Gender bias in fame judgments: Implicit gender stereotyping or matching study phase fame?

MELANIE C. STEFFENS
Friedrich-Schiller-Universität Jena, Jena, Germany

AXEL BUCHNER
Heinrich-Heine-Universität, Düsseldorf, Germany

and

SILVIA MECKLENBRÄUKER
Universität Trier, Trier, Germany

Familiarized names are falsely judged *famous* more often than nonfamiliarized names. Banaji and Greenwald (1995) demonstrated a gender bias in this *false fame effect,* with the effect being larger for male than for female names. This effect was interpreted as reflecting the operation of a gender stereotype. However, the famous male names were, in fact, better known than the famous female names. Thus, the presence of more famous male names during study may have contributed to the observed male–famous association. If so, there should be no gender bias if the studied famous male and female names are equally famous, and a reversed gender bias should emerge if the famous female names are more famous than the male names. In two experiments, these predictions were corroborated. A "classical" gender bias was found only when the famous males were more famous than the famous females. These findings are consistent with the hypothesis that the gender bias in fame judgments, rather than showing implicit gender stereotyping in the sense of a transsituational judgment bias, reflects the fact that, in test, participants select a proportion of fame judgments to male and female names so that it matches the relative degree of fame of male and female names encountered during study.

Stereotypes encompass automatic, uncontrollable, or implicit components. Banaji and Greenwald (1995) developed a unique procedure for measuring implicit gender stereotyping. They investigated the so-called *false fame effect* that was originally obtained in the data of Neely and Payne (1983) and was extended into a research paradigm by Jacoby and colleagues (Dywan & Jacoby, 1990; Jacoby, Kelley, Brown, & Jasechko, 1989; Jacoby, Woloshyn, & Kelley, 1989; Jennings & Jacoby, 1993). A false fame effect is demonstrated if familiarized names of nonfamous persons are judged *famous* with a higher probability than are the same names when they have not been previously familiarized. Apparently, in the absence of explicit knowledge that the names' familiarity results from the study phase, familiarity is misattributed to the names' fame (Steffens, Buchner, Martensen, & Erdfelder, 2000).

According to Banaji and Greenwald (1995), the gender stereotype implies a closer cognitive association between *male* and *famous* than between *female* and *famous,* which, in turn, results in a gender bias in the process of

attributing fame: A lower criterion supposedly is used when male, as opposed to female, names are judged. Banaji and Greenwald empirically showed such a gender bias. The familiarity-induced increase in fame for nonfamous names was larger for male than for female names. In addition to the nonfamous names, however, the names of famous persons were present in the study and test phases of their experiments. Critically, "despite our effort to equate objective fame, the famous male names used in the research were better known to participants than were the famous female names" (Banaji & Greenwald, 1995, p. 195). The fact that famous male names were more famous than famous female names during the study phase may have contributed to, or may even have fully determined, the observed male–famous association if the participants simply selected the proportion of fame judgments to male and female names in test so that it matched the relative degree of fame of the male and female names encountered during study (see also Buchner, Steffens, & Berry, 2000). The gender bias might thus be found only if studied male names are more famous than studied female names. This hypothesis was tested in Experiment 1.

EXPERIMENT 1

Experiment 1 contrasted a condition in which the studied famous males were much more famous than the famous females (henceforth, the *more famous males* condi-

tion) with a condition in which the famous male and female names were selected to be equally famous (henceforth, the *equal fame* condition). These two conditions were created by holding the famous and nonfamous male names and the nonfamous female names constant across conditions, while manipulating the degree of fame of the famous female names. If the gender bias reported by Banaji and Greenwald (1995) reflected a stereotype in the sense of a transsituational cognitive structure that affects social judgments, a gender bias should be found irrespective of the relative fame of the male and the female names in the study phase. If, in contrast, participants during the test phase simply adjust their fame judgments to match the relative degree of fame of male and female names encountered during study, Banaji and Greenwald's original gender bias should be replicated only in the more famous males condition, and no gender bias should be found in the equal fame condition.

