Gender and Letters of Recommendation for Academia: Agentic and Communal Differences

Juan M. Madera
University of Houston

Michelle R. Hebl and Randi C. Martin
Rice University

In 2 studies that draw from the social role theory of sex differences (A. H. Eagly, W. Wood, & A. B. Diekman, 2000), the authors investigated differences in agentic and communal characteristics in letters of recommendation for men and women for academic positions and whether such differences influenced selection decisions in academia. The results supported the hypotheses, indicating (a) that women were described as more communal and less agentic than men (Study 1) and (b) that communal characteristics have a negative relationship with hiring decisions in academia that are based on letters of recommendation (Study 2). Such results are particularly important because letters of recommendation continue to be heavily weighted and commonly used selection tools (R. D. Arvey & T. E. Campion, 1982; R. M. Guion, 1998), particularly in academia (E. P. Sheehan, T. M. McDevitt, & H. C. Ross, 1998).

Keywords: gender stereotypes, letters of recommendation, academia, social role theory

The problem of pipeline shrinkage for women in academia is a well-known and researched phenomenon (Bellas & Toutkoushian, 1999; Camp, 1997; Olsen, Maple, & Stage, 1995; Taylor, 2007; Windall, 1988). This phenomenon refers to the fact that women enter graduate school at about the same rate as do men, but women are less likely to enter and succeed in academia at the same rate as their male counterparts, particularly in science and engineering disciplines. In fact, the National Science Foundation (2008) has reported that women comprise about 29% of science and engineering faculty at 4-year colleges and universities and comprise only 18% of full professors. One contributing factor to this gender disparity may be gender differences in letters of recommendation. In particular, there is little research that addresses whether letters of recommendation for academia are written differently for men and women and whether potential differences influence selection decisions in academia. The present study addresses this issue.

The focus on letters of recommendation is justified because they are an important and commonly used selection tool that provides information on applicants' past performance and qualifications (McCarthy & Goffin, 2001), confirms or supplements information

provided by applicants (Brem, Lampman, & Johnson, 1995), and describes applicants' motivation (Tommasi, Williams, & Nordstrom, 1998). In fact, Cascio and Aguinis (2004) stated, "the fact is, decisions are made on the basis of letters of recommendations" (p. 278). In particular, they have been found to be among the most important criterion used to screen and evaluate applicants for internships (Lopez, Oehlert, & Moberly, 1996), graduate programs (Landrum, Jeglum, & Cashin, 1994), medical schools (Johnson et al., 1998), military training programs (McCarthy & Goffin, 2001), and psychology faculty positions (Sheehan, McDevitt, & Ross, 1998).

Regardless of the reasons for using letters of recommendation, research has shown that letters of recommendation can be written differently for women than for men (McCarthy & Goffin, 2001; Trix & Psenka, 2003). In qualitative studies, researchers have reported that letters of recommendation for college (LaCroix, 1985) and graduate school (Watson, 1987) contained stereotypical gender-related words and phrases, describing female applicants as feminine and male applicants as masculine. Using discourse analysis, Trix and Psenka (2003) analyzed over 300 letters of recommendation for doctors applying for medical faculty positions and found that letters were longer for male than female applicants. In addition, letters for men contained more standout adjectives, such as *superb, outstanding*, and *remarkable,* and contained more research-related descriptors than did letters for women. In a study that replicated Trix and Psenka's (2003) study, Schmader, Whitehead, and Wysocki (2007) examined letters of recommendation for science faculty positions and found that letters for male compared to female applicants also contained more standout adjectives. However, there were no statistical differences in length, positive and negative language, and research- and teaching-related words for male and female applicants.

There are, however, important limitations to the previously conducted research studies. First, many of these findings were based on descriptive rather than on inferential statistics. Second, many studies used subjective rather than objective means for

scoring gender differences, with the authors of the study carrying out the scoring rather than relying on more objective alternatives (i.e., the Linguistic Inquiry and Word Count [LIWC] program; Pennebaker, Francis, & Booth, 2001). Third, many studies did not use statistical procedures (e.g., hierarchical linear modeling) that address the fact that letters of recommendation are nested within applicants. Fourth, there were important variables that were not controlled for in these previous studies that might have affected the results. Fifth, these prior studies did not examine whether gender differences in letters actually affected judgments about hireability.

Thus, to improve on the methodology of these earlier studies, we examined gender differences in letters of recommendation with objective methods (i.e., language content analysis; Pennebaker et al., 2001), with statistical procedures appropriate for nested data, including indicators of productivity as control variables (e.g., publications, teaching experience, and honors), and by assessing the effects of gender differences on judgments of hireability. We also situated our examination of potential differences within contemporary theorizing about gender bias, specifically drawing on the social role theory of sex differences (Eagly, Wood, & Diekman, 2000).

Gender Stereotypes: Agentic and Communal Characteristics

According to social role theory, behavioral sex differences arise from the division of labor—the differential social roles inhabited by women and men (Eagly et al., 2000). Historically, men have been more likely to engage in tasks that require speed, strength, and the ability to be away from home for expanded periods of time, whereas women were more likely to stay home and engage in family tasks, such as child rearing. Accordingly, men are perceived and expected to be agentic, and women are perceived and expected to be communal. Agency includes descriptions of aggressiveness, assertiveness, independence, and self-confidence (Eagly & Johannesen-Schmidt, 2001). Agentic behaviors at work include speaking assertively, influencing others, and initiating tasks. Communal behaviors at work include being concerned with the welfare of others (i.e., descriptions of kindness, sympathy, sensitivity, and nurturance), helping others, accepting others' direction, and maintaining relationships (Eagly & Johannesen-Schmidt, 2001).

