Knobloch-Westerwick, S., Glynn, C. J., & Huge, M. (2013). The Matilda Effect in science communication: An experiment on gender bias in publication quality perceptions and collaboration interest. *Science Communication, 35*, 603-625.

The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest

An experiment with 243 young communication scholars tested hypotheses derived from role congruity theory regarding impacts of author gender and gender-typing of research topics on perceived quality of scientific publications and collaboration interest. Participants rated conference abstracts ostensibly authored by females or males, with author associations rotated. The abstracts fell into research areas perceived as gender-typed or gender-neutral to ascertain impacts from gender-typing of topics. Publications from male authors were associated with greater scientific quality, in particular if the topic was male-typed. Collaboration interest was highest for male authors working on male-typed topics. Respondent sex did not influence these patterns.

The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest

Female scholars experience many disadvantages around the globe and are drastically underrepresented in higher academic ranks (e.g., European Commission, 2012; NSF, 2006). Causes for this striking underrepresentation are debated until today (e.g., Ceci, Williams, & Barnett, 2009). One likely contributing factor resides in the communication among scholars, as part of science communication. Although no full consensus exists regarding the definition of science communication, several conceptualizations have included scholarly communication among scientists. For example, Burns, O'Connor, and Stocklmayer (2003, p. 191) noted that "Science communication may involve science practitioners, mediators, and other members of the general public, either peer-to-peer or between groups." A common approach to study scholarly communication is citation analysis (e.g., Knobloch-Westerwick & Glynn, in press; Tai, 2009), showing two-way scholarly communication, as one scholar's work is received and utilized (i.e., cited) by another. More frequently, however, scholarly communication may take the shape of one-way communication in which a scientist is merely the recipient and perceiver of another scientist's publication (as in the empirical study presented below).

Although ideally scientists would communicate with each other in unbiased fashion, living up to an ideal marketplace of ideas, patterns of stereotyping may still apply (e.g., Greenwald & Banaji, 1995). Communication among scientists has been described as a social system (e.g., Garvey & Griffith, 1967) and thus is likely subject to such biases. The present work pertains to a possible communication bias along the lines of a Matilda effect and is, based on an extensive review of the literature that was accessible to us, the first experimental investigation of a gender bias in science communication among scholars of which we are aware.

About twenty years ago, Rossiter (1993) introduced the term *Matilda effect* for a systematic under-recognition of female scientists. The term has been coined with reference to the well-known ‘Matthew effect’ (Merton, 1968)—over-recognition of those at the top of the scientific profession, even credit misallocation to scientists that are already well-known. The present work extends the related empirical evidence by drawing on role congruity theory (Eagly & Karau, 2002) and presents an experiment to test whether author gender and gender stereotypes associated with research topics affect perceptions of publication quality and interest in scholarly collaboration. In the following, we briefly review work related to the Matilda effect as a bias against women in science and then introduce role congruity theory (Eagly & Karau, 2002). Hypotheses derived from this background will be tested with a sample of young scholars in communication. The academic field of communication presents a relevant discipline to study in this context because its topics and methods are diverse and eclectic. The stimuli all pertained to communication as a social science (as opposed to critical studies or humanities) and participants were recruited from institutions with a predominantly social science focus on communication.

Gender Biases in Science

A number of studies have found evidence for a bias against female scientists. Compared to their male counter-parts, they receive grants less often and receive smaller grant allocations (Bornmann, Mutz, & Daniel, 2007; RAND, 2005; Wenneras & Wold, 1997), fewer citations (e.g., Knobloch-Westerwick & Glynn, in press), and fewer scientific awards—for example, “men were more than eight times more likely than women to win a scholarly award and almost three times more likely to win a young investigator award” (Lincoln, Pincus, Koster, & Leboy, 2012). Further, among recipients of career development grants, women are significantly less likely than men to obtain subsequent academic success for a number of criteria, such as receiving major grants, getting promoted, and holding academic leadership positions (Jagsi, DeCastro, Griffith, Rangarajan, Churchill, Stewart, et al., 2011). Evaluation biases against women may play a role in this context. In fact, faculty recommendation letters have been found to differ in language use and praise by sex of the evaluated individual, favoring males (Trix, Frances, & Penska, 2003). Yet women likely benefit from a blind-peer review process, as it reduces gender biases through author anonymity (Budden, Tregenza, Aarssen, Koricheva, Leimu, & Lortie 2008). Taken together, the data suggest a "pervasive culture of negative bias—whether conscious or unconscious—against women in academia" (Ledin, Bornmann, Gannon, & Wallon, 2007, p. 986).

However, the empirical evidence on causes for women's underrepresentation in science has gaps and inconsistencies. Ceci, Williams, and Barnett (2009) provided a recent comprehensive review of the existing evidence, along with their judgment on major causes for gender gaps in science. They believe that women's choices are the most important explanatory factor, yet concede that these choices are subject to constraints and sociocultural influences. Further, specifically for evaluation biases, they conclude the evidence to be limited but point out that even a small bias accumulates to multiplicative impacts on women's representation in academia.

On a methodological note, the majority of research about discrimination against female scientists has utilized data on naturally occurring actions or events such as grant allocations, promotion decisions, or citations and thus carries great ecological validity. The disadvantage, however, is that many statistical controls need to be applied to account for the many competing explanations and confounding variables (see Ceci, Williams, & Barnett, 2009, for a discussion of the various potential influences). In contrast, experimental work can establish clear causal impacts resulting from scholars' gender. But this type of research is very scarce, as summarized below.

