E. M., & Lord, R. G. (1997). Implicit theories, self-schemas, and leader-member exchange. *Academy of Management Journal*, 40, 988-1010.
- Fidell, L. S. (1970). Empirical verification of sex discrimination in hiring practices in psychology. *American Psychologist*, 25, 1094-1098.
- Finkelstein, L. M., Frautschy Demuth, R. L., & Sweeney, D. L. (2007). Bias against overweight job applicants: further explorations of when and why. *Human Resource Management*, 46, 203-222.
- Fisher v. UT-Austin, 631 F.3d § 09-50822 (2011).

- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878-902.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum model of impression formation from category-based to individuating processes: influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, *23*, 1-74.
- Fiske, S. T., Xu, J., Cuddy, A., & Glick, P. (1999). (Dis)respecting versus (dis)liking: status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues*, *55*, 473-491.
- Fleishman, E. A. (1953). The measurement of leadership attitudes in industry. *Journal of Applied Psychology*, *37*, 153-158.
- Foti, R. J., Fraser, S. L., & Lord, R. G. (1982). Effects of leadership labels and prototypes on perceptions of political leaders. *Journal of Applied Psychology*, *67*, 326-333.
- Fox, H., & Lefkowitz, J. (1974). Differential validity: ethnic group as a moderator in predicting job performance. *Personnel Psychology*, *27*, 209-223.
- Fraser, S. L., & Lord, R. G. (1988). Stimulus prototypicality and general leadership impressions: their role in leadership and behavioral ratings. *Journal of Psychology*, *122*, 291-303.
- Goldberg, L. R. (1999). 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). Tilburg, Netherlands: Tilburg University Press.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., ... Gough, H. C. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*, 84-96.
- Goldberg, P. A. (1968). Are women prejudiced against women? *Transactions*, *5*, 28-30.
- Goldstein, H. (2011). *Multilevel statistical modeling*. West Sussex, UK: Wiley & Sons.

- Goodman, J. K., Cryder, C. E., & Cheema, A. (2012). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, forthcoming.
- Gordon, R. A., & Arvey, R. D. (2004). Age bias in laboratory and field settings: a meta-analytic investigation. *Journal of Applied Social Psychology*, *34*, 1-27.
- Gordon, L. V., & Medland, F. F. (1965). The cross-group stability of peer ratings of leadership potential. *Personnel Psychology*, *18*, 173-177.
- Graves, L. M., & Powell, G. N. (1995). The effect of sex similarity on recruiters' evaluations of actual applicants: a test of the similarity-attraction paradigm. *Personnel Psychology*, *48*, 85-98.
- Harris, M. B., Harris, R. J., & Bochner, S. (1982). Fat, four-eyed, and female: stereotypes of obesity, glasses, and gender. *Journal of Applied Social Psychology*, *12*, 503-516.
- Haslam, S. A., & Ryan, M. K. (2008). The road to the glass cliff: differences in the perceived suitability of men and women for leadership positions in succeeding and failing organizations. *Leadership Quarterly*, *19*, 530-546.
- Hebl, M. R. (1997). Nonstigmatized individuals' reactions to the acknowledgment and valuation of a stigma by overweight individuals and physically disabled individuals. *Dissertation Abstracts International*, *58*, 3367.
- Hebl, M. R., & Heatherton, T. F. (1998). The stigma of obesity in women: the difference is black and white. *Personality and Social Psychology Bulletin*, *24*, 417-426.
- Heilman, M. E. (1983). Sex bias in work settings: the lack of fit model. *Research in Organizational Behavior*, *5*, 269-298.
- Heilman, M. E. (2001). Description and prescription: how gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, *57*, 657-674.
- Heilman, M. E., Block, C., J., Martell, R. F., & Simon, M. C. (1989). Has anything changed? Current characterizations of men, women, and managers. *Journal of Applied Psychology*, *74*, 935-942.

