Running Head: GENDER DIFFERENCES IN HUMOR

Who's Funny: Gender Stereotypes, Humor Production, and Memory Bias

Laura Mickes[1], Drew E. Hoffman[1], Julian L. Parris[1], Robert Mankoff[2] and Nicholas Christenfeld[1]

[1]University of California, San Diego, [2]New Yorker Magazine[TM]

Keywords: humor effect; humor; stereotype; gender; source memory

Address Correspondence to: Laura Mickes, Ph.D.
UC San Diego
9500 Gilman Drive
La Jolla, CA 92093-0109
Tel: (858) 822-3264
Email: lmickes@ucsd.edu

Abstract

It has often been asserted, by both men and women, that men are funnier. We explore two possible explanations for such a view, first testing whether men, when instructed to be as funny as possible, write funnier cartoon captions than do women, and second examining whether there is a tendency to falsely remember funny things as having been produced by men. Thirty two participants, half of each gender, wrote captions for 20 cartoons. Raters then indicated their humor success. Raters of both genders found the captions written by males funnier, though this preference was significantly stronger among the male raters. In the second experiment, male and female participants were presented with the funniest and least funny captions from the first experiment along with the caption author's gender. On a memory test, both females and males disproportionally misattributed the humorous captions to males and the nonhumorous captions to females. Men may think men are funnier because they actually find them so, and while women rated the captions written by males slightly higher, our data suggest they may regard men as funnier more because they falsely attribute funny things to them.

Who's Funny: Gender Stereotypes, Humor Production, and Memory Bias

It seems a truth, while perhaps not universally acknowledged, at least widely shared, that men are funnier than women (e.g., Lewis, 2000). Such a view has been expressed by men and women, and often in conjunction with firm assertions that men's humor advantage, if such it be, is not part of any general intellectual superiority (Hitchens, 2007; Greer, 2009). Presuming a reliable gender difference in humor production, various theories have been offered, including suggestions that humor, like the head butting of elk, is done to impress potential mates (Bressler, Martin, & Balshine, 2006). Consistent with such a notion, females indicate a preference for mates who makes them laugh, whereas males prefer a mate who laughs at their humor (Li, et al., 2009). There is also evidence that both genders comply, with women laughing more, and men making people laugh more (Provine, 2000, p. 27; but see Kothoff, 2006). However, this evidence does not require that men actually are more capable of being funny, as it could be due to some combination of emotional responsivity, differential effort, and pity. There are no direct tests of assertions about gender differences in the ability to be funny.

In this paper we explore explanations for the impression that men are funnier than women. It could be that the stereotype exists because it is true, and people have correctly observed the world. The impression could also exist without its being true, if people's view of the world is systematically biased. In two studies, we explore these two possibilities. In the first study, male and female participants wrote, or at least tried to write, funny captions to accompany cartoon images, and raters, also male and female, evaluated their success. The second study examined whether, in the context of a memory experiment, people would be more likely to recall

funny things as having been produced by men. Previous work has shown that stereotypes, about

occupations for example, can influence memory (e.g., Mather, Johnson, & De Leonardis, 1999;

Marsh, Cook, & Hicks, 2006). People giving credit for funniness to men because they expect

men to be funny, while it could not explain the origin of the stereotype, could explain its

perpetuation.

Study 1

Phase 1: Humor Production

Participants

Thirty-two undergraduates (16 male and 16 female) from the University of California,

San Diego (UCSD) participated for course credit.

Materials

Twenty New Yorker Magazine^{TM} cartoons from their caption contest were compiled

together with a questionnaire that asked gender as well as other demographic questions.

Procedure

After viewing two sample New Yorker Magazine^{TM} cartoons with captions, participants

wrote a caption for each of the 20 cartoons in their packet. Participants were encouraged to be as

funny as possible, and were told that others would be rating their captions. They worked alone in

a quiet room, and were given 45 minutes to produce the 20 captions. Afterwards, they completed

a questionnaire which asked them how funny they thought others would find their captions, and

also whether they thought men or women were funnier.

Phase 2: Humor Rating

Participants

Eighty-one UCSD undergraduates (34 male and 47 female) received class credit for participating.

Materials

The twenty New Yorker Magazine[TM] cartoon images with new captions written by the 32 participants in Phase 1 were used. The stimuli were presented on a monitor and participants responded with keyboard presses.   These participants also indicated at the end whether they thought men or women were funnier.

Procedure

Participants rated the funniness of the captions using a five-round knock-out tournament style system. In a given round of a tournament, one cartoon image was displayed with two corresponding captions, as shown in Figure 1. At their own pace, raters chose the funnier of the two captions with a key-press. Next, two new captions appeared. This process was repeated for all 32 captions corresponding to each cartoon. The 16 captions chosen as winners in the first round were randomly pitted against each other in the next round. By the end, participants made 620 choices. How many rounds the writer's caption survived before being knocked out determined the humorousness score that writer earned during that tournament from that rater, with one point for each round their caption won; thus, a writer could earn zero to five points per tournament. This approach may be able to provide a more sensitive measure of relative funniness than Likert-type ratings, especially since many captions were crowded at the not-

funny-at-all end of the continuum.