An experiment by Steinpreis and colleagues (1999) revealed that among psychology scholars, both sexes were more likely to vote to hire a male scientist rather than a female scientist, based on otherwise identical curriculum vitae information (Steinpreis, Anders, & Ritzke, 1999). Most recently, a very similar experiment was conducted with faculty in the biological and physical sciences to examine reactions to student applications for a lab manager position (Moss-Racusin, Dovidio, Brescoll, Graham, and Handelsman (2012). The results showed that both male and female faculty found female applicants less hireable and less competent and offered them lower salaries and less mentoring. Similar experimental evidence also emerged beyond a science context (e.g., for hiring of musicians; Goldin & Rouse, 2000), evaluation of work teams (Heilman & Haynes, 2005), as well as leadership perceptions (Eagly & Karau, 2002). But very little experimental evidence on scholars' gendered perceptions of peers is at hand, as the work presented by Steinpreis et al. (1999) and Moss-Racusin et al. (2012) appears to stand alone in this regard. Yet these two experiments simulated only hiring decisions, while a bias against women may apply much more broadly to everyday science communication among scholars, such as conference presentations and publications. None of the existing work on social biases has looked at such science communication among scholars. The present work aims to address this gap with an experiment on young scholars' perceptions of the quality of scholarly contributions and draws on role congruity theory to develop hypotheses.

Role Congruity Theory

Role congruity theory (e.g., Eagly & Karau, 2002) is a suitable framework to explain a gender bias against women scientists along the lines of a Matilda effect. The theory originated in social role theory (Eagly, 1987), which conceptualizes gender through the construct of gender roles. Social roles are socially shared expectations regarding individuals holding specific social position or from a certain social category. Gender roles represent consensual beliefs about attributes of women and men but also include normative expectations in that they prescribe qualities or tendencies thought to be desirable for each sex (Eagly, 1987). A key assumption of social role theory suggests that most of the beliefs about the sexes pertain to *communal* and *agentic* characteristics (Eagly, 1987). Communal characteristics, more often associated with women, include being concerned with others' well-being (e.g., sympathetic, helpful). On the other hand, agentic attributes, linked more often to men, refer to an assertive and self-assured disposition (e.g., self-confident, objective, ambitious, competitive, leadership skills) that match up well with stereotypes of scientists. As an application of this theoretical reasoning, cross-national research by Nosek, Smyth, Sriram, Lindner, Devos, Ayala, et al. (2009) examined implicit gender–science stereotyping and found it predicts sex differences in science and math achievement

From this background, role congruity theory (e.g., Eagly & Karau, 2002) evolved and was applied to women in leadership roles, which are commonly associated with men (Koenig, Eagly, Mitchell, & Ristikari, 2011). However, it has been extended to women in scientist roles (Knobloch-Westerwick & Glynn, in press). Role congruity theory extends social role theory by emphasizing the congruity between gender roles and other roles such as the scientist role. It outlines influences and dynamics that impact congruity perceptions and subsequent consequences for perceptions and behaviors. Role congruity theory postulates that bias against female scientists originates in differences between a female gender role and the common expectations toward individuals in a scientist role (e.g., Nosek at al., 2009). Such prejudice emerges when perceivers evaluate women in scientists' roles because of perceived inconsistency between the primarily communal qualities that perceivers usually associate with women and the rather agentic and masculine qualities commonly believed to apply to scientists.

Hence, beliefs about scientists and women tend to be dissimilar, whereas beliefs about scientists and men tend to be similar. As a result, observing someone in a scientist role may instigate gender-based expectations in competition with scientist role-based expectations. When judging female scientists, this process leads to an incongruity in perceptions. This incongruity affects the judgment of female scientists negatively with regard to performance in that scientist role. This social-psychological phenomenon of role incongruity accounts for a Matilda effect of a bias against women when their performance is evaluated in a scientific context. Specifically among scholars, this negative evaluation bias should subsequently reduce interest in connecting or working with the evaluated individual. Based on these considerations, we propose the following two hypotheses.

H1: Male authors' scholarly contributions are associated with greater scientific quality than female authors' scholarly contributions.

H2: Male authors are perceived as more attractive for scholarly exchange and collaboration.

Additionally, the topics that scientists focus on in their work may be connected to gender role expectations as well and subsequently increase or decrease role congruity perceptions, depending on the gender norms associated with the topics. For example, the fields of mathematics and physics are more often associated with scientist attributes often linked to men (e.g., objective, skilled in math), whereas the humanities and social sciences may involve greater focus on relationships, congruent with the communal qualities commonly associated with women. Accordingly, a female scientist studying the male-typed area of physics should induce more incongruity between the female gender role and the scientist role than a female scholar in social work as a more female-typed field.

Moreover, research topics differ in the extent to which they are associated with stereotypically male or female characteristics. For instance, within communication research, the area of media effects on children could induce more stereotypically female attributes such as nurturing, caring, etc., whereas political communication research might be more linked to stereotypically male characteristics such as competitiveness, assertiveness, etc. Hence, a female communication researcher who works on political communication may induce more incongruity between female gender role and scientist role than a female scholar who works on children and media. Such stereotyping of research topics in communication research has been demonstrated on empirical grounds and is also reflected in the proportions of the sexes among authors in communication research publications (Knobloch-Westerwick & Glynn, in press). These considerations lead to the third hypothesis.

H3a: The effects suggested in H1 for scientific quality depend on the research topic that an author is associated with.

H3b: The effects suggested in H2 for scholarly exchange and collaboration depend on the research topic that an author is associated with.

Also based on role congruity theory, we further predict that the gender gaps suggested in H1 and H2 will depend on individuals' gender roles attitudes. This perceiver characteristic should act as an important moderator on the suggested effects. For onlookers with more traditional gender role attitudes, incongruity and subsequent prejudice should be greater.

H4a: The effect suggested in H1 for scientific quality is more pronounced among perceivers with traditional gender roles attitudes.

H4b: The effect suggested in H2 for scholarly exchange and collaboration is more pronounced among perceivers with traditional gender roles attitudes.

Accordingly, gender bias against women scholars will be more pronounced among men, who overall hold more traditional gender roles attitudes (e.g., Bolzendahl & Myers, 2004). Plausibly, the sex difference in gender role ideology commonly found in the general population also exists among scientists. Although higher education leads to more egalitarian gender ideology, it does not reduce the sex difference pertaining to gender ideology (Bryant, 2003). Indeed, prior research on the Matilda effect has yielded sex differences among scholars in that male researchers exhibited a bias toward citing same-sex authors more frequently, whereas female researchers cited authors of both sexes in proportion to the pool of publications available for being referenced (Knobloch-Westerwick & Glynn, in press). However, other work has shown that both sexes tend to evaluate women less favorably in performance contexts and women may prefer a male boss over a female boss even more strongly than men (Elsesser & Lever, 2011). Thus we examine how the sex of the perceiver influences biases against female scholars with a research question.