Central to understanding gender stereotypes is that they are both descriptive and prescriptive (Burgess & Borgida, 1999; Rudman & Glick, 2001). Stereotypes suggest not only how men and women do behave (i.e., descriptive), but also how men and women should behave (i.e., prescriptive). The prescriptive nature of gender stereotypes also specifies what women should not do—often leading to penalties for women who do not conform to their respective norm. As a result, women are expected to engage in a feminine gender role that reflects communal qualities but not agentic ones (Wood & Eagly, 2002). The descriptive and prescriptive nature of these stereotypes can affect women's entrance and mobility in certain jobs. For example, agency has been found to be associated with roles of leadership (Eagly & Karau, 2002; Heilman, Block, & Martell, 1995). Managerial and executive level jobs are usually considered to be a masculine role—thought to require agentic qualities, such as ambition, aggressiveness, and achievement. Therefore, attitudes are often less positive toward female than male leaders, and it is more difficult for women to become and succeed

as leaders (Eagly & Karau, 2002). This occurs because women are perceived and expected to be communal, but leaders are perceived and expected to be agentic. It is this *lack of fit,* or mismatch, between attributes of gender or their prescriptive components of shoulds and work roles that can affect women in the workplace (Heilman, 2001). Thus, it is important to examine whether women are being described in letters of recommendation as less agentic and more communal than men because agency is related to higher status and success in the workplace.

Study 1

Overview and Hypotheses

To examine whether social role theory might explain gender differences in letters of recommendation, we analyzed letters of recommendation written for applicants for faculty positions in a psychology department at a Research I university (as designated by the Carnegie Classification of Institutes of Higher Education). According to social role theory (Eagly et al., 2000), because men are usually ascribed agentic characteristics, we expected that men would be described in more agentic terms than would women in letters of recommendation and that women would be described in more communal terms than would men.

> *Hypothesis 1.* Women are more likely than men to be described in communal terms in letters of recommendation.

> *Hypothesis 2.* Men are more likely than women to be described in agentic terms in letters of recommendation.

Method

Sample

The sample consisted of 624 letters of recommendation and 194 applicants for eight junior faculty positions from 1998 to 2006 at a Southern university in the United States. Of those whose sex could be identified, 46% ($n = 89$) of the applicants were women, and 54% ($n = 103$) were men; 30% ($n = 193$) of the recommenders were women, and 70% ($n = 477$) were men.[1] Applicants' ages ranged from 26 to 40 years, with a mean of 32 ($SD = 3.69$). The mean number of letters per applicant was 3.23.

Procedure

This study used archival data obtained from the psychology department. After receiving institutional review board approval, we transcribed the original letters of recommendation to electronic form (Microsoft Word document) and used a computer text analysis program, the LIWC program (Pennebaker et al., 2001). The LIWC program analyzes text files and computes the percentage of words from a file that fall into each of 74 possible linguistic

---

[1] The faculty positions were in the follow areas: applied experimental ($n = 2$), applied psychology ($n = 49$), cognitive ($n = 22$), cognitive/neuroscience ($n = 43$), cognitive/neuroscience developmental ($n = 6$), health ($n = 32$), industrial/organizational ($n = 37$), and social ($n = 3$). Male recommenders wrote 262 letters for male applicants and 194 letters for female applicants; female recommenders wrote 78 letters for male applicants and 109 letters for female applicants.

categories, such as negative emotion, self-reflection, causation, and physical issues. The program dictionary is composed of 2,300 words and word stems and was developed with emotion rating scales (e.g., the Positive Affect Negative Affect Scale; Watson, Clark, & Tellegen, 1988). The LIWC dictionary was validated by having judges rate the content of hundreds of text files, comparing their results to those of the computer program (Pennebaker & Francis, 1996; also see Pennebaker & King, 1999).[2]

Measures

Communal adjectives. We created a dictionary for communal adjectives derived from Eagly's work in communal and agentic characteristics (Eagly & Johannesen-Schmidt, 2001; Eagly & Karau, 2002; Eagly et al., 2000; Wood & Eagly, 2002).[3] The final list included terms such as *affectionate, helpful, kind, sympathetic, sensitive, nurturing, agreeable, tactful, interpersonal, warm, caring,* and *tactful.* The average percentage of words for communal adjectives in each letter of recommendation was 0.69%.

Social–communal orientation. The social orientation index in LIWC counts the number of words that deal with other people. Psychologically, it reflects how much letter writers referred to other people when writing about the applicant. Words in this category include *husband, wife, kids, babies, brothers, children, colleagues, dad, family, they, him,* and *her.* The average percentage for social–communal words in each letter of recommendation was 8.22%.

Agentic adjectives. We created a dictionary for agentic adjectives. The list of words was also derived from Eagly's work on communal and agentic characteristics (Eagly & Johannesen-Schmidt, 2001; Eagly & Karau, 2002; Eagly et al., 2000; Wood & Eagly, 2002). This list includes *assertive, confident, aggressive, ambitious, dominant, forceful, independent, daring, outspoken,* and *intellectual.* The average percentage for communal adjectives in each letter of recommendation was 0.81%.

Agentic orientation. The cognitive mechanism, motion, and achieve indexes in LIWC count the number of words that deal with other peoples' cognitive processes, achievements, and actions. Psychologically, these indexes reflect how much letter writers referred to the applicants as active, dynamic, and achievers. As such, we constructed a single category by using a composite score of the three indexes. Words in this category included *earn, gain, do, know, insight,* and *think.* The average percentage for agentic orientation words in the letters of recommendation was 7.42%.