Method

Participants. The participants were 120 students at the Universität Trier. All the students took part in the experiment voluntarily and received a small incentive in exchange. Each participant took part on two occasions, separated by an interval of 48 h. Four participants were selected randomly and excluded from all the analyses, to balance for Name Sets A and B and for experimental condition (see below). Of the remaining participants, 66 were female, and 50 were male. They ranged in age from 19 to 32 years ($M = 23$).

Materials. The stimuli were 144 names of four types: famous male, famous female, nonfamous male, and nonfamous female. The 36 famous male names were selected to be quite famous. Briefly, 54 pilot study participants had rated, on a 5-point scale, the fame of the names of 725 famous persons who had achieved fame in different fields, such as politics, arts, or sports. The instructions specified that a rating of 1 would indicate that *a name seems nonfamous*, a rating of 3 that *one knows a name but does not know what the person has done to become famous*, and a rating of 5 that *a name is definitely famous*. Examples of the names of famous males used in Experiment 1 (average fame rating, $M = 3.48$) are Franz Vranitzky, Alwin Schockemöhle, and Gustav Klimt. In the more famous males condition, 36 female names were selected to be much less famous ($M = 1.5$); examples are Mary Robinson, Lina Wertmüller, and Miriam Makeba. In the equal fame condition, 36 female names were selected that were as famous as the 36 famous male names ($M = 3.48$). Examples are Dolly Parton, Utta Danella, and Doris Lessing.

Nonfamous names were constructed to parallel the famous names with regard to gender, ethnicity, number of letters, and repetition of first letters in first and last names. Each first or last name occurred only once in the entire name list. Examples of nonfamous names are Norbert Mail, Ricardo Angeli, Erwin Teschke, Bärbel Wolferding, Betty Hruby, and Michelle Molitor. The 36 names in each category were randomly divided into Sets A and B.

The 72 names in each set (18 famous male, 18 famous female, 18 nonfamous male, and 18 nonfamous female names) were brought into a random order, numbered, and printed on the study phase sheet with a 5-point scale next to each name, on which pronounceableness was to be rated in an effort to veil the purpose of the name presentation. On the study phase sheet, age, sex, and an individual code for each participant were to be indicated. The code consisted of a combination of the participants' and their parents' name letters and birthday numbers, in order to allow for the correct assignment of test lists to conditions. In addition, the study phase sheet contained the instructions for the pronounceableness rating. These instructions stressed that the participants were expected to work through the list quickly.

The test sheets contained 146 names, arranged on two pages, each accompanied by two boxes in which to check whether the name was famous or nonfamous. The first two names, Thomas Gottschalk and Claudia Schiffer, were buffers, and both were very famous in Germany. We encouraged the participants to use familiarity as the judgment criterion, as is typical in false fame experiments (see Jacoby, Kelley, et al., 1989; Jacoby, Woloshyn, & Kelley, 1989). The 144 critical names (all the names in Sets A and B) were arranged in a random order. Half of them were from the study phase and, hence, were familiarized for a given participant, and half were not. Again, age, sex, and the individual code had to be indicated on the test sheet.

Procedure. Students were invited to take part in the experiment if they were able to return for the test phase 2 days later. Study phase sheets were handed to the students, and they were to sit down quietly and fill out their questionnaires—that is, rate the names for pronounceableness. The procedure was repeated 2 days later in the test phase, with the exception that fame ratings were required. After that, the participants were debriefed.

Design. Independent variables were the relative fame of the male and female names (more famous males or equal fame; between subjects), as well as name gender (male or female; within subjects) and name familiarization (familiarized or nonfamiliarized; within subjects). The fame judgments of the names represented the dependent variable. Study phase name set (A or B) was balanced.