RQ1: Are the effects suggested in H1 for scientific quality (RQ1a) and H2 for scholarly exchange and collaboration (RQ1b) more pronounced among male scholars than female scholars?

Method

Overview

A national sample of 243 graduate students enrolled in communication programs evaluated abstracts taken from the 2010 International Communication Association's (ICA) annual conference. These abstracts were presented as being authored by either female or male authors, with this association carefully rotated for research areas categorized as female-typed, male-typed, or gender-neutral (this categorization of research topics was adopted from a citation analysis by Knobloch-Westerwick & Glynn, in press). Participants also completed questions on gender roles attitudes.

Recruitment and Respondents

A national sample of graduate students enrolled in communication programs at both the PhD and MA level was recruited. Email requests were sent to 39 different graduate chairs or program directors asking them to forward a recruiting email and link for the study to their current graduate students. Of the 243 individuals from 20 different institutions that completed the online session, 70% were female. The average age was 28.55 years (*SD* = 5.97). Most respondents were American (75%), followed by Chinese (4%) and Korean (3%). Further, most participants were Caucasian (73%) or Asian (16%). The majority was enrolled in a Ph.D. program (72%) while 25% were M.A. students. The average number of years respondents had been enrolled in a graduate program was 2.9 (*SD* = 1.59). Participants were affiliated with graduate programs at Ohio State (16%), Texas-Austin (11%), Missouri (9%), Indiana (8%), University of Southern California (7%), Utah (6%), Arizona State (5%), Illinois (5%), and others.

Procedure

Students who received the invitation email were asked to participate in an online study. They were provided with a hyperlink to the study, set up in Qualtrics (a widely used platform for online surveys and experiments). The purpose of the research was described as follows: "We are interested in how social scientists/students perceive academic papers (based on the abstract). For the present survey, we compiled some research abstracts from a conference program of a communication research convention." Then participants read and evaluated 15 abstracts taken from the 2010 International Communication Association's (ICA) annual conference in Singapore. Six target abstracts were displayed with author names indicating either female or male authors, with the sequence of abstracts, association with gender-typed or gender-neutral research topics (see Knobloch-Westerwick & Glynn, in press) carefully balanced, while the other nine abstracts served as distracters. Further, participants completed questions pertaining to gender roles attitudes and a question regarding their recall of displayed authorship information. As an incentive, respondents were awarded their choice of a $25 Amazon.com gift card or a $25 Paypal credit.

Stimuli Pretest

To establish that the presented research abstracts were comparable with regard to overall impressions while falling into gender-typed research areas, a pretest was conducted. Six assistant professors in communication were asked to rate abstracts based on their personal impressions, explained as a preliminary step for research into how impressions of scholarly work are formed. Three women and three men agreed to help with the project. Their research interests spanned across news/journalism, health/mass communication, political/intercultural communication, intergroup communication, media effects, and health/communication technologies.

Forty-five abstracts taken from the 2010 International Communication Association conference, all between 135 and 150 words long, were presented on a paper questionnaire. None of the raters attended this conference; data were collected before the conference. Each abstract was rated for *interesting*, *relevant*, *rigorous*, and *publishable*, on a 0-10 scale with *not at all* and *extremely* as anchors. Moreover, each abstract was categorized into research areas, with up to two categorizations considered. The 15 research area categories were adopted from Knobloch-Westerwick and Glynn (in press), as they had been empirically shown to be more or less associated with communication scholars of the two sexes or as gender-neutral (see further details below and specifically reported by Knobloch-Westerwick & Glynn, in press).

The data were condensed by averaging each of the four rating dimensions across raters, resulting in a data set with the 45 abstracts as units. A factor analysis (principal-component) with the four rating dimensions yielded just one factor that explained 81% of the variance; the Cronbach's alpha was .92. Thus the four dimensions were collapsed into an 'evaluation' variable.

Stimuli

Abstracts. Based on the stimuli pretest evaluations and categorizations into research areas, the 15 abstracts listed by headings in Table 1 were selected to serve as stimuli in the main experiment. All these abstracts fell into the narrow range of 6.41-7.46 on the evaluations variable, which was based on a 0-10 scale. Furthermore, the abstracts were selected so that four were mostly categorized into the female-typed research areas (communication pertaining to children, parenting, and body image), four fell into male-typed research areas (such as political communication, computer-mediated communication, news, and journalism), and four were mostly categorized into gender-neutral research areas (such as health communication or intercultural communication).

One challenge emerged due to the fact that much fewer research areas are primarily associated with female scholars and one of these few areas, namely body image, was often also categorized as health. However, to avoid having only research related to children and parenting as work associated with female scholars, we decided to include one abstract on body image communication as well, even though it had often been categorized into both 'body image' (female-typed) and 'health' (gender-neutral).

Displayed author names. Respondents were presented with two neutral-typed abstracts (one with two female authors, one with two male authors), two female-typed abstracts (one with two female authors, one with two male authors), and two male-typed abstracts (one with two female authors, one with two male authors). To vary the target abstracts and base participants' ratings on a larger stimuli set, six abstracts were drawn for each individual participant from the twelve abstracts listed as female-typed, male-typed, or gender-neutral in Table 1 based on a rotation across participants. However, across the sample, all twelve abstracts were utilized, though in different combinations of six each for individual participants. The remaining nine abstracts from the set of 15 were displayed as distracters. Six of them featured a female and a male author, with author sequence and research areas counter-balanced. In addition, two distracter abstracts from gender-neutral research areas and one from a male-typed area were presented as single-authored by a male or with four authors (two females and two males, with genders alternating in the author sequence), in contrast to all abstracts being presented with two authors. The abstract with four authors was always presented last and served as basis for a question on recall of the author information.