Gender. Gender for both the applicants and the recommenders was coded female (1) or male (2).

Control variables. We used seven control variables. These were the number of years in graduate school, the number of total publications, the number of first author publications, the number of honors, the number of postdoctoral years of education, the position applied for, and the number of courses taught.

Results

Given that the letters of recommendation were nested within applicants, we used the HLM 6 program (Raudenbush, Bryk, Cheong, & Congdon, 2004) to analyze the data by conducting hierarchical linear modeling. We used full maximum likelihood estimation procedures and included random effects. Hypothesis

testing involved three steps: (a) the control variables were entered into the model, (b) the main effects variables (applicant gender and letter writer gender) were entered in the equation, and (c) the interaction variable (Applicant Gender × Letter Writer Gender) was entered into the model. Thus, for the analyses, the intercepts of the Level 1 variables, communal adjectives, social–communal index, agentic adjectives, and agentic orientation index, were predicted by the Level 2 variable, gender of the applicant.[4] For exploratory reasons, we also included the gender of the letter writer. Descriptive statistics and intercorrelations for the control, independent, and dependent variables are reported in Table 1.

Hypothesis 1

As shown in Table 2, gender of applicant significantly predicted both communal adjectives ($\beta = -.16$, $p < .05$) and the social–communal index ($\beta = -.12$, $p < .05$). More specifically, the results show that women were described by more communal terms than were men. In addition, letters written for women mentioned more social–communal terms than letters for men. Thus, the results supported Hypothesis 1. The gender of the letter writer significantly predicted the social–communal index ($\beta = .11$, $p < .05$), but not the communal adjectives ($\beta = .08$, $p < .05$). The interaction between gender of the applicant and gender of the letter

---

[2] Despite the published evidence for the validity of the LIWC program (e.g., Pennebaker & Francis, 1996; Pennebaker et al., 2001; Pennebaker & King, 1999), we further validated our LIWC measures with three coders (who were unaware of the hypotheses) who rated the extent to which letter writers described the applicants as communal and agentic. More specifically, the coders used a 9-point scale (1 = *not at all* to 9 = *very much*) to respond to the following two questions: "To what extend was the applicant described as agentic (assertive, independent, aggressive)?" and "To what extend was the applicant described as communal (kind, nurturing, caring)?" The coders were provided with the definition and examples of agency and communal descriptors. The coders' ratings of agentic description were significantly related to agentic adjectives ($r = .17$, $p < .05$) and agentic orientation ($r = .12$, $p < .05$). Similarly, the coders' ratings of communal description were significantly related to communal adjectives ($r = .19$, $p < .05$) and social–communal orientation ($r = .12$, $p < .05$).

[3] We also used the program WordNet (http://wordnet.princeton.edu/) to search for synonyms of the word in the list. WordNet is an online lexical reference system, designed with psycholinguistic theories of human lexical memory, which organizes English nouns, verbs, adjectives, and adverbs into synonym sets, each representing one underlying lexical concept. We used it for both the Communal Adjectives and Agentic Adjectives measures.

[4] Analysis for each outcome variable (i.e., the communal and agentic measures) begins with fitting an unconditional model (i.e., with no predictors) to estimate or partition the variance components. This essentially tests the within-applicant and between-applicant variance. If no within-applicant variance exists in the dependent variables, then hierarchical linear modeling is not appropriate because there is only one level of analysis (i.e., no within-applicant variance). The results of the unconditional (null) models indicated that there was significant between-applicant variance in each dependent variable ($p < .01$ for all variables) and that a substantial proportion, $P = r^2/(r^2 + t_{00})$, of the total variance in these dependent variables was within applicants. That is, 26% of communal adjectives, 23% of social–communal, 20% of agentic adjectives, and 10% of agentic orientation variance was within applicant. Thus, there is substantial variance between and within applicants that warrants the use of hierarchical linear modeling to examine Level 1 and Level 2 dependent and independent variables.

Table 1
Means, Standard Deviations, and Intercorrelations for Study 1

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Gender of applicant	1.51	0.50	—												
2. Gender of writer	1.71	0.28	.23*	—											
3. Communal adjectives	0.0069	0.30	−.16*	−.13	—										
4. Social–communal	0.082	1.32	−.28*	−.29*	.27*	—									
5. Agentic adjectives	0.0081	0.29	.11	−.07	.04	.03	—								
6. Agentic orientation	0.074	1.04	.11	.19*	.01	−.05	.23*	—							
7. Years in graduate school	4.17	2.0	.10	.11	−.10	−.11	−.11	−.06	—						
8. No. of publications	4.24	3.5	.09	.10	−.11	−.33*	−.11	−.01	.06	—					
9. No. of first-author publications	1.93	2.2	.13	.10	−.05	−.35*	−.11	−.01	.01	.75*	—				
10. No. of honors	0.91	1.4	−.10	−.10	−.02	.05	.01	.15*	−.01	.05	.06	—			
11. Postdoctoral years	1.10	1.5	.15*	−.08	−.11	−.24*	−.15*	−.01	.05	.39*	.44*	.18*	—		
12. No. of courses taught	5.45	3.3	−.09	−.01	.11	.18*	−.01	−.02	.12	−.10	−.11	−.03	−.02	—	
13. Position applied	3.38	1.7	.01	.09	.10	.02	.02	.16*	−.01	.07	.03	.09	.09	.08	—

Note. Gender was coded as female = 1, male = 2. Level 1 variables were aggregated: communal adjectives, social–communal, agentic orientation, and agentic adjectives.
* $p < .05$.

writer was not significant for communal adjectives or social–communal orientation.