Results

Hit rates (*famous* responses to famous names) and false alarm rates (*famous* responses to nonfamous names) were computed for familiarized male, familiarized female, nonfamiliarized male, and nonfamiliarized female names. The Type I error probability α was set at .05 for all the analyses that follow. As a measure of effect size, R_p^2 is reported (Cohen, 1977), which represents the proportion of variance explained by a given variable, relative to the variance not explained by any other variable in the design. Preliminary analyses showed that participant sex did not qualify any of the reported effects. Therefore, this factor is not included in the analyses of Experiments 1 and 2.

Hits. In the more famous males condition, the hit rates for male names look much the same as in the equal fame condition, but the hit rates for female names are much lower (see Figure 1). This indicates that the fame manipulation was successful, in that the selected famous male names were, indeed, much more famous than the famous female names. However, despite our efforts to select equally famous males and females for the equal fame condition, the hit rates were somewhat higher for male than for female names. In both conditions, the hit rates for familiarized names were higher than the hit rates for nonfamiliarized names.

A $2 \times 2 \times 2$ analysis of variance (ANOVA) showed main effects of relative fame of male and female names [$F(1,114) = 40.24$, $R_p^2 = .26$], of name gender [$F(1,114) = 356.30$, $R_p^2 = .76$], and of name familiarization [$F(1,114) = 41.00$, $R_p^2 = .26$; all other Fs ≤ 2.41]. Importantly, the name gender effect depended on the relative fame [$F(1,114) = 200.45$, $R_p^2 = .64$]. Analyses of simple main effects showed an effect of name gender in the more
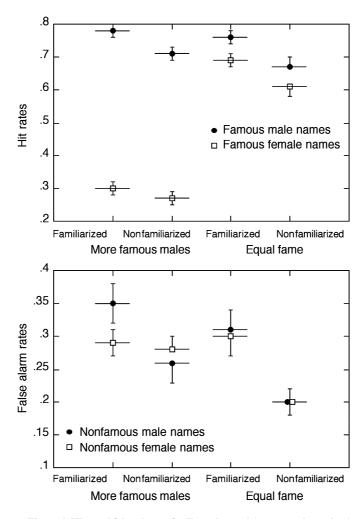
Figure 1. Hits and false alarms for Experiment 1 (means and standard errors of the mean) for familiarized male and female names and nonfamiliarized male and female names in two experimental conditions, the more famous males and the equal fame conditions.

famous males condition [$F(1,114) = 545.63$, $R_p^2 = .83$] and also in the equal fame condition [$F(1,114) = 11.13$, $R_p^2 = .09$]. A comparison of the effect sizes shows that the fame difference between male and female famous names was larger in the more famous males condition.

False alarms. The lower panel of Figure 1 shows that the false alarm rates were higher for familiarized than for nonfamiliarized names, which represents the well-known false fame effect. More important, this false fame effect seems to have been larger for male than for female names only in the more famous males condition. Indeed, a 2 × 2 × 2 ANOVA showed the expected interaction between relative fame, name gender, and name familiarization [$F(1,114) = 3.94$, $R_p^2 = .03$]. In addition, there was a false fame effect [$F(1,114) = 57.90$, $R_p^2 = .38$] that interacted with relative fame [$F(1,114) = 7.19$, $R_p^2 = .06$] and with name gender [$F(1,114) = 6.13$, $R_p^2 = .05$].

A separate analysis of the more famous males condition showed a false fame effect [$F(1,57) = 12.06$, $R_p^2 = .17$]. This false fame effect—the *increase* in assigned fame due to familiarization—was clearly larger for male ($M = .09$) than for female ($M = .01$) names [$t(57) = -2.84$, $R_p^2 = .12$], replicating the gender bias in the false fame effect reported by Banaji and Greenwald (1995). A separate analysis of the equal fame condition also showed an overall false fame effect [$F(1,57) = 53.32$, $R_p^2 = .48$], but this false fame effect was the same size for male ($M = .11$) and female ($M = .10$) names [$t(57) < 1$]. Thus, there was no gender bias in the equal fame condition.