The author names were generated by randomly sampling from the U.S. Social Security Administration's database of popular names for people born in the 1960s in the United States. We sampled from the top 50 first names and top 200 last names and randomly paired the results together to generate a full name for the authors. These names were also checked against the Comm Abstracts database of names to ensure that we did not include any names of actual communication researchers. Table 2 lists the names utilized in the experiment. First and last names were randomly paired together and then assigned to the experimental abstracts. These names were created to avoid the possibility that using real author names would introduce experimental noise in our study.

Measures

Abstract ratings. After viewing an abstract, respondents indicated their impressions on a ten-point scale, with the anchors "Not at all" and "Very," based on the following items: Rigorous, by competent authors, influential, important, innovative, publishable in a prestigious journal, high-quality contribution, reflective of expertise. Respondents were also asked the degree to which, aside from their own specific topic interests, they would be interested in discussing research with the authors as well as collaborating with the authors.

To condense the ratings that participants provided for all abstracts, each of the ten dimensions (e.g., 'high quality contribution') was averaged across all presented 15 abstracts. The collapsed data with the resulting ten variables were subjected to a factor analysis (principal component, varimax), which yielded two factors. The first factor showed high loadings ($\geq .831$) on the first eight items, while these items had low loadings ($\leq .386$) on the second factor, and explained 77.3% of the variance. The second factor featured high loadings ($\geq .880$) on the two remaining items ("interested in collaborating with the author(s)" and "interested in discussing research with the author(s)") while these had low loadings ($\leq .340$) on the first factor, and explained 10.5% of the variance. Based on this factor structure, two condensed variables were created for the abstracts of interest—one with the average score for the first eight ratings, labeled *Scientific Quality* ($M = 5.46$, $SD = 1.28$) and one with the average score for the last two ratings, labeled *Collaboration Interest* ($M = 4.13$, $SD = 1.56$).

Gender roles attitudes. Fourteen items adopted from Davis and Greenstein (2009) served to capture attitudes toward gender roles, for example "It is more important for a wife to help her husband's career than to have one herself" and "A wife's most important task is caring for her children." Eleven items from a dogmatism scale (Trodahl & Powell, 1965; e.g., "Most people just don't know what's good for them") were randomly interspersed to veil the purpose of the research. The items for gender roles attitudes were subjected to an item analysis, after reversing three of the items. Given satisfactory a Cronbach's alpha of .835, the items were collapsed in a mean score (M = 1.91, SD = .55), with lower scores indicating greater gender equality support.

Attention to author information. All participants saw the same last abstract with two female and two male authors and were asked on the following screen page how many authors total (37% correct responses) and how many female authors (36% correct responses) and male authors (32% correct responses) had been associated with it. Only 26% gave correct responses to all three questions.

Familiarity with the original abstracts. Five percent indicated that they had attended the ICA conference in Singapore (from which the abstract texts were taken); 17% said they had browsed the program. The average rating for the question "Did any of the abstracts seem familiar to you?" was 3.09 (SD = 2.58) on a 9-point scale, with 47% on the option for 'not at all.'

Results

Impacts of Author Gender and Research Topic on Perceived Quality

To address H1, H3a, and RQ1a, an analysis of variance using *Scientific Quality* ratings as repeated measures was conducted, with *topic* (gender-neutral vs. female-typed vs. male-typed) and *author gender* (female authors vs. male authors) as within-group factor and respondent sex as between-group factor. In addition, respondent age and ethnicity (Caucasian vs. Asian vs. other) served as control variables and were included as covariates.

First off, the analysis yielded a number of impacts that are not relevant for the hypotheses. Respondent sex ($F(1, 226) = 12.63$, $p < .001$, *partial* $\eta^2 = .053$; $M_{\text{women}} = 5.58$, $SD = 1.24$, vs. $M_{\text{men}} = 5.18$, $SD = 1.34$), respondent age ($F(1, 226) = 10.24$, $p = .002$, *partial* $\eta^2 = .043$; $r = -.18$, $p = .005$), and ethnicity ($F(2, 226) = 3.07$, $p = .048$, *partial* $\eta^2 = .026$; $M_{\text{Caucasian}} = 5.49$, $SD = 1.25$, vs. $M_{\text{Asian}} = 5.26$, $SD = 1.30$, vs. $M_{\text{Other}} = 5.54$, $SD = 1.40$) affected the general ratings level for *Scientific Quality*. An interaction between ethnicity and topic ($F(4, 452) = 3.11$, $p = .015$, *partial* $\eta^2 = .027$) reflected that Asian participants rated abstracts on female-typed topics particularly low.

More importantly, the effect suggested in H1 was found to be significant ($F(1, 226) = 4.52$, $p = .035$, *partial* $\eta^2 = .020$), because abstracts from male authors ($M_{\text{estimated}} = 5.33$, *S.E.* = .12) were associated with greater *Scientific Quality* than abstracts from female authors ($M_{\text{estimated}} = 5.26$, *S.E.* = .12). Moreover, an interaction between research topic and authors' gender in line with H3a emerged as significant ($F(2, 452) = 3.08$, $p = .047$, *partial* $\eta^2 = .013$) (see Figure 1). In line with H3a, the perceived *Scientific Quality* of a research contribution depended both on author gender and research topic. Separate repeated measures analyses of variance were conducted to examine each simple effect and revealed that abstracts received particularly high ratings if they were associated with male authors and pertained to male-typed topics. Specifically, abstracts on male-typed topics received significantly higher ratings if they were associated with male authors as opposed to female authors. Further, abstracts associated with male authors received significantly higher ratings if they pertained to male-typed topics as opposed to female-typed topics. The significant differences are indicated in Figure 1. Regarding RQ1a, respondent sex did not affect the impacts suggested in H1.