Hypothesis 2

As shown in Table 2, the results revealed that gender of applicant significantly predicted agentic adjectives ($\beta = .16$, $p < .05$); men were described by more agentic terms than were women. However, gender of applicant was not related to the agentic orientation index ($\beta = .04$, $p > .05$). Thus, Hypothesis 2 was partially supported. The gender of the letter writer was not a significant predictor of agentic adjectives and agentic orientation. As depicted in Figure 1, there was a significant interaction of gender of the applicant and gender of the writer on agentic orientation ($\beta = .12$, $p < .05$). Analyses revealed that for female applicants, male writers used more agentic orientation terms than did female writers ($\beta = 3.02$, $p < .05$), whereas the pattern for male applicants was the opposite but was not significant ($\beta = .63$, $p > .05$).

Discussion

With respect to social role theory, our results confirmed our hypotheses, demonstrating that female applicants were more likely to be described with communal terms (e.g., *affectionate, warm,* *kind,* and *nurturing*) than male applicants. Letters of recommendation for female applicants also mentioned more social–communal terms, such as *student(s), child, relative,* and *mother.* In contrast, male applicants were more likely to be described in agentic terms (e.g., *ambitious, dominant,* and *self-confident*) than were female applicants. Thus, we found support for our hypotheses. It is important to note that these differences were obtained even though we included objective measures of performance from applicants' curriculum vitae. That is, unlike past research (e.g., LaCroix, 1985; Trix & Penska, 2003; Watson, 1987), the current study included productivity factors, such as the number of publications, teaching experience, postdoctoral years, and honors. Such factors could affect the quality of letters of recommendation and the use of agentic and communal descriptions.

The interaction between applicant gender and letter writer gender showed an interesting pattern; for female applicants, male writers used more agentic orientation terms than did female writers. Although we did not hypothesize an interaction, it could be the case that men are more likely to emphasize agency than women when writing letters of recommendation. An alternative explanation could be that women focus more on communality and deemphasize the agency of women. This idea supports the universality of gender norms and the prescriptive stereotypes of agency and

Table 2
Hierarchical Linear Modeling Results With Applicant Gender, Writer Gender, and Their Interaction as Predictors

Dependent variable	Applicant gender			Writer gender			Interaction		
	Estimate	SE	t	Estimate	SE	t	Estimate	SE	t
Communal adjectives	**−.16**	.19	−3.02*	.08	.03	−1.44	.11	.03	1.71
Social–communal orientation	**−.12**	.05	−1.98*	**.11**	.09	−2.10*	.10	.09	1.16
Agentic adjectives	**.16**	.06	2.46*	.02	.03	−0.77	.06	.03	−0.15
Agentic orientation	.04	.18	0.64	.04	.08	0.67	**.12**	.08	−2.26*

Note. Applicant gender was coded as female = 1, male = 2. Results are after controlling for the number of years in graduate school, the number of total publications, the number of first-author publications, the number of honors, the number of postdoctoral years, the applied position, and the number of courses taught. Values in bold are statistically significant.
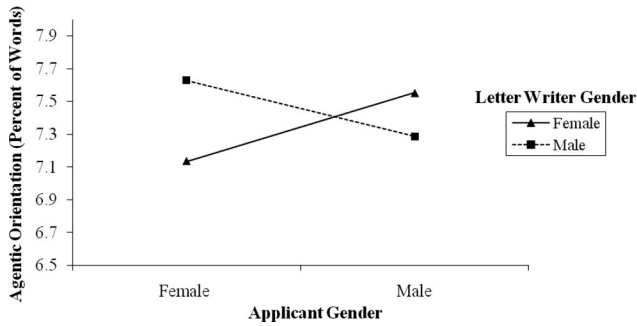* $p < .05$.

Figure 1. Applicant Gender × Letter Writer Gender interaction on agentic orientation from Study 1.

communion across both men and women, which is consistent with the broader literature on stigma (Crocker, Major, & Steele, 1998; Goffman, 1963; Hebl, Tickle, & Heatherton, 2000) and, more specifically, with the literature on sex differences (e.g., see Heilman & Okimoto, 2007; Heilman, Wallen, Fuchs, & Tamkins, 2004).

One question that emerges is whether the behavior of men and women differs along agentic and communal lines, as suggested by the frequency with which these terms were used, or whether the frequency of these terms reflects the writers' perception of the individuals colored by social role stereotyping or perhaps even by what they felt was appropriate for describing the individual. That is, a letter writer might perceive a male and female applicant to be equal in warmth, kindness, and helpfulness but may feel more comfortable about commenting on all of these things for the female than for the male applicant. Our data cannot speak directly to these issues.[5]

Study 2

Overview and Hypotheses

Study 1 showed that there are gender differences in communal and agentic characteristics in letters of recommendation. Therefore, it is important to examine whether differences in agentic and communal characteristics influence hiring decisions in academia. Central to understanding how agentic and communal characteristics are related to hiring decisions in academia is the gender typing of occupations as male or female on the basis of (a) job responsibilities believed to be gender linked or (b) the sex of the usual job holder (Heilman, 1995; Lyness & Heilman, 2006). Agency has been related to leadership and high-status occupations, such as academia, more than communal characteristics have (Eagly & Johannesen-Schmidt, 2001; Eagly & Karau, 2002; Heilman, 2001). Academia remains a male-dominated occupation (Bellas & Toutkoushian, 1999; Camp, 1997; Olsen et al., 1995; Windall, 1988). Thus, academic positions for research-oriented universities can be perceived to be masculine, and as a result, agentic characteristics might be positively related to hiring ratings. In contrast, communal characteristics might be negatively related to hiring ratings. As such, we predicted the following:

> *Hypothesis 1.* Agentic characteristics included in letters of recommendation will be positively related to hiring ratings, but communal characteristics will be negatively related to hiring ratings.