For each tournament, a preference score was computed by subtracting the average points the rater allocated to female writers from the average points allocated to male writers. An average over the 20 tournaments of these preference scores was calculated for each rater, providing a measure of the degree to which that rater preferred captions produced by male writers across the entire experiment. A score of 0 for any particular rater thus indicates an absence of average preference for either male or female produced humor, and any positive deviation from 0 indicates a relative preference for humor produced by males.

Results

Questionnaire

Participants were in broad agreement with the stereotype that men are funnier, with (aggregating across writers and raters) 89% of the woman so indicating, and 94% of the men (1 female and 2 males did not respond to the question). This rate of endorsing the stereotype is not significantly different by gender, χ^2 (1, N=110) = .58, p > .45. The caption writers also indicated that this would apply to them, with the male writers predicting an average funniness for their own creations of 2.3 (on a 1-5 scale), while the female writers predicted a significantly more modest 1.5 (t(30) = 3.50, p = .001).

Humor Ratings

To assess the degree to which raters exhibited a preference for captions written by males or females, a one-way analysis of variance was computed on the preference scores calculated for each rater, with rater gender as a between-subjects factor. Because an average preference score

of zero indicates no average rater preference for males or female writers, the relevant analysis is the test of the model intercept. Raters on average do show a preference for male writers, $t(79) = 7.85$, $p < .001$, allocating 0.11 more points to them ($SD = 0.46$, $d = 0.24$). Both female and male raters show this preference for male writers, although it is significantly greater among male raters, $t(79) = 3.63$, $p < 0.001$. Female raters allocated male writers on average 0.06 more points ($t(46) = 3.22$, $p < 0.01$, $d = 0.13$); male raters allocated them on average 0.16 more points ($t(33) = 7.66$, $p < 0.001$, $d = 0.35$).

Content Analysis

Given differences, albeit slight, in the success of male and female writers, and the differences, again slight, in their success with male and female raters, we performed a rough investigation of the content of the cartoon captions. Two research assistants independently flagged each of the 640 captions for the presence of each of 25 different categories of content, drawn from theories of humor (e.g., puns, self-deprecation (Long & Graesser, 1988); benign violations (McGraw & Warren, 2010)). With the data from these two coders, aggregate measures of category usage were calculated for male and female authors. For only two categories were there gender differences: male authors use more sexual humor and profanity. The base usage of profanity and sexual humor was low, so these categories were combined to calculate for each of the 32 caption authors a percentage of their 20 captions that used either. Male authors used these categories of humor in more of their captions ($t(30) = 2.17$, $p = .038$, males $= 4.30\%$, females $= 1.95\%$). This modest difference in usage, however, does not appear to explain the humor advantage for males, and in no follow-up analysis was there any humor advantage for such

captions, nor were they preferred by male raters.

Discussion

These results are consistent with the widespread belief that men are funnier than women: both male and female raters judged captions written by men to be funnier. Males demonstrated an even stronger preference for captions written by males, indicating a uniquely strong appreciation of male humor by male raters. While males did, in a way that might surprise few, produce more profanity and sexual content, this was not the basis of their slightly greater humor success, nor of their slightly greater appeal to men.

Experiment 2

This experiment was designed to test the idea that memory is affected by the belief that males are funnier than females. Stereotypes can impact source attributions (e.g., Bayen, Nakamura, Dupuis, & Yang, 2000; Hicks & Cockman, 2003, Marsh, Cook & Hicks, 2006), and here we explore whether people will tend to recall funny things as having been produced by men rather than women. Such errors could enable people to maintain the view that men are funnier, even in a world where they are not.

Method

Participants

Seventy-two UCSD undergraduates (36 male and 36 female) participated for class credit.

Materials

Study items comprised the 20 cartoons from the first study, coupled with 100 captions selected from those generated by the Study 1 participants. These were the 50 captions that the raters had rated as most humorous (25 written by females and 25 written by males), and the 50 that had been rated as least humorous (25 written by females and 25 written by males). The cartoons were presented on a monitor. Each image was presented multiple times, but captions were only presented once. Below the image and caption appeared the caption writer's gender. In total, there were 80 targets (shown during presentation and test) and 20 lures (shown only during test) that were balanced for humorousness and writers' gender.

Procedure

Participants were told that the same cartoon images would appear more than once and with different captions. Participants were instructed to remember the captions and the writers' gender for a memory test. The 80 targets appeared individually for 12 seconds, in random order. Following a 2-minute distractor task of simple math problems, participants took the 100-item memory (80 targets and 20 lures) test.

In the memory test, participants first made an old/new decision, indicating whether they believed a given caption had been displayed before. Participants key-pressed "o" to indicate "old" (the correct response for targets), or "n" for "new" (the correct response for lures). This allows an examination of whether funny captions are better remembered. Next, regardless of the old/new response, participants made a source (writer's gender) decision, by pressing "f" for female and "m" for male. This allows an examination of whether funny captions are remembered as having been produced by men.