- Heilman, M. E., & Eagly, A. H. (2008). Gender stereotypes are alive and well, and busy producing workplace discrimination. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 393-398.
- Homa, D. (1984). On the nature of categories. *Psychology of Learning and Motivation, 18*, 49-94.
- Hox, J. (1998). Multilevel modeling: when and why. In I. Balderjahn, R. Mather, & M. Schader (Eds.). *Classification, data analysis and data highways* (pp. 147-154). New York: Springer.
- Huang, J., Curran, P., Keeney, J., Poposki, E., DeShon, R. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business Psychology, 27*, 99-114.
- Hu, L.T. and Bentler, P.M. (1999), "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," *Structural Equation Modeling, 6*, 1-55.
- International Personality Item Pool: a scientific collaborator for the development of advanced measures of personality traits and other individual differences Retrieved from <http://ipip.ori.org/>
- Jasper, C. R., & Klassen, M. L. (1990). Perceptions of salespersons' appearance and evaluation of job performance. *Perceptual and Motor Skills, 71*, 563-566.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory-versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
- Judd, C. M., Hawkins, L. J., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology, 89*, 899-913.
- Judge, T. A., & Cable, D. M. (2011). When it comes to pay, do the thin win? The effect of weight on pay for men and women. *Journal of Applied Psychology, 96*, 95-112.
- Kane, T. J. (2003). The long road to race-blindness. *Science, 302*, 571-573.
- Kennedy, D. B., & Homant, R. J. (1984). Personnel managers and the stigmatized employee. *Journal of Employment Counseling, 21*, 89-94.
- Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin, 137*, 616-642.

- Kormanik, M. B. (2009). Sexuality as a diversity factor: an examination of awareness. *Advances in Developing Human Resources, 11*, 24-36.
- Landy, F. J. (2008). Stereotypes, bias, and personnel decisions: strange and stranger. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 379-392.
- Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin, 87*, 72-108.
- Larkin, J. C., & Pines, H. A. (1979). No fat persons need apply: experimental studies of the overweight stereotype and hiring preferences. *Sociology of Work and Occupations, 6*, 312-327.
- Larwood, L. (1995). Attributional effects of equal employment opportunity: theory development at the intersection of EEO policy and management practice. *Group and Organization Management, 20*, 391-408.
- LeBreton, J. M., Wu, J., & Bing, M. N. (2008). The truth(s) on testing for mediation in the social and organizational sciences. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 107-141). New York: Taylor & Francis Group.
- Lee Badgett, M. V. (1995). The wage effects of sexual orientation discrimination. *Industrial and Labor Relations Review, 48*, 726-739.
- Leigh, J. P., & Berger, M. C. (1989). Effects of smoking and being overweight on current earnings. *American Journal of Preventative Medicine, 5*, 8-14.
- Lord, R. G., Brown, D. J., Harvey, J. L., & Hall, R. J. (2001). Contextual constraints on prototype generation and their multilevel consequences for leadership perceptions. *Leadership Quarterly, 12*, 311-338.
- Lord, R. G., De Vader, C., & Alliger, G. (1986). A meta-analysis of the relation between personality traits and leadership perceptions: an application of validity generalization procedures. *Journal of Applied Psychology, 71*, 402-410.

- Lord, R. G., Foti, R. J., & Phillips, J. S. (1982). A theory of leadership categorization. In H. G. Hunt, U. Sekaran, & C. Schriesheim (Eds.), *Leadership: Beyond establishment views* (pp. 104-121). Carbondale: Southern Illinois University Press.
- Lord, R. G., & Maher, K. J. (1991). *Leadership and information processing: Linking perceptions and performance*. Boston: Unwin Hyman.
- Lyness, K. S., & Heilman, M. E. (2006). When fit is fundamental: performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology, 91*, 777-785.
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: thinking categorically about others. *Annual Review of Psychology, 51*, 93-120.
- Macrae, C. N., Stegner, C., & Milne, A. B. (1994). Activating social stereotypes: a functional analysis. *Journal of Experimental Social Psychology, 30*, 370-389.
- Martell, R. F., Lane, D. M., & Emrich, C. (1996). Male-female differences: a computer simulation. *American Psychologist, 51*, 157-158
- Martinko, M. J. & Gardner, W. L. (1983). A methodological review of sex-related access discrimination problems. *Sex Roles, 9*, 825-839.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38*, 379-387.
- Morris, S. (2006). Body mass index and occupational attainment. *Journal of Health Economics, 25*, 347-364.
- Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 87-108.
- Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: response strategies, selective attention, and stimulus generalization. *Learning, Memory, and Cognition, 28*, 924-940.