Discussion

The results in the more famous males condition show that when the famous male names were clearly more famous than the famous female names, a gender bias in the false fame effect could be observed. When, as in the equal fame condition, the famous male names were only slightly better known to the participants than the famous female names, no gender bias was observed in the false fame ef-

fect. To be sure, the male-to-female fame manipulation in the more famous males condition was quite drastic, even to the degree of statistically eliminating the false fame effect for the nonfamous female names, indicating that we may have overshot a little with this manipulation. It is therefore important that when the fame associated with famous male and female names was similar, a false fame effect of equal magnitude occurred for both male and female names.

From a theoretical viewpoint, it is critical that the pattern of results observed in this experiment should be expected if, in test, participants simply select a proportion of fame judgments to male and female names so that it matches the relative degree of fame of the male and female names encountered during study. This pattern is unexpected if the gender bias results from some transsituational gender stereotype. The matching hypothesis makes the additional prediction that a reversed gender bias should emerge if studied females names are more famous than studied male names. Buchner et al. (2000) did not corroborate this hypothesis. However, their materials and design differed from Banaji and Greenwald's (1995) in important respects. For instance, Buchner et al. used photographs of moderately famous and, occasionally, very famous men and women, but they did not include nonfamous people. Therefore, Experiment 2 was designed to test the prediction of a reversed gender bias when studied female names are more famous than studied male names, while at the same time being modeled closely after the original experiments of Banaji and Greenwald and Experiment 1.

EXPERIMENT 2

If, as is predicted by the matching hypothesis, the gender bias in fame judgments depends on the different degrees of fame of studied male and female names, a reversed gender bias should be observed if the female names presented during study are more famous than the male names.

Method

Participants. The participants were 85 students at the Universität Trier, who attended a lecture on two subsequent occasions, 24 h apart. They received a small incentive. In order to balance study phase Name Sets A and B among the participants, the data of 15 randomly drawn participants were excluded from all the analyses. Of the remaining participants, 54 were female, and 16 were male. They ranged in age from 18 to 28 years ($M = 21$).

Materials. The stimuli were identical to those in Experiment 1, with the following exceptions. The sets of famous names consisted of 36 names of famous females (average fame rating, $M = 3.90$) and of 36 names of less famous males ($M = 3.60$). Examples are Camille Claudel, Isabel Allende, Arabella Kiesbauer, Roald Dahl, Lion Feuchtwanger, and Georg Büchner. The buffer names on the test sheets were Marilyn Monroe and Helmut Kohl.

Procedure. The procedure was identical to that in Experiment 1, with the following exceptions. The experiment was carried out as a group experiment. At the beginning of a lecture, study phase sheets were handed to all the students in the class. They were asked to complete the sheets in exchange for a small incentive. Study phase sheets were afterward collected, and the students learned that the incentives would be distributed later. After 24 h, test sheets were distributed at the beginning of another lecture, and the students were asked to follow the instructions. The students were debriefed after the lecture, and incentives were handed out.

Design. Independent variables were name gender and name familiarization (within subjects). The fame judgments of the names represented the dependent variable. Study phase name set (A or B) was balanced.

Results

Hits. As is shown in Figure 2, fame was more readily assigned to familiarized than to nonfamiliarized famous names. More important, the participants assigned fame more readily to famous female than to famous male names. A 2×2 ANOVA on the hit rates revealed an effect of familiarization [$F(1,69) = 31.44$, $R_p^2 = .31$] and of name gender [$F(1,69) = 46.22$, $R_p^2 = .40$] and no interaction [$F(1,69) < 1$]. These results corroborate that famous female names were more famous than famous male names, but the magnitude of the fame difference was considerably smaller than in the more famous males condition in Experiment 1. Therefore, we would expect the reversed gender bias in the false fame effect to be smaller than the gender bias was in that condition.