Heilman's research (Heilman, 1995, 2001; Heilman & Okimoto, 2007; Heilman et al., 2004) demonstrated that communal characteristics can hinder women in the workplace, because women are expected to be communal. However, when women are perceived to be communal, they can be evaluated negatively when performing in occupations in which agency is perceived to be important. With reference to the present study, this body of work led to the prediction that communal characteristics in letters of recommendation would be negatively related to hiring ratings, and because women are more likely to be described with communal characteristics, women would be rated as less hireable than men for academic positions at a research university. Conversely, men would be rated as more hireable than women, because men are more likely than women to be described with agentic characteristics, which would be positively related to hiring ratings. According to this reasoning, gender differences in hiring ratings for academia would be mediated by gender differences in communal and agentic characteristics in letters of recommendation. More formally, we predicted the following:

> *Hypothesis 2.* Men will be rated as more hireable than women for academic positions at a research university.

> *Hypothesis 3.* Communal and agentic characteristics will mediate the relationship between gender and hireability in selection decisions for academia at a research university.

Method

Participants and Procedure

Six psychology professors served as subject matter experts (SMEs). The SMEs were three professors in industrial–organizational psychology and three in cognitive psychology. After receiving institutional review board approval, the SMEs were provided with the letters of recommendation from Study 1 and were instructed to rate each applicant on hireability (i.e., hiring decisions in academia based on letters of recommendation). The SMEs were provided with a random sample of 100 letters of recommendation. They were provided with an additional 25 letters that were the same across raters to get an estimate of interrater reliability. In other words, SMEs reviewed 100 unique letters and 25 letters that all six SMEs reviewed. The letters of recommendation were modified by removing names of applicants and recommenders, names of schools, and the gender of the applicants and recommenders (i.e., gender identifying information was replaced by *he/she, his/her,* and *him/her*). The SMEs were instructed, "imagine that you are reviewing and rating applicants for a tenure-track assistant professor position in the psychology department" and asked to "rate the applicant based on this letter of recommendation."

---

[5] Given that we controlled for aspects of the curriculum vita that indicated research productivity, at least the agentic behaviors that underlie this productivity could not have been the source of the discrepancy in the use of the agentic terms. On the social–communal side, we had less in the way of objective measures that might have related to these terms; thus, we had fewer grounds for distinguishing among behavior, perception, or the letter writers' conformity to social norms as the basis of the differences.

Measures

Agency and communion. The scales from Study 1 were used to measure agentic and communal characteristics: Communal Adjectives, Social–Communal Orientation, Agentic Adjectives, and Agentic Orientation. For the sake of simplicity, we developed composites of the agentic and communal measures by standardizing the scales and taking their respective means. Thus, for Study 2, we had an agentic composite (mean of the Agentic Adjectives and Agentic Orientation scales) and a communal composite (mean of the Communal Adjectives and Social–Communal Orientation scales).[6]

Gender. Participant gender was coded female (1) or male (2).

Hireability. Participants rated the hireability of the applicants with a 9-point scale (1 = *not at all,* 9 = *very much*) on four items (see the Appendix). The intraclass correlation coefficient (ICC1; interrater reliability) was .98 for the 25 overlapping letters. The intraclass correlation coefficient (ICC2; group mean reliability) was .97, and the alpha coefficient for the measure was .99. Thus, there was sufficient evidence for the reliability of the raters and the scale.

Control variables. The same control variables from Study 1 were used. We also included the gender of the letter writer as a control variable because the results from Study 1 showed some gender differences in letter writers.

Results

Table 3 shows the means, standard deviations, and correlations for the control variables, applicant gender, letter writer gender, the communal composite, the agentic composite, and the outcome variable—hireability. To test Hypotheses 1 and 2, we conducted hierarchical linear modeling with the HLM 6 program (Raudenbush et al., 2004). We used full maximum likelihood estimation procedures and included random effects. For this hierarchical linear modeling analysis, the control variables were entered in the first step, and the communal composite, agentic composite, and applicant gender were entered in the second step.[7]

Table 4 shows the regression results. Supporting Hypothesis 1, communal characteristics were negatively related to hireability ($\beta = -.28$, $p < .05$), indicating that a greater proportion of communal characteristics in the letters of recommendation was related to lower ratings of hireability. However, the proportion of agentic characteristics was not significantly related to hireability ($\beta = .09$, $p > .05$). Thus, Hypothesis 1 was partially supported. Gender of the applicant was not related to ratings of hireability ($\beta = .04$, $p > .05$), not supporting Hypothesis 2.

Because the regression analysis showed that gender and the hiring ratings were not related, we could not use the traditional Baron and Kenny (1986) mediation analysis to test Hypothesis 3. However, James, Mulaik, and Brett (2006) proposed structural equation modeling or path analysis as an alternative to test mediation. This approach does not require the distal variable to correlate with the outcome variable (i.e., gender and hireability in the current study). In fact, scholars have questioned whether it is necessary to provide evidence for a significant path from the distal variable to the outcome variable to establish mediation (Collins, Graham, & Flaherty, 1998; James et al., 2006; MacKinnon, 2000; MacKinnon, Krull, & Lockwood, 2000; Shrout & Bolger, 2002).

