Testing for Backlash in Hiring:
A Field Experiment on Agency, Communion, and Gender

Rickard Carlsson, Jens Agerström, Fredrik Björklund, Magnus Carlsson, and Dan-Olof Rooth

Author note:
Correspondence concerning this article should be addressed to Rickard Carlsson, Department of Psychology, Linnaeus University, 391 82 Kalmar, Sweden. Email: Rickard.Carlsson@lnu.se.
Abstract

It has been suggested that women (men) who appear agentic (communal) when applying for jobs suffer a backlash in the form of reduced chances of being hired. However, the evidence of backlash is mainly restricted to simulated hiring decisions with undergraduates as participants.

To examine whether backlash occurs when men and women apply for real jobs in the labor market, we conducted a field experiment. Gender, agentic and communal traits were manipulated in the applications. Whether or not the applications resulted in a job interview invitation constituted the dependent variable. We find no evidence of backlash, suggesting that women are not punished for presenting themselves as agentic in their job applications, nor are men punished for appearing communal.
Women are generally perceived to be more communal (e.g., caring) than men, whereas men are generally perceived to be more agentic\(^1\) (e.g., assertive) than women (Prentice & Carraranza, 2002). Intuitively, this suggests that in order for women to have the same career opportunities as men in male dominated areas, they need to make up for the stereotype of being less agentic. However, it has been proposed that women who do so will experience a backlash, i.e., negative evaluations for violating prescriptions of feminine niceness, putting them at risk for social and economic penalties (Rudman, 1998). For example, a woman who presents herself as a cutthroat career-person will be frowned upon and perceived as socially deficient, compared to a man who acts in an identical way. Conversely, men might risk backlash for expressing communal (counter-stereotypical) behavior (Moss-Racusin, Phelan & Rudman, 2010).

Research on backlash in hiring has primarily been conducted within the confines of the laboratory, with the typical finding that agentic female applicants receive lower hireability ratings (e.g., Rudman & Glick, 1999, 2001). Although useful, results from laboratory studies do not always extend to real-life treatment, simply because there tend to be many more variables operating in real-life settings against which the effect must prevail. In hiring contexts such variables may include recruiter experience, incentives for hiring the best candidate, accountability, and anti-discrimination laws.

A few studies have probed for a backlash effect beyond the laboratory. In a longitudinal study, O’Neill and O’Reilly (2011) found that agentic women do not receive fewer promotions overall than non-agentic women. Furthermore, an Austrian field experiment, in which job applications were submitted to job openings, found no evidence of backlash in the recruiters’ decisions about whom to invite for an interview (Weichselbaumer, 2004). However, the Weichselbaumer study suffers from several limitations that prevent us from drawing any strong
conclusions concerning backlash. First, there were only three applicants (a male, a communal female, and an agentic female candidate), which greatly limits the generalizability of the results. Second, there was no manipulation of agency vs. communion for the male candidate, making it impossible to ascertain if males are subjected to backlash. Additionally, the communal and agentic female applicant differed in a number of respects such as physical appearance and clothing, confounding these variables with agency and communion.

The overall aim of the present study is to test for a backlash effect when fictive applicants apply for real jobs. Specifically, are women punished for appearing agentic rather than communal, and are men punished for appearing communal rather than agentic?

**Method**

**Design**

The most common approach to test for discrimination in field experiments is to randomly assign the independent variable (e.g., gender) to a few matched applications (see Riach & Rich, 2002 for an overview). However, this approach severely limits the generalizability of the experiment, since the effect of the independent variable is limited to a few matched applications. Although matched, the applications will each signal a number of things. For example, the applicants may (both) be highly agentic, love sports, and have an excellent CV, or they may (both) be communal, theater-loving applicants with mediocre CVs. Because, the independent variables are randomly assigned to the matched applications, internal validity is achieved. Yet, it is possible that the discrimination effect found may have been different for another set of matched applications.
In the present study, we seek to remedy the aforementioned problem with external validity by randomly generating each application through a random assignment of a large number of variables outside our research focus: previous employments, years of unemployment, time studied abroad, type of leisure activities commonly engaged in, and the residential address. These variables were not analyzed, but served as background noise to enhance ecological validity. Importantly, because all variables were randomly assigned independently of each other, they cannot confound any effect.

Data selection

The current dataset come from a large field experiment (8,466 applications) designed to answer a number of different research questions, ranging from effects of previous unemployment spells to the role of agency and communion in ethnic discrimination (see Agerström, Björklund, Carlsson & Rooth, 2012; Eriksson & Rooth, in press). For efficiency reasons, and because of the nature of our design (random assignment of variables), we conducted one (large) field experiment rather than a separate field experiment for each research question.