False alarms. The false fame effect, the increase in assigned fame due to familiarization, was larger for female ($M = .09$) than for male ($M = .05$) names. This difference was statistically significant [$t(69) = 1.88$, $R_p^2 = .05$], albeit relatively small in terms of the standardized effect size. Aside from this critical result, there was a false fame effect [$F(1,69) = 29.55$, $R_p^2 = .30$], and fame was more readily assigned to female than to male nonfamous names [$F(1,69) = 15.25$, $R_p^2 = .18$].

Discussion

Famous female names were more famous than male names, although the difference was clearly not as large as the fame advantage of the famous male over the famous female names in the more famous males condition in Experiment 1, in which we had observed the "classical" gender bias in the false fame effect. Even so, a small but statistically significant *reversed* gender bias was found. A gender stereotype in the form of a male–famous association is not a sufficient explanation for this pattern of results. The presence of more famous study phase names of one gender seems to be a crucial constituent of the observed gender bias.

GENERAL DISCUSSION

Banaji and Greenwald (1995) reported the "classical" gender bias in terms of a familiarity-induced increase in judged fame that was larger for nonfamous male than for nonfamous female names. The present research replicates and extends their finding by showing that the di-
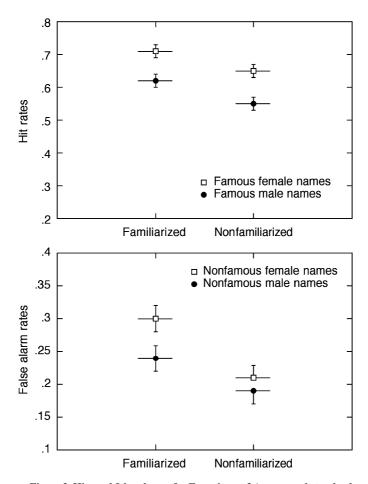
Figure 2. Hits and false alarms for Experiment 2 (means and standard errors of the mean) for familiarized male and female names and nonfamiliarized male and female names.

rection of the gender bias depends on the relative fame of the male and female names presented *within the context* of the experiment. This context-dependent gender bias—the signed difference between the fame increase for nonfamous male and nonfamous female names—is illustrated in Figure 3. We replicated Banaji and Greenwald's "classical" gender bias, but only when, during study, the famous male names were clearly more famous than the famous female names. Two new findings of our experiments were, first, no gender bias when the fame of the famous studied male and female names was very similar and, second, even a reversed gender bias in fame judgments when the famous studied female names were more famous than the male names.

Taken together, this pattern of findings is compatible with the matching hypothesis—that is, the hypothesis that participants simply assign fame to male and female names during test to match the relative fame of the famous male and female names encountered during study. The present results pose problems for Banaji and Greenwald's (1995) original explanation of the gender bias in terms of gender stereotyping—that is, a transsituational cognitive structure that affects social judgments. Rather,

our findings imply that males are more often misjudged famous in contexts in which there are more famous males than famous females, whereas females are more often misjudged famous in contexts in which there are more famous females. Our notion is, however, that currently, in most societies, there are only few contexts with more famous females than males. We cannot exclude, in principle, the possibility of a transsituational gender bias on top of the situation-induced gender bias that we found. However, the finding of no gender bias in the equal fame condition in Experiment 1 raises problems for Banaji and Greenwald's explanation. A more general implication of the present findings is the need to carefully evaluate whether an empirically found bias is due to a transsituational (stereotyping) process or, rather, could have been induced by the specific situational circumstances.

Although the data pattern reported here seems relatively clear-cut, we want to express our concerns that this may not be representative for research on the gender bias in the false fame effect. First of all, the false fame effect reported by Banaji and Greenwald (1995) was already quite small, with about half a name (out of 18) for female names and about 1 name for male names, and so
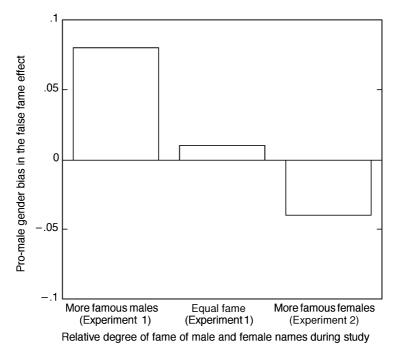
Figure 3. Mean gender bias in the false fame effect for the two conditions in Experiment 1 and for Experiment 2. See the text for details.