- Nye, J. L., & Forsyth, D. R. (1991). The effects of prototype-based biases on leadership appraisals: a test of leadership categorization theory. *Small Group Research, 22*, 360-379.
- Olian, J. D., Schwab, D. P., & Haberfeld, Y. (1988). The impact of applicant gender compared to qualifications on hiring recommendations: a meta-analysis of experimental studies. *Organizational Behavior and Human Decision Processes, 41*, 180-195.
- Pichler, S., Varma, A., & Bruce, T. (2010). Heterosexism in employment decisions: the role of job misfit. *Journal of Applied Social Psychology, 40*, 2527-2555.
- Polinko, N. K., & Popovich, P. M. (2001). Evil thoughts but angelic actions: responses to overweight job applicants. *Journal of Applied Social Psychology, 31*, 905-924.
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*, 209-233.
- Puhl, R. M., & Brownell, K. D. (2001). Bias, discrimination, and obesity. *Obesity Research, 9*, 788-805.
- Puhl, R. M., & Heuer, C. A. (2009). The stigma of obesity: a review and update. *Obesity, 17*, 941-964.
- Pulakos, E. D., White, L. A., Oppler, S. H., & Borman, W. C. (1989). Examination of race and sex effects on performance ratings. *Journal of Applied Psychology, 74*, 770-780.
- Register, C. A., & Williams, D. R. (1990). Wage effects of obesity among young workers. *Social Science Quarterly, 71*, 130-141.
- Roehling, M. V. (1999). Weight-based discrimination in employment: psychological and legal aspects. *Personnel Psychology, 52*, 969-1016.
- Roehling, M. V., Pichler, S., & Bruce, T. A. (2013). Moderators of the effect of weight on job-related outcomes: a meta-analysis of experimental studies. *Journal of Applied Social Psychology, 43*, 237-252.
- Rosen, R., & Jerdee, T. H. (1974). Effects of applicant's sex and difficulty of job on evaluations of candidates for managerial positions. *Journal of Applied Psychology, 59*, 511-512.
- Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology, 9*, 283-294.



- Rothblum, E. D. (1992). The stigma of women's weight: social and economic realities. *Feminism and Psychology, 2*, 61-73.
- Rothblum, E. D., Miller, C. T., & Garbutt, B. (1988). Stereotypes of obese female job applicants. *International Journal of Eating Disorders, 7*, 277, 283.
- Roth, P., Bobko, P., McFarland, L., & Buster, M. (2008). Work sample tests in personnel selection: a meta-analysis of black-white differences in overall and exercise scores. *Personnel Psychology, 61*, 637-662.
- Roth, P. L., Purvis, K. L., & Bobko, P. (2012). A meta-analysis of gender group differences for measures of job performances in field studies. *Journal of Management, 38*, 719-739.
- Rudolph, C. W., Wells, C. L., Weller, M. D., & Baltes, B. B. (2009). A meta-analysis of empirical studies of weight-based bias in the workplace. *Journal of Vocational Behavior, 74*, 1-10.
- Shapiro, J. R., King, E. B., & Quinones, M. A. (2007). Expectations of obese trainees: How stigmatized trainee characteristics influence training effectiveness. *Journal of Applied Psychology, 92*, 239-249.
- Schein, V. E. (1973). The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology, 57*, 95-100.
- Schmidt, F. L., & Johnson, R. H. (1973). Effect of race on peer ratings in an industrial setting. *Journal of Applied Psychology, 57*, 237-241.
- Shapiro, J. R., King, E. B., & Quinones, M. A. (2007). Expectations of obese trainees: how stigmatized trainee characteristics influence training effectiveness. *Journal of Applied Psychology, 92*, 239-249.
- Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998). Stereotype efficiency reconsidered: encoding flexibility under cognitive load. *Journal of Personality and Social Psychology, 75*, 589-606.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. New York: Oxford Press.