Because of space restrictions, we choose to limit our reporting in two ways. First, the method section omits some details regarding the materials (e.g., pre-testing) and procedure (e.g., details on how the different text versions constituting the independent variable were created), and focuses on the unique aspect of the present research instead (see Agerström et al., 2012 for more details about the methodology). Second, since conducting statistical analyses on the full-scale field experiment would entail a lengthier and more complex results section, as well as some redundancy, we limited our analyses to the part of the dataset that directly pertains to our research question. Accordingly, we specifically examined male and female applicants whose
personal letter either emphasized agency over communion or communion over agency. Our statistical analyses therefore omitted data from 2,804 applications that contained a manipulation the male candidate’s ethnicity, as well as 2,791 applications in which communion and agency were equally expressed. Importantly, an analysis based on the full-scale field experiment yields virtually identical results.

**Materials and procedure**

2,871 applications were submitted to 2,271 job openings posted by the Swedish Employment Agency, between March and November, 2007. Our dependent variable was whether the job applicant received a job interview invitation. Invitations were received either by email or voicemail. The job openings included female-dominated (cleaners, restaurant workers, accountants, nurses, primary school teachers, shop sales assistants), gender-balanced (high school teachers), and male-dominated (business sales assistants, construction workers, motor-vehicle drivers and computer professionals) occupations.

The general appearance of the job applications used templates provided by the Swedish Employment Agency and contained a letter of interest and a CV. The letter of interest contained the gender manipulation (name: Anna vs. Eric), as well as the manipulation of agency over communion (see Appendix). These paragraphs have been pre-tested to strongly signal different levels of agency and communion. Further, they have been found to moderate ethnic discrimination in another sub-sample of our field experiment, confirming that the signal is sufficiently strong to affect callback rates (see Agerström et al., 2012).
Results

Because both our dependent and our independent (gender, agency vs. communion) variables were dichotomous, we ran linear probability model regressions with clustered robust standard errors (using STATA 12). This allowed us to directly interpret the results in terms of the probability of being called to an interview. However, due to clustering on the organization, we could not calculate standardized beta coefficients.

We coded gender as female = 0 and male = 1, and the agency variable as 0 meaning more communal than agentic and 1 meaning more agentic than communal. We also included the interaction between these dummy variables.

The constant reveals (Table 1) that communal female applicants were invited to interviews in 30% of the cases. The probability does not change for agentic women, and there is no significant interaction with applicant gender. Hence, there is no backlash towards agentic women or communal men. Further, there was no significant difference between the male and female applicant. To rule out occupation as a moderating factor, we ran the same regression separately for the male- and female-dominated job openings. These analyses yielded virtually identical results, with no indication of a backlash effect. Finally, we conducted a series of more complex analyses based on the full-scale field experiment (i.e., including the letters of interest where agency and communion were balanced). We also tested for the interaction with male-typed jobs. The results remained unchanged.
Table 1. Linear probability regression model with applicant gender, agency over communion, and their interaction, predicting probability of job interviews

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male applicant</td>
<td>-0.02</td>
<td>0.02</td>
<td>.424</td>
</tr>
<tr>
<td>Agentic</td>
<td>0.00</td>
<td>0.02</td>
<td>.961</td>
</tr>
<tr>
<td>Male applicant X Agentic</td>
<td>0.01</td>
<td>0.03</td>
<td>.802</td>
</tr>
<tr>
<td>Constant</td>
<td>.30</td>
<td>0.02</td>
<td>.000</td>
</tr>
</tbody>
</table>

Note. All variables are dummy coded. Hence, the constant is the callback for a female applicant who is communal rather than agentic, and the effect of being agentic rather than communal is the simple effect for female applicants. Standard errors are robust and adjusted for 2271 clusters (job openings). Standardized beta cannot be calculated when clustering.