was the gender bias *within* this small false fame effect. In fact, the gender bias was statistically significant in only two out of four experiments. It should thus not be surprising if this effect was difficult to replicate. Indeed, Buchner and Wippich (1996, Experiment 1) did not replicate the gender bias in the false fame effect, and when they found a gender bias in their Experiment 2, it depended on the sex role orientation of their participants (but not on their sex). Other, as yet unpublished results from our lab (Steffens et al., 1999) also nourish the suspicion that the gender bias in fame judgments may be a rather fleeting phenomenon, which, when found, may take different forms, depending on the precise study and test phase materials used (for details, see Buchner et al., 2000).

The present results are in line with those of Steffens, Mecklenbräuker, Buchner, and Mehl (2004), who found that the direction of the gender bias mirrored the *relative number* (rather than the relative degree of fame) of female and male names learned prior to the experiment: When participants had been exposed to a larger number of famous males than famous females, they later exhibited a false fame effect favoring males. In contrast, when they had initially been exposed to a larger number of famous females, they later showed a false fame effect favoring females. In detail, fame was more readily assigned to all nonfamous names of the gender overrepresented in the list learned before the experiment. A female–famous association was thus established as easily as a male–famous association.

Given Steffens et al.'s (2004) finding and the present data pattern, we think it reasonable to conclude that the gender bias in fame judgments reflects a flexible match-ing process that enables quick adaptations to the characteristics of different judgment situations, rather than the operation of a transsituational gender stereotype.

REFERENCES

BANAJI, M. R., & GREENWALD, A. G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of Personality & Social Psychology*, **68**, 181-198.

BUCHNER, A., STEFFENS, M. C., & BERRY, D. C. (2000). Gender stereotyping and decision processes: Extending and reversing the gender bias in fame judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1215-1227.

BUCHNER, A., & WIPPICH, W. (1996). Unconscious gender bias in fame judgments? *Consciousness & Cognition*, **5**, 197-220.

COHEN, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). Hillsdale, NJ: Erlbaum.

DYWAN, J., & JACOBY, L. [L.] (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. *Psychology & Aging*, **5**, 379-387.

JACOBY, L. L., KELLEY, C., BROWN, J., & JASECHKO, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality & Social Psychology*, **56**, 326-338.

JACOBY, L. L., WOLOSHYN, V., & KELLEY, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, **118**, 115-125.

JENNINGS, J. M., & JACOBY, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology & Aging*, **8**, 283-293.

NEELY, J. H., & PAYNE, D. G. (1983). A direct comparison of recognition failure rates for recallable names in episodic and semantic memory tests. *Memory & Cognition*, **11**, 161-171.

STEFFENS, M. C., BUCHNER, A., MARTENSEN, H., & ERDFELDER, E. (2000). Further evidence on the similarity of memory processes in the process dissociation procedure and in source monitoring. *Memory & Cognition*, **28**, 1152-1164.

STEFFENS, M. C., MECKLENBRÄUKER, S., BUCHNER, A., BACHEM, P., DECKER, C., KLEIS, A., LUX, V., & PHILIPPI, R. (1999). Haben implizite Geschlechterstereotype einen Einfluß auf Berühmtheitsurteile? [Do implicit gender stereotypes influence fame judgments?] In A. Widmann (Ed.), *Experimentelle Psychologie; Beiträge zur 41. Tagung experimentell arbeitender Psychologen, Leipzig, 28. März–1. April 1999* (p. 299). Lengerich, Germany: Pabst.

STEFFENS, M. C., MECKLENBRÄUKER, S., BUCHNER, A., & MEHL, B. (2004). On the bounded rationality of implicit gender stereotyping in fame judgments. *European Journal of Social Psychology*, **34**, 397-406.