- Singh, D., & Young, R. K. (1995). Body weight, waist-to-hip ratio, breasts, and hips: role in judgments of female attractiveness and desirability for relationships. *Ethology and Sociobiology*, *16*, 483-507.
- Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: the early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1411-1436.
- Swim, J., Borgida, E., Maruyama, G., & Myers, D. G. (1989). Joan McKay versus John McKay: do gender stereotypes bias evaluations? *Psychological Bulletin*, *105*, 409-429.
- Tunceli, K., Li, K., & Williams, L. K. (2006). Long-term effects of obesity on employment and work limitations among U.S. adults, 1986-1999. *Obesity*, *14*, 1637-1646.
- Turnage, J. J., & Muchinsky, P. M. (1984). A comparison of the predictive validity of assessment center evaluations versus traditional measures in forecasting supervisory job performance: interpretive implications of criterion distortion for the assessment paradigm. *Journal of Applied Psychology*, *69*, 595-602.
- U.S. Bureau of Labor Statistics. (2004). Annual average tables from the January 2004 issue of Employment and Earnings (Table 14: Employed persons by detailed industry and sex, 2004 annual averages). Retrieved from: <http://www.bls.gov/cps/wlf-table14-2005.pdf> on October 16, 2012.
- van Gils, S., van Quaquebeke, N., & van Knippenberg, D. (2010). The x-factor: on the relevance of implicit leadership and followership theories for leader-member exchange agreement. *European Journal of Work and Organizational Psychology*, *19*, 333-363.
- Vanhove, A., & Gordon, R. A. (in press). Weight discrimination in the workplace: a meta-analytic examination of the relationship between weight and work-related outcomes. *Journal of Applied Social Psychology*.
- Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance appraisal: a quantitative review. *Journal of Occupational and Organizational Psychology*, *67*, 189-205.

- Wojciszke, B. (2005). Morality and competence in person and self perception. *European Review of Social Psychology, 16*, 155-188.
- Zebrowitz, L. A. (1996). Physical appearance as a basis of stereotyping. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 79-120). New York: Guilford.
- Zhang, Q., & Wang, Y. (2004). Socioeconomic inequality of obesity in the United States: do gender, age, and ethnicity matter? *Social Science & Medicine, 58*, 1171-1180.

APPENDIX I  
OVERVIEW OF THE TASK

## Study overview

Greetings! For this study you will be playing the role of a decision maker in charge of conducting an initial screening of applicants applying for jobs at a fictitious company. Specifically, you will be evaluating eight different individuals who recently applied for one of two job openings. Some basics are provided below regarding the context of your role as an applicant screener and your task.

**The organization:** The company for which you will be working is StayActive, Inc. StayActive is a healthcare company that provides medical equipment (e.g., motorized wheelchairs, chair lifts, etc.) that allow elderly and disabled individuals to remain active and independent.

StayActive is a company with two different overarching divisions: (a) product sales, and (b) product and customer services.

The sales division is just that – in charge of product sales. Sales personnel are tasked with selling StayActive, both in StayActive stores and on the road, travelling to various healthcare providers and assisted living sites.

The product and customer service division is also just that – in charge of assisting clients in using StayActive products. Service personnel are tasked with handling customer questions and concerns by phone and traveling to the homes of customers to teach them to use the products StayActive provides.

**A Review of Your Task:** StayActive uses a multi-stage hiring process. First, an initial screening is conducted by the human resources department, which includes a basic background check of each applicant's information available through social media. Here a decision is made regarding whether applicants should be interviewed by a hiring manager within the specific department (e.g., sales, customer service) for which they are applying.

For this study, you will be playing the role of a human resources employee conducting the initial screening of applicants for two different job openings. To begin, you will be presented with a description of the first job. It's important that you pay close attention to the job descriptions because afterwards

you'll be asked to describe the ideal applicant for that job. You'll then view the applicants for the first job, answer some questions about each, and finally be asked to rank order the applicants. Importantly, only the top two ranked applicants will move on to the interview stage. You'll then repeat this process for the second job. Finally, you'll be asked to respond to a few brief questions before the study is complete. At that point you will be presented with a code that you will need to enter into the MTurk website to receive compensation for your participation.

APPENDIX II  
JOB DESCRIPTIONS

## **Sales Position**

### ***Tasks***

Interact with potential customers.

Establish business relationships with potential customers.

Understand the product needs of potential customers and use this information to maximize sales.

Self-manage time and selling-related activities.

Negotiate product prices with potential customers.

Work with other salespersons to maximize sales and company profit.

### ***Required Knowledge & Skills***

**Persuasion** — Convince potential customers that StayActive products are right for them.

**Negotiation** – Get customers to agree to the highest possible selling price of StayActive products, providing a greater profit for the company and the salesperson, through own sales commission.

**Quick and Effective Problem Solving** – Use logic and reasoning, for example, to identify what the customer needs and provide the most appropriate StayActive product.

**Judgment and Decision Making** – Consider costs and benefits of business interactions; e.g., providing a greater discount to a new customer who may purchase more StayActive products in the future.



## **Customer Service Position**

### ***Tasks***

Respond by phone to StayActive customer questions, concerns, and complaints regarding StayActive products, and keep accurate records of these interactions.