Discussion

The current research set out to investigate whether agentic women and communal men would face a backlash when applying for real jobs. Thus, we conducted a field experiment, using an improved design compared to the previous field experiment by Weischelbaumer (2004). The results suggest that neither women nor men face backlash, regardless if the job is male- or female-dominated. While this finding is inconsistent with studies conducted on American university students in simulated hiring contexts (Rudman & Glick, 1999; 2001), it mirrors the results from other studies conducted in more ecologically valid settings (O’Neil & O’Reilly, 2011; Weischelbaumer, 2004). Importantly, these three applied studies complement each other. Weischelbaumer's (2004) field experiment in the Austrian labor market had strong manipulations (matched applications that included photos) but suffered from potential confounds, and low levels of generalizability. In contrast, the strength of the longitudinal study on real-life promotions by O’Neil and O’Reilly (2011) is its ecological validity. However, its non-
experimental approach casts some doubt on the internal validity. The present field experiment on the Swedish labor market complements these two previous studies, combining strong internal validity with high levels of generalizability, due to high (randomly generated) variation in the applications as well as in the jobs applied for. In isolation, a single null effect may not be very informative, since it could have been the result of specific characteristics of that study. However, taken together, these three applied studies on the backlash effect on the European and US labor market suggest that there is no, or very weak, effect of backlash in hiring on these labor markets.

It is important to emphasize that this lack of backlash in these three applied studies does not constitute failed replications of the original laboratory studies, nor is it a rebuttal of the theoretical concept of backlash. Rather, one has to consider the different role of basic and applied research. Laboratory experiments are useful to identify the existence of an effect, and applied studies are useful in gauging the size of that effect in a certain applied setting, informing us about its applied importance. However, the consistent lack of backlash in these applied studies does not suggest that backlash is entirely irrelevant. First, even taken together these three studies are, of course, not entirely exhaustive. For example, although O’Neil and O’Reilly (2011) found no backlash for promotions, their findings cannot be reliably generalized to the highest level of positions (i.e., CEO), simply because so few people ever advance to such positions. Further, although these studies are informative regarding the size of the effect, applied research is, for obvious reasons, not particularly suited at answering questions regarding the underlying process. Indeed, it is possible that employers perceive agentic women and communal men differently, but act professionally and treat them fairly. Interestingly, although O'Neil and O'Reilly (2011) found no main effect of backlash against agentic women, they discovered that it interacted with self-monitoring. This suggest that agentic women are potentially subject to backlash in promotions on
the US labor market, but are able to correct for this by monitoring their own behavior. In sum, although we are, based on the combined evidence from three studies that complement each other, quite confident in that there is, in general, little or no backlash in hiring on the Swedish, Austrian, or US labor markets, we are in the dark as to why this is the case. Hence, our suggestion for future research in backlash is to examine its underlying processes.

In conclusion, although backlash against agentic women/communal men appears to be of minor concern in the hiring process, we are certainly not suggesting that organizations should cut-down on their work towards gender equal hiring practices. On the contrary, the lack of backlash could be viewed as testimony to that the struggle toward a more gender equal labor market is starting to pay off.

References


Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn’t be, are allowed to be, and don’t have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly, 26*, 269-2281. DOI: 10.1111/1471-6402.t01-1-00066


These dimensions are commonly referred to as the “big two” in social cognition, and are also commonly called warmth (communion) and competence (agency) instead, but the underlying construct is essentially the same (see Wojciszke & Abele, 2008). However, we prefer the terms agency and communion in the present study, since these are the labels typically used in the backlash literature.

The lack of overall gender discrimination closely mirrors that of previous studies on the Swedish (Carlsson, 2011) and US (Bertrand & Mullainathan, 2004) labor markets.
Appendix

Following are examples of the paragraphs we used to signal higher agency than communion, or higher communion than agency. Three different text versions of the agency and communion manipulation were used.

**Example of applicant that is higher in communion than agency:**

My friends and former colleagues think that I am a warm and social person who gets along great with others, both at work and elsewhere.

Also, I think it is important to consider other people’s needs, not just money. I have a strong sense of empathy with people who are less fortunate than myself and I do some charity work.

Regarding work I really like working but at the same time I think it is important to keep a balance between work and leisure. The best days are the ones when I feel that I have done my share at work and yet have the energy to be active in my spare time. It is not important for me to be the best at work and my colleagues would probably describe me as a pretty relaxed.
Example of applicant that is higher in agency than communion

As a person I usually do not sit and keep my opinions to myself but rather say what I think straight out. Some of my former colleagues would probably call me a bit stubborn, but I believe it is important that my point is seen and that the job gets done. And I often prefer to complete my work tasks alone, since then it is easier to concentrate on what you are supposed to do.

Concerning work I am used to put great effort into it and I always try to do my best. I strive to be precise so that the work does not need to be redone. My old co-workers would probably say that I am a person who always manages to get the job done. In addition, I would describe myself as a hardworking and tenacious person who handles stress well.