The initial ANOVA model described above was then extended by using the condensed score for *Gender Roles Attitudes* as additional covariate. While most effects reported above were also significant with this model, the ethnicity effect on overall rating levels did not approach significance ($p = .15$). *Gender Roles Attitudes* also showed an impact on the general rating level for *Scientific Quality* ($F(1, 225) = 4.69$, $p = .031$, *partial* $\eta^2 = .020$; $r = -.11$, $p = .083$). More importantly, the authors' gender impact per H1 vanished as a result of controlling for *Gender Roles Attitudes* ($p = .308$). However, an interaction between research topic and authors' gender remained significant ($F(2, 450) = 3.13$, $p = .045$, *partial* $\eta^2 = .014$) even when controlling for *Gender Roles Attitudes*.

Impacts of Author Gender and Research Topic on Collaboration Interest

To address H2, H3b, and RQ1b, an analysis of variance using *Collaboration Interest* ratings as repeated measures was conducted, with *topic* (gender-neutral vs. female-typed vs. male-typed) and *author gender* (female authors vs. male authors) as within-group factor and respondent sex as between-group factor. Again, respondent age and ethnicity (Caucasian vs. Asian vs. other) served as control variables and were included as covariates.

While irrelevant for the hypotheses, respondent sex ($F(1, 226) = 4.38$, $p = .037$, *partial* $\eta^2 = .019$; $M_{\text{women}} = 4.23$, $SD = 1.55$, vs. $M_{\text{men}} = 3.90$, $SD = 1.59$) and respondent age ($F(1, 226) = 5.94$, $p = .016$, *partial* $\eta^2 = .026$; $r = -.12$, $p = .057$) both influenced general *Collaboration Interest* levels.

More importantly, the interaction between author gender and research topic suggested in H3b emerged as significant ($F(2, 452) = 3.27$, $p = .039$, *partial* $\eta^2 = .014$) and is illustrated in Figure 2. Female authors evoked greater *Collaboration Interest* if they worked on female-typed topics than gender-neutral topics. Male authors induced greater *Collaboration Interest* if they worked on male-typed topics compared to female-typed or gender-neutral topics. In the realm of male-typed topics, *Collaboration Interest* was greater in male authors than in female authors. However, the effect suggested in H2 fell short of significance (p = .108). No support for H4b emerged. Regarding RQ1b, respondent sex did not affect the impacts suggested in H2.

The same ANOVA as reported above, just with *Gender Roles Attitudes* as additional covariate, yielded the same effect pattern.

Impacts of Gender Roles Attitudes on Perceived Quality and *Collaboration* Interest

To address H4 specifically, regression analyses were conducted with the following dependent variables: ratings for perceived *Scientific Quality* of abstracts from (a) female authors (averaged across topics) and (b) male authors (averaged across topics) and (c) the difference between these two averaged rating scores, as well as ratings for *Collaboration Interest* in (d) female authors (averaged across topics) and (e) male authors (averaged across topics) and (f) the difference between the two averaged rating scores. These analyses were run separately to avoid violation of regression analysis assumptions pertaining to dependent measures being independent from each other, as the ratings individual participants provided for abstracts associated with a *female* author would not be independent from ratings these participants provided for abstracts associated with *male* authors and subject to autocorrelation. The predictor variables were gender roles attitudes, respondent sex, and distractor ratings as control variable.

Aside from impacts of the control variable, these analyses yielded only an impact of gender roles attitudes on perceived quality of abstract associated with female authors, *beta* = -.13, p = .005. Hence, stronger support for gender equality led to higher ratings of perceived quality of female authors' contribution. The control variable was significant in this model as well, *beta* = .71, $p < .001$; the regression model had an adjusted R^2 of .529. Thus H4a was supported, while H4b was not.

Discussion

The present investigation is, to our knowledge, the first experimental study of a bias in science communication among scholars. It examined whether future scholars' ratings of scientific contributions depend on their own sex, the gender of the persons who authored the scientific contribution, and/or the topic area of the research. Hypotheses were derived from role congruity theory (Eagly & Karau, 2002).

H1 was supported—male authors' contributions were indeed associated with greater scientific quality, even though the actual content had been carefully rotated to avoid any confounding impacts. Furthermore, contributions from male authors were perceived as having particularly high scientific quality if they pertained to male-typed topics such as political communication or communication technologies, in line with H3a. Interestingly, the main effect of author sex on perceived quality vanished when the analysis controlled for gender roles attitudes, which speaks to the impact suggested in H4a. Indeed, a regression analysis showed that stronger support for gender equality led to higher ratings of perceived quality of female authors' contributions, supporting H4a. Thus, most likely, the favoritism for male authors is a consequence of conservative gender norms.

A bias against female scholars extended further to young scholars' interest in exchange and collaboration. Overall, male authors were perceived as more attractive for such interpersonal connections if they worked on male-typed topics compared to gender-neutral or female-typed topics. On the other hand, female authors fostered greater *Collaboration Interest* if they were associated with work on female-typed topics compared to gender-neutral topics. While these patterns support H3b, no evidence for a generally higher interest in collaboration with male authors per H2 and no significant impact of gender role attitudes per H4b emerged.

With regard to the research question, no impacts of perceivers' sex on the hypothesized patterns emerged. It is possible that a limitation in our study—the small portion of male participants—made it impossible to demonstrate an effect along these lines. Another interpretation is that, among young future communication scholars, both sexes fall back to gender-typed notions when asked to give an impression of others' scholarly work.

The present findings converge well with gender role congruity theory. However, some limitations and directions for future work need to be outlined. The research design does not allow for unequivocal testing of what theory is best suited to explain the demonstrated effects. To put gender role congruity theory to a more rigorous test, future research should measure role perceptions pertaining to both females' and scientists' roles. Individuals who perceive greater incongruity should give female-authored contributions lower ratings, according to the theory. Methodologically, we believe that great strides forward could be made by disentangling role congruity more specifically by measuring individual expectations for women, men, and scientists to compute perceived incongruities, possibly based on implicit measures (Gawronski & LeBel, 2007) that reflect the accessibility of these attitudes as well. In conceptual terms, drawing on research on gender-based discrimination in general employment contexts (e.g., Davison & Burke, 2000; Roth, Purvis, & Bobko, 2012) would help to advance this line of research further.