Make home visits to direct customers in using their new StayActive equipment.

Make home visits and/or check in by phone on existing customers to ensure their satisfaction with the products.

Resolve customers' service or billing complaints by performing activities such as refunding money or adjusting bills.

### ***Required Knowledge & Skills***

**Customer and Personal Service Knowledge** — This includes determining customer needs/concerns, understanding how customers should be treated, and evaluating customer satisfaction.

**Service Orientation** — Actively look for ways to help people; this involves active listening and communication skills.

**Social Perceptiveness** – Being aware of customers' emotions; this includes picking up on customers' tone and emotional state.

**Problem Sensitivity** – The ability to tell when something is wrong or is likely to go wrong. This involves recognizing and understanding customer concerns, and being sensitive to customer needs.

## APPENDIX III

### PROTOTYPE AND PERCEPTIONS OF APPLICANT MEASURES

### **Open-ended Prototype Description.**

*Directions as stated to participants:* Please provide a description of a person who you believe to be an ideal applicant for this job.

[Text box for entering response]

### **Prototype and Perceptions of Applicant Ratings**

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content:

Competence and warmth follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878-902.

*Directions as stated with regard to participants' prototype ratings:* Please indicate how important each of the following descriptors is to the employee you just described.

*Directions as stated with regard to participants' perceptions of the target applicant:* Please rate the applicant on each of the following characteristics:

(1 = "not at all"; 7 = "extremely")

Competent

Confident

Competitive

Independent

Intelligent

Friendly

Warm

Good-natured

Sincere

### **Open-ended Perceptions of Applicant**

*Directions as stated to participants:* Please provide any additional thoughts on your perceptions of the applicant, not captured by the ratings you just made.

[Textbox for entering response]

APPENDIX IV  
STIMULUS MATERIALS

## Normal-weight Applicant Stimuli



www.shutterstock.com - 43885225



www.shutterstock.com - 3215819



www.shutterstock.com - 101761816



k1540866 www.fotosearch.com

## Overweight Applicant Stimuli



freedigitalphotos.net



www.shutterstock.com - 23000287



k10096299 fotosearch.com



www.shutterstock.com - 44739991



www.shutterstock.com · 3215819

## Taylor Smith

457 8<sup>th</sup> St.

Little Rock, CO, 76548

(970)-555-5968

Tsmith@charter.com

### Education

---

**Bachelor's degree, Business Management**

**Metropolitan State University, 2013 (expected cumulative GPA: 3.0)**

### Experience

---

**Business owner, Amway**

(August 2010-Present)

**Part-time customer service, Johnson's Auto Sales (car dealership)**

(September 2009-Present)

### Activities

---

**Student President, Future Business Leaders of America (FBLA), Wofford University chapter**

(September 2012-Present)

APPENDIX V  
CONTROL MEASURES



### **Conscientiousness scale in Big Five Inventory**

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory-versions 4a and 5a*.

Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

(1 = “disagree strongly” and 7 = “agree strongly”):

*Directions as stated to participants:* please indicate the extent to which you agree with each of the following statements.

I see myself as...

Someone who does a thorough job.

Can be somewhat careless.

Is a reliable worker.

Tends to be disorganized.

Tends to be lazy.

Perseveres until the task is finished.

Does thing efficiently.

Makes plans and follows through with them.

Is easily distracted.

### **Demographics/Background Questions**

(1 = “yes”; 0 = “no”)

*Directions as stated to participants:* please respond to each of the following.

What is your gender?

Do you have any previous professional experience...

In evaluating job applicants?

In sales?

In customer service?

In the healthcare or pharmaceutical industry?

APPENDIX VI  
CRITERION MEASURES

### **Criterion Questions Presented after Viewing Each Applicant**

*Directions as stated to participants:* please respond to the following questions in reference to the applicant you just viewed.

(1=Not at all, 7=Extremely):

#### *Work-related "Interpersonal" Factor*

To what extent...

Would you, personally, want to work with this person on a daily basis?

Would you be concerned that this person would make your job harder? (reverse)

Would you expect this person to be a valuable coworker?

Do you think this person's performance would cause problems for you and your coworkers, if hired? (reverse)

#### *Work-related "Success" Factor*

To what extent do you think this person...

Would be successful in this job?

Will thrive in this job?

Is capable of handling this job?

Would fail if hired for this job? (reverse)