Rather, the simultaneous test of the significance of both the path from the distal variable to a mediator and the path from the mediator to the outcome variable (i.e., the structural equation modeling approach) provides, relative to other approaches (e.g., Baron & Kenny's, 1986, steps), the best balance of Type I error rates and statistical power (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). In light of this literature, we used the structural equation modeling/path analysis approach for testing the mediation hypothesis.

We used path analysis with Mplus (Muthen & Muthen, 1998) to test Hypothesis 3—that communal and agentic characteristics would mediate the relationship between gender and hireability in selection decisions for academia. To determine the adequacy of model fit, we used four fit indices: (a) chi-square and degrees of freedom, (b) the comparative fit index (CFI), (c) the incremental index of fit (IFI), and (d) the root mean square error of approximation (RMSEA). It is suggested that good fit indices for CFI and IFI are greater than .90 and that a good fit index for RMSEA is less than .08 (Byrne, 2001; Steiger, 1990; Vandenberg & Lance, 2000). The model we tested included gender as the distal predictor, with the communal and agentic composites as the mediators of the gender–hireability relationship and with the control variables in the model. This model demonstrated adequate fit, $\chi^2(54) = 52.66$, $p > .05$; CFI = .97; IFI = .98; RMSEA = .044 (see Figure 2). Women were described as more communal than men, $\beta = -.28$, $p < .05$, and men were described as more agentic than women, $\beta = .19$, $p < .05$. The communal composite was negatively related to hireability ($\beta = -.26$, $p < .05$). The agentic composite was not significantly related to hireability ($\beta = .16$, $p > .05$). The variables in the model explained 27% of the variance in hireability ratings.[8] Thus, the results partially supported Hypothesis 3.

Discussion

The results of Study 2 revealed that communal characteristics were negatively related to hireability ratings and that the communal ratings mediated the relationship between applicant gender and hireability ratings for a research-oriented university. The results for agency, however, were not as clear. Although we expected a positive relationship between hireability and agency, the results did not show a significant effect. It might be the case that agency is expected from applicants in academia, and therefore the agentic characteristics did not have an impact on hireability. It also might

---

[6] Bivariate correlations showed that the two composites had stronger relationships with the outcome variable—hireability—than the four scales, suggesting that the composites might be more appropriate than the individual scales.

[7] The results of the unconditional (null) models indicated that there was significant between-applicant variance in the hireability dependent variable ($\chi^2 = 340.39$ $p < .01$) and that a substantial proportion, $P = r^2/(r^2 + t_{00})$, of the total variance in hireability was within applicants; that is, 22% of hireability variance was within applicant.

[8] Applicant gender explained 7% and 3% of the communal and agentic composites, respectively. As recommended by MacKinnon et al. (2002), we used the Sobel test to examine the significance of the indirect effect of gender and communal characteristics on hireability. The test indicated that gender had an indirect effect on hireability through its direct effect on the communal composite ($z = 3.41$, $p < .05$).

Table 3

Means, Standard Deviations, and Intercorrelations for Study 2

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
1. Years in graduate school	4.17	2.0	—											
2. No. of publications	4.24	3.5	.06	—										
3. No. of first-author publications	1.93	2.2	.01	.75*	—									
4. No. of honors	0.91	1.4	−.01	.05	.06	—								
5. Postdoctoral years	1.10	1.5	.05	.39*	.44*	.18*	—							
6. No. of courses taught	5.54	3.3	.12	−.10	−.11	−.03	−.02	—						
7. Position applied	3.38	1.7	−.01	.07	.03	.09	.09	.08	—					
8. Gender of letter writer	1.71	0.28	.11	.10	.10	−.10	−.08	−.01	.09	—				
9. Gender of applicant	1.51	0.50	.10	.09	.13	−.10	.15*	−.09	.01	.01	—			
10. Agentic composite	−0.003	0.77	−.11	−.07	−.09	.11	−.10	−.02	.11	.09	.14*	—		
11. Communal composite	−0.001	0.75	−.14	−.28*	−.26*	.02	−.24*	.19*	.19*	−.27*	−.27*	.03	—	
12. Hireability	5.85	1.4	−.07	.46*	.38*	.07	.10	−.02	.18*	.10	.09	.02	−.29*	—

Note. Gender was coded as female = 1, male = 2. Level 1 variables were aggregated: communal adjectives, social–communal, agentic orientation, letter length, grindstone, research terms, teaching terms, doubt raisers, negative intensifiers, and positive intensifiers.
* $p < .05$.

be the case that the number of publications is a better indicator than the adjectives and descriptions provided by the letter writers. Correlations showed that the number of publications was significantly related to hireability. Although the raters did not have access to curriculum vitae, letter writers might have mentioned the total number of publications or discussed papers that might be under review or other papers being prepared for publications. If so, then the raters may have relied on this more direct information about productivity than on agency descriptors, thus accounting for the lack of an effect of agency on hireability decisions. The communal characteristics used by letter writers might be a better indicator of communality than other possible indicators in letters (i.e., the control variables from Study 1 and Study 2), because communion infers interpersonal information about exhibiting nurturing and socially sensitive attributes, which is information not necessarily conveyed in a vita.