Regarding the present sample, it needs to be acknowledged that it resulted from willingness of graduate programs to cooperate with the researchers and may thus not be viewed as representative or random. We worked with a sample of young communication scholars—future work may aim to examine whether these effects are more pronounced in a sample of more ‘senior’ scholars who may hold more traditional gender roles attitudes (Bettencourt, Vacha-Haase, & Byrne, 2011). It is desirable to replicate the study with scholars already advanced in their careers, despite likely greater challenges in recruiting participants from this population. In addition, the present results can only speak to scholars from one particular discipline, communication. Given that biases against women have been shown to vary by scientific discipline (e.g., Ferber, 1988), it seems likely that examination of different academic fields would show biases of different extent. For example, along the lines of role congruity theory, the roles of a mathematician and a female may induce greater perceived incongruity than the roles of a scholar in English literature and a female, resulting in greater bias against women in scholarly communication among mathematicians than among English literature scholars.

The demonstrated impacts were relatively subtle, which needs to be considered with the following two aspects in mind. First, the stimuli abstracts were presented to the participants in a text-only format in which the authors’ sex was only to be inferred from the first names. A recall question regarding author information revealed that participants paid little conscious attention to the author information, as only about a quarter could retrieve the correct information. Thus the reported impacts, in line with the hypotheses, must result from processes that the participants were unaware of and did not invest much cognitive capacity in. In other scholarly contexts, author sex is much more evident—for instance in a bio note (“her research interests include…”), from a portrait photo on a web site, or even more so in personal encounters. Such contexts that make author gender more salient should produce larger gender biases.

Second, even a seemingly small impact of 1-3% as found in the present study will accumulate to dramatic consequences. Martell, Lane, and Emrich (1996) demonstrated with a computer simulation how a seemingly small bias translates into large impacts on career progress over time. Their model started out with equal proportions of women and men. It examined the impact of a one-percentage point difference in explained variance in work performance ratings. After successive promotion rounds, the proportion of women dropped significantly and was reduced to 35%. Thus small biases cannot be dismissed as irrelevant because they ultimately have drastic multiplicative impacts on women's representation in academia.

The present investigation highlights how science communication between scholars can affect the scientific knowledge production and reception process. Given the enormous importance of peer evaluations across all domains of individual academic careers, the demonstrated gender bias will have important implications. In contexts such as grant proposals, promotion and tenure reviews, hiring decisions etc., a scholar's sex will be a relevant factor, according to our findings. Even though the effects may seem small from a statistical perspective, they are significant and will add up greatly across individual academic careers. Moreover, certain personal characteristics such as motherhood may further accentuate gender bias. The overall conclusion is that male scholars will have a much smoother ride, especially if they work on male-typed topics.

References

Bettencourt, K.E.F, Vacha-Haase, T., & Byrne, Z.S. (2011). Older and younger adults' attitudes toward feminism: The influence of religiosity, political orientation, gender, education, and family. *Sex Roles, 64* (11-12), 863-874. DOI: 10.1007/s11199-011-9946-z

Bolzendahl, C.I., & Myers, D.J. (2004). Feminist attitudes and support for gender equality: opinion change in women and men, 1974–1998. *Social Forces,* 83, 759–89.

Bornmann, L., Mutz, R., & Daniel, H.D. (2007). Gender differences in grand peer review: A meta-analysis. Journal of Infometrics, 1, 226-238.

Budden, A.E., Tregenza, T., Aarssen, L.W., Koricheva, J., Leimu, R, & Lortie C.J. (2008). Double-blind review favors increased representation of female authors. *Trends in Ecology and Evolution, 23* (1), 4-6. DOI: 10.1016/j.tree.2007.07.008.

Burns, T. W., O'Connor, D. J., & Stocklmayer, S. M. (2003). Science communication: a contemporary definition. *Public Understanding of Science*, *12*(2), 183-202. DOI: 10.1177/09636625030122004

Ceci, S.J., Williams, W.M., Barnett, S.M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, *135*(2). 218-261. doi: 10.1037/a0014412

Davis, S.N., & Greenstein, T.N. (2009). Gender ideology: Components, predictors, and consequences. *Annual Review of Sociology*, *35*, 87-105.

Davison, H. K., & Burke, M. J. (2000). Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior, 56*, 225–248.

Eagly, A. (1987). *Sex differences in social behavior: A social-role interpretation*, Hillsdale, NJ: Erlbaum.

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

Elsesser, K.M. & Lever, J. (2011). Does gender bias against female leaders persist? Quantitative and qualitative data from a large-scale survey. *Human Relations, 64* (12), 1555-1578. DOI: 10.1177/0018726711424323

European Commission (2012). http://www.genderandscience.org/doc/synthesis_report.pdf female and male medical faculty. *Discourse and Society, 14* (2), 191-220.

Ferber, M.A. (1988). Citations and networking. *Gender& Society, 2*(1), 82-89.

Garvey, W. D., & Griffith, B. C. (1967). Scientific communication as a social system. *Science, 157*(3792), 1011-1016.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What Do Implicit Measures Tell Us? Scrutinizing the Validity of Three Common Assumptions. *Perspectives On Psychological Science (Wiley-Blackwell)*, *2*(2), 181-193. doi:10.1111/j.1745-6916.2007.00036.x

Goldin, C. & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review, 90*(4), 715-741.DOI: 10.1257/aer.90.4.715.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4-27.

Heilman, M.E., & Haynes, M.C. (2005). No credit where credit is due: Attributional rationalization of women's success in male-female teams. *Journal of Applied Psychology, 90* (5), 905-916. DOI: 10.1037/0021-9010.90.5.905.

Jagsi, R., DeCastro, R., Griffith, K.A., Rangarajan, S., Churchill, C., Stewart, A., & Ubel, P. (2011). Similarities and differences in the career trajectories of male and female career development award recipients. *Academic Medicine, 86* (11), 1415-1421. DOI: 10.1097/ACM.0b013e3182305aa6.