Table 4

Hierarchical Linear Modeling Predicting Hireability

	Model 1		Model 2		Pseudo R^2
Variable	Estimate	SE	Estimate	SE	
Years in graduate school	−.07	.04	−.10	.05	
No. of publications	.25*	.03	.21*	.03	
No. of first-author publications	.03	.07	.10	.07	
No. of honors	.05	.03	.03	.07	
Postdoctoral years	−.03	.04	−.09	.08	
No. of courses taught	.05	.01	.09	.03	
Gender of letter writer	.03	.10	−.05	.14	
Position applied (block)					.07
Gender of applicant			.04	.12	
Agentic attribution composite			.09	.11	
Communal attribution composite			−.28*	.12	

Note. Gender was coded as female = 1, male = 2. Standardized estimates and the standard errors are reported.
* $p < .05$.

General Discussion

The studies presented in the current article replicate and extend past research by showing (a) that there are gender differences in letters of recommendation—women are described as more communal and less agentic than are men (Study 1)—and (b) that communal characteristics have a negative relationship with hiring decisions in academia (Study 2). These results can be understood within the social role theory framework (Eagly et al., 2000). The data suggest that female applicants are described in accordance with communal gender norms, which are both descriptive and prescriptive (Eagly et al., 2000; Heilman et al., 1995). In addition, the results suggest that there is a lack of fit between the attributes of communality and the work role of academia. Such findings are particularly important because letters of recommendation are important and commonly used selection tools (Cascio & Aguinis, 2004; Sheehan et al., 1998).

This research not only has important implications for women in academia but also for women in management and leadership roles. A large body of research suggests that communality is not perceived to be congruent with leadership and managerial jobs (e.g., Eagly & Johannesen-Schmidt, 2001; Eagly & Karau, 2002; Heilman, 2001). In particular, Heilman's (2001) lack of fit model, suggests that "fit-derived performance expectations, whether positive or negative, can profoundly affect evaluation processes" (p. 660). Thus, for occupations in which agency is linked to success or perceived as more important than communality, the perception of
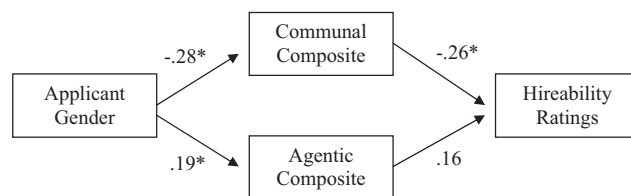

Figure 2. Communal and agentic composites as the mediators of the gender–hireability relationship from Study 2. Standardized estimates are shown. Results are with control variables in the model. * $p < .05$.

lack of fit between a female applicant and the job requirements can arise as a result of women being described as more communal and less agentic than men. It is important to take caution, however, because letters of recommendation are not heavily weighted in some organizations and occupations.

As with most research, there are limitations to the current research. Although we used archival data and not hypothetical letters of recommendation in Study 1 (which is a strength), we cannot rule out the possibility that the differences in communal and agentic descriptions in the letters were based on real gender differences. Future research might try to disentangle true differences from perceived differences in agency and communion. Another potential limitation is that the hireability measure from Study 2 was based on perceived intentions to hire and not on actual hiring; therefore, we take caution with claims about possible career development for women in academia. However, the participants in Study 2 were faculty members from universities and therefore have experience in reading letters of recommendation and making selection decisions.

In addition, the percentages of communal and agentic adjectives in the letters of recommendation were low, such that one or two words could make a difference. However, the gender differences were large enough to be statistically significant, and although the effect was small, research by Martel, Lane, and Willis (1996) has shown that seemingly small gender differences may have enormous impact when compounded over time. Thus, small differences or what seem like molehills of disparity can become mountains of disparity over time and experiences (see Valian, 2000). Though quantifying the content of the letters of recommendation is an objective method and strength, the use of LIWC was also a potential limitation. In particular, LIWC does not take into account the context of the meaning of words. For example, a letter writer who mentioned that the applicant worked with a conscientious research assistant was scored the same as a letter writer who mentioned that the applicant was conscientious.

Despite these limitations, the potential implications of the current research for the use of letters of recommendation are important. The importance of letters of recommendation for academia is well established (Arvey & Campion, 1982; Guion, 1998; Sheehan et al., 1998), and the current research demonstrates differences in how men and women are described in letters according to gender norms of communality and agency (Eagly et al., 2000; Eagly & Johannesen-Schmidt, 2001; Wood & Eagly, 2002). This research showed that communal characteristics mediate the relationship between gender and hiring decisions in academia, suggesting that gender norm stereotypes—and not necessarily the sex of applicants—can influence hireability ratings of applicants.

References

Arvey, R. D., & Campion, T. E. (1982). The employment interview: A summary and review of recent research. *Personnel Psychology, 35,* 281–322.

Baron, R. M., & Kenny, D. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Bellas, M. L., & Toutkoushian, R. K. (1999). Faculty time allocation and research productivity: Gender, race, and family effects. *Review of Higher Education, 22,* 367–390.

Brem, C., Lampman, C., & Johnson, M. (1995). Preparation of applications for academic positions in psychology. *American Psychologist, 50,* 533–537.

Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy and Law, 5,* 665–692.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* Mahwah, NJ: Erlbaum.

Camp, T. (1997). The incredible shrinking pipeline. *Communications of the ACM, 40,* 103–110.

Cascio, W. F., & Aguinis, H. (2004). *Applied psychology in human resource management* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Collins, L. M., Graham, J. W., & Flaherty, B. P. (1998). An alternative framework for defining mediation. *Multivariate Behavioral Research, 33,* 295–312.

Crocker, J., Major, B., & Steele, C. (1998). Social stigma. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (Vol. 2, 4th ed., pp. 504–553). New York: McGraw-Hill.