Knobloch-Westerwick, S., & Glynn, C. J. (in press). The Matilda Effect—role congruity effects on scholarly communication: A citation analysis of Communication Research and Journal of Communication articles. *Communication Research.*Koenig, A.M., Eagly, A.H., Mithcell, A.A., & Ristikari, T. (2011). Are leader types masculine? A meta-analysis of three research paradigms. *Psychological Bulletin, 137* (4), 616-642. DOI: 10.1037/a0023557

Ledin, A., Bornmann, L., Gannon, F., & Wallon, G. A persistent problem: Traditional gender roles hold back female scientists. *Embo Reports, 8*(11), 982-987. DOI: 10.1038/sj.embor.7401109

Lincoln, A. E., Pincus, S., Koster, J. B., & Leboy, P. S. (2012). The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s. *Social Studies of Science, 42,* (2), 307-320. DOI: 10.1177/0306312711435830.

Martell, R. F., Lane, D. M., & Emrich, C. (1996). Male-female differences: A computer simulation. *American Psychologist, 51*, 157-158.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56-63.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474-16479.National Science Foundation [NSF]. (2006). *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*. Arlington, VA, USA: National Science Foundation

Nosek, B.A., Smyth, F.L., Sriram, N., Lindner, N.M., Devos, T., Ayala, A., et al. (2009).

National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences, 106* (26), 10593-10596.

RAND. (2005). *Is there gender bias in federal grant programs?* (RAND Infrastructure, Safety, and Environment Research Brief RB-9147-NSF). Retrieved April 1, 2010, from http://rand.org/pubs/research_briefs/RB9147/RAND_RB9147.pdf

Rossiter, M.W. (1993). The Matilda effect in science. *Social Studies of Science, 23*(2), 325-341.

Roth, P. L., Purvis, K. L., & Bobko, P. (2012). A meta-analysis of gender group differences for measures of job performance in field studies. *Journal of Management, 38*, 719-739.

Steinpreis, R.E., Anders, K.A., & Ritzke, D. (1999). The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles,* 41(7-8), 509-528. DOI: 10.1023/A:1018839203698

Tai, Z. (2009). The structure of knowledge and dynamics of scholarly communication in agenda setting research, 1996–2005. *Journal of Communication*, *59*(3), 481-513. doi:10.1111/j.1460-2466.2009.01425.x

Trix, F. & Psenka, C. (2003). Exploring the color of glass: Letters of recommendation for

Trodahl, V. C., & Powell, F. A. (1965). A short-form dogmatism scale for use in field studies. *Social Forces, 44*, 211-214.

Twenge, J.M. (1997). Attitudes toward women, 1970-1995: A meta-analysis. *Psychology of Women Quarterly, 21* (1), 35-51. DOI: 10.1111/j.1471-6402.1997.tb00099.x

Wennerås, C. & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature, 387*, 341-343.

Table 1

Pretest Results for Stimuli Selected for Main Experiment

Abstract Title	**Categorizations in %**			**Evaluation**
	Female-Typed	**Male-Typed**	**Neutral**	**(*M*, *SD*)**
Female-Typed Research Areas				
Infants' Visual Attention to Videos as a Function of Program Pacing	80	10	10	6.42 (2.33)
The (Mis)perceivers: Gender Differences in Self-Other Body Image Discrepancies and Body Dissatisfaction	60	0	40	7.00 (1.58)
Peer Pressure Among Parents? Understanding Parents' Decisions to Use Very Young Children's Television Media	82	0	18	7.00 (.96)
Correlations of Media Habits Between Generations: Parental Influence on Children's TV Watching	83	8	8	7.00 (.88)
Male-Typed Research				
Trust Games: Impact of Seller Photo and Reputation on Trust in Computer-Mediated Transactions	0	100	0	6.96 (1.26)
The Role of Communication in Political Participation: Exploring the Social Normative/Cognitive Processes Related to Political Behavior	0	100	0	6.54 (1.47)
Democratizing Journalism: Realizing the Citizen's Agenda for Local News	0	100	0	6.79 (.71)
Online Political Participation in the 2008 U.S. Presidential Election: Mobilizing or Reinforcing?	0	100	0	6.85 (1.24)
Gender-Neutral Research				
Health Communication and Face-Negotiation Theory in the Operating Room	0	0	100	6.41 (1.87)
A Mixed-Method Study Examining Patient Expectations in a Tertiary Eye Care Center	0	0	100	6.50 (2.04)
Information Seeking and the Age Disparity in Receiving Adjuvant Chemotherapy Among Colorectal Cancer Patients	0	0	100	7.46 (1.50)
Comparing Media Content Across Cultures	0	29	71	6.46 (.97)
Distracters				
The Relationship Between Sensation Seeking and Intercultural Communication Competence […]	0	0	100	4.96 (1.81)
A Transmodern Perspective on Intercultural Communicative Competence	0	0	100	5.25 (1.41)
Share or Not to Share […]: Examining Psychological Effects of Heuristic Cues on Users' Attitudes on a Product Review Website	0	0	100	6.58 (2.35)

Table 2

Author Names Assigned to Stimulus Abstracts

Male Authors	**Female Authors**
Andrew Stone	Brenda Collins
Douglas Burns	Jennifer Peters
Craig Harrison	Michelle Arnold
Jeffrey Crawford	Patricia Warren
Stephen Murphy	Melissa Jordan
Robert Phillips	Amy Bell
Timothy Kelley	Anne Cooper
Brian Stevens	Janet Henry
Matthew Webb	Lisa Sanders
James Nichols	Elizabeth Hunter
Joseph Lewis	Lori Ellis
David Hicks	Christine Russell
Gary Sullivan	Pamela Richardson
John Moore	Susan Scott
Donald Elliott	

Note. Popular first and last names from the Social Security Administration's database of people born in the 1960s were randomly selected and paired together, based on http://www.ssa.gov/OACT/babynames/decades/names1960s.html.