Eagly, A. H., & Johannesen-Schmidt, M. C. (2001). The leadership styles of women and men. *Journal of Social Issues, 57,* 781–797.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109,* 573–598.

Eagly, A. H., Wood, W., & Diekman, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Traunter (Eds.), *The developmental social psychology of gender* (pp. 123–174). Mahwah, NJ: Erlbaum.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity.* New York: Prentice Hall.

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions.* Mahwah, NJ: Erlbaum.

Hebl, M. R., Tickle, J., & Heatherton, T. F. (2000). Awkward moments in interactions between nonstigmatized and stigmatized individuals. In T. Heatherton, R. Kleck, M. Hebl, & J. Hull (Eds.), *The social psychology of stigma* (pp. 275–306). New York: Guilford Press.

Heilman, M. E. (1995). Sex stereotypes and their effects in the workplace: What we know and what we don't know. *Journal of Social Behavior and Personality, 10,* 3–26.

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. (1995). Sex stereotypes: Do they influence perceptions of managers? *Journal of Personality and Social Psychology, 10,* 237–252.

Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks? The implied communality deficit. *Journal of Applied Psychology, 92,* 81–92.

Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male tasks. *Journal of Applied Psychology, 89,* 416–427.

James, L. R., Mulaik, S. A., & Brett, J. M. (2006). A tale of two methods. *Organizational Research Methods, 9,* 233–244.

Johnson, M., Elam, C., Edwards, J., Taylor, D., Heldberg, C., Hinkley, R., & Comeau, R. (1998). Medical school admission committee members' evaluations of and impressions from recommendation letters. *Academic Medicine, 73,* S41–S43.

LaCroix, P. P. (1985, fall). Sex in recs. *Journal of College Admission,* 24–26.

Landrum, R., Jeglum, E., & Cashin, J. (1994). The decision-making process of graduate admissions committees in psychology. *Journal of Social Behavior and Personality, 9,* 239–248.

Lopez, S., Oehlert, M., & Moberly, R. (1996). Selecting criteria for American Psychological Association-accredited internship programs: A survey of training directors. *Professional Psychology: Research and Practice, 27,* 518–520.

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,* 777–785.

MacKinnon, D. P. (2000). Contrasts in multiple mediator models. In J. S. Rose, L. Chassin, C. C. Presson, & S. J. Sherman (Eds.), *Multivariate applications in substance use research: New methods for new questions* (pp. 141–160). Mahwah, NJ: Erlbaum.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effects. *Prevention Science, 1,* 173–181.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7,* 83–104.

Martel, R., Lane, D. M., & Willis, C. (1996). Male–female differences: A computer simulation. *American Psychologist, 51,* 157–158.

McCarthy, J. M., & Goffin, R. D. (2001). Improving the validity of letters of recommendation: An investigation of three standardized reference forms. *Military Psychology, 13,* 199–222.

Muthen, L. K., & Muthen, B. O. (1998). *Mplus user's guide.* Los Angeles: Muthen & Muthen.

National Science Foundation. (2008). *ADVANCE: Increasing the participation and advancement of women in academic science and engineering careers* (NSF 07–582). Retrieved June 23, 2008, from http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf07582

Olsen, D., Maple, S. A., & Stage, F. K. (1995). Women and minority faculty job satisfaction: Professional role interests, professional satisfaction, and institutional fit. *Journal of Higher Education, 66,* 267–293.

Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion, 10,* 601–626.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC 2001): A computerized text analysis program.* Mahwah, NJ: Erlbaum.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77,* 1296–1312.

Raudenbush, S. W., Bryk, A., Cheong, Y. F., & Congdon, R. (2004). *HLM6: Hierarchical linear and nonlinear modeling.* Chicago: Scientific Software International.

Rudman, L. A., & Glick, P. (2001). Perspective gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57,* 743–762.

Schmader, T., Whitehead, J., & Wysocki, V. H. (2007). A linguistic comparison of letter of recommendation for male and female chemistry and biochemistry job applicants. *Sex Roles, 57,* 509–514.

Sheehan, E. P., McDevitt, T. M., & Ross, H. C. (1998). Looking for a job as a psychology professor? Factors affecting applicant success. *Teaching of Psychology, 25,* 8–11.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods, 7,* 422–445.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173–180.

Taylor, D. (2007). Employment preferences and salary expectations of students in science and engineering. *Professional Biologist, 57,* 175–185.

Tommasi, G. W., Williams, K. B., & Nordstrom, C. R. (1998). Letters of recommendation: What information captures HR professionals' attention? *Journal of Business and Psychology, 13,* 5–18.

Trix, F., & Psenka, C. (2003). Exploring the color of glass: Letters of recommendation for female and male medical faculty. *Discourse and Society, 14,* 191–220.

U.S. Equal Employment Opportunity Commission. (1978). Uniform guidelines on employee selection procedures. *Federal Register, 43,* 38290–38315.

Valian, V. (1998). *Why so slow? The advancement of women.* Cambridge, MA: MIT Press.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4–69.

Watson, C. (1987). Sex-linked differences in letters of recommendation. *Women and Language, 10,* 26–28.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Windall, S. E. (1988). AAAS presidential lecture: Voices from the pipeline. *Science, 241,* 1740–1745.

Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin, 128,* 699–727.

Appendix

Hireability Scale Items

1. How likely would you be willing to hire this candidate?

2. To what extent is this a "top-notch" candidate?

3. Is it likely that this candidate will make an effective academician?

4. How "excellent" is this candidate based on this letter?