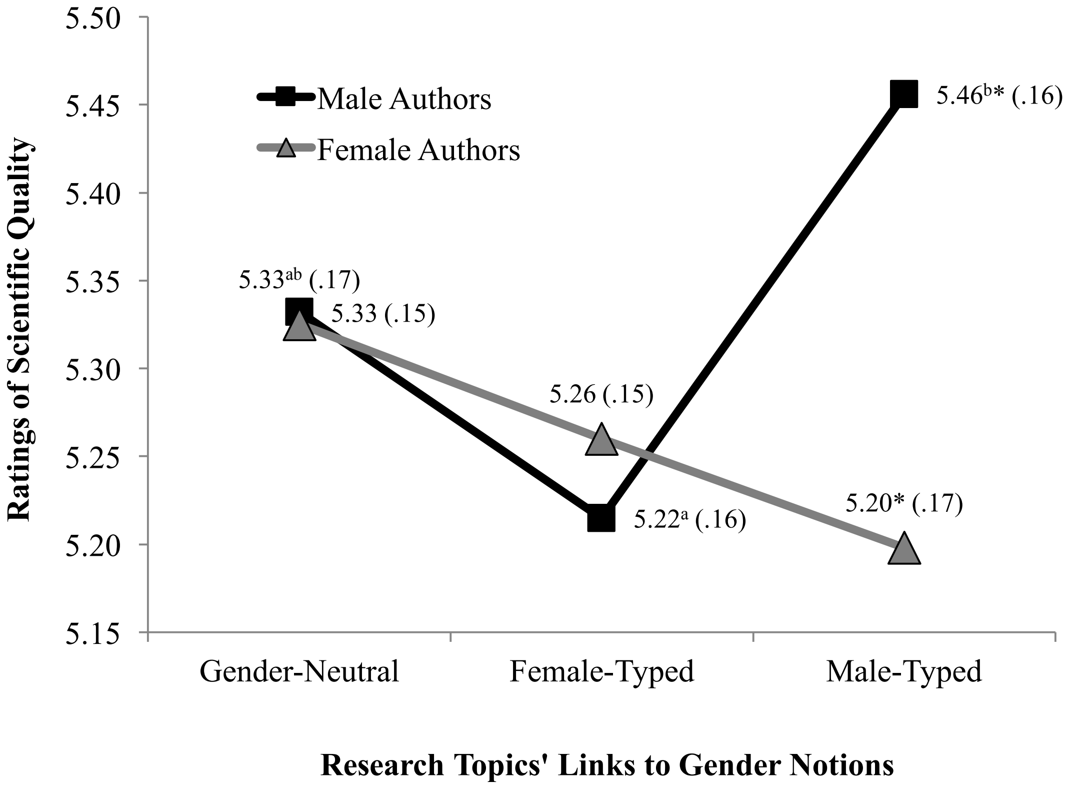

Figure 1: Perceived Scientific Quality as a Function of Research Topic and Author Gender

Note. Graph reports estimated means with standard error in brackets. Means within a research topic category with asterisks and means in a data series with different superscripts differ at $p < .05$.

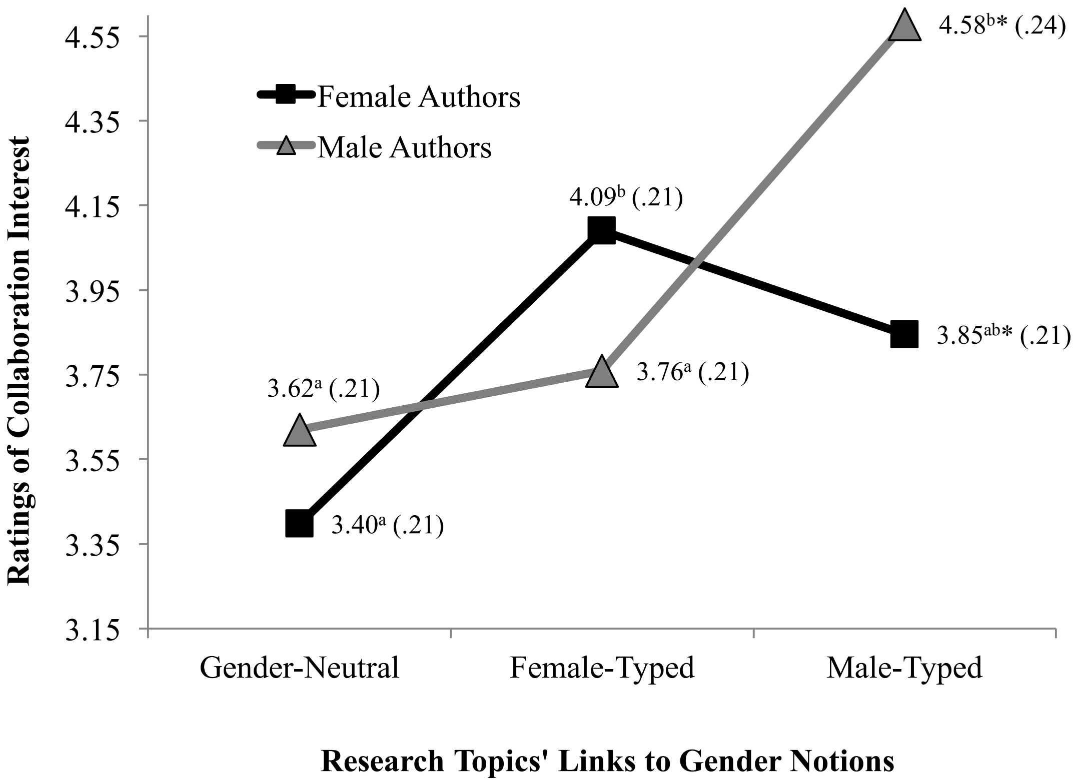

Figure 2: Collaboration Interest as a Function of Research Topic and Author Gender

Note. Graph reports estimated means with standard error in brackets. Means within a research topic category with asterisks and means in a data series with different superscripts differ at $p < .05$.