Results

Old/New and Source Accuracy

Because female and male participants did not differ significantly in their accuracy scores, the data were combined to analyze overall old/new and source accuracy. The proportion correct (88%) on the old/new decision was significantly above chance, $t(71) = 38.04$, $p < .001$. The average d' (a measure of sensitivity, which is zero at chance; Macmillan & Creelman, 2005) = 2.98, $sd = .71$. The proportion correct (65%) on the source decision was also significantly above chance, $t(71) = 13.42$, $p < .001$. The average d' for source decisions was .45, $sd = .38$.

The Humor Effect

For both new/old judgments as well as source memory, there was a significant effect of humor (as measured by d'), with the funny captions ($M = 2.95$, $sd = .73$) remembered better than the non funny captions ($M = 2.63$, $sd = .65$), $t(71) = 4.63$, $p < .001$, and the source of the funny captions ($M = .57$, $sd = .54$) remembered better than the non funny captions ($M = .34$, $sd = .51$), [$t(71) = 2.85$, $p = .006$].

The Effect of Humor on Source Memory

For both funny and non funny target captions, the proportion of items for which the correct gender source was identified was calculated. This provides conditional source identification measure (CSIM) scores, and is the standard way to assess response bias (Murnane & Bayen, 1996). For example, for the funny captions the number correctly identified as male was divided by the number of funny captions written by males and females. To illustrate further, say one subject correctly identifies 18 out of the 20 targets that were funny captions written by females.

This subject attributed 12 out of the 18 to a female writer. Thus, this subject's CSIM score for funny targets written by females would be .67. This was repeated for each subject for each humor type and gender of the caption writer. A three-way ANOVA with subjects as random effects was computed using these CSIM scores with writer gender, humor type, and subject gender as factors. As shown in Figure 2, participants preferentially attributed the funny captions, but not the unfunny ones, to males. The interaction was significant, $F(1,210) = 9.29$, $p = .003$. Follow-up t-tests indicate that the non funny captions were more often attributed to females ($M = .68$, $sd = .14$) compared to males ($M = .63$, $sd = .17$), $t(210) = 2.44$, $p = .016$, and there was a trend in attributing the funny captions to males ($M = .70$, $sd = .14$) compared to females ($M = .66$, $sd = .14$), $t(210) = 1.87$, $p = .063$. The degree of memory distortion was not significantly different for the male and females raters ($F(1,210) = 2.36$, $p = .126$).

The same source attribution analysis was conducted on the 20 new items. For both funny and non funny lure captions, the proportion of items for which each gender source was identified was calculated. The interaction humor type and gender of writer was significant, $F(1,210) = 4.10$, $p = .044$. Follow-up t-tests indicate no difference in the attribution of non funny captions to females ($M = .49$, $sd = .28$) compared to males ($M = .50$, $sd = .26$), $t(210) = 0.13$, $p = .899$, but did show funny captions were more often attributed to males ($M = .61$, $sd = .26$) than compared to females ($M = .47$, $sd = .26$), $t(210) = 2.99$, $p = .003$.

Discussion

Humorous items often afford a memorial benefit. The findings replicate the humor effect (Kaplan & Pascoe, 1977; Schmidt, 1994, 2002; Schmidt & Williams, 2001; Takahashi & Inoue,

2009), with the funny captions not only remembered better, but also with the gender of their

authors better remembered. The analyses also provide evidence for a humor-based retrieval bias;

individuals of both genders tend to misattribute humorous captions to male writers. This was true

both for misremembering captions whose author gender they had seen and for attributions for

new captions whose author gender they were only guessing. This finding is consistent with

previous research on stereotypes influencing source memory decisions (e.g., Hicks & Cockman,

2003).

<div align="center">Conclusion</div>

We explored two possible components underlying the stereotype that males are funnier,

and both received some empirical support. Men, at least under the conditions and constraints of

the present experimental situation, were funnier. In addition, funny captions were preferentially

attributed to male authors, a bias that was present in both men and women.

The data are not entirely consistent with a view of male humor being favored

evolutionarily to impress women (Miller, 2000), as male humor (perhaps like the male sports

car) appeals most especially to other men. Nor are they consistent with humor being gender

specific, as we found that, while men prefer male humor, women, albeit very slightly, also prefer

it. Our data did reveal slight differences in the content of the humor produced by men and

women, with men inclined slightly more to sex and profanity, but the factors that made their

captions slightly more successful, and especially so to other men, were elusive.

The data do suggest that men's view that men are funnier could be a result of their

actually finding the humor they produce funnier, and also their biased recall of funny things as

having sprung from the mind of man. Women, perhaps slightly less strong in their conviction that men are funnier, also showed less of an effect of actually finding them funnier, though they did show equally biased recall. Women laughing more at men (Provine, 2000), when the gender is known, may be due largely not to superior humor, but to more subtle social influences, which are known to impact laughter (e.g., social status (Coser, 1960; Robinson & Smith-Lovin, 2001)).

Other factors will likely contribute to the impression of male funniness. In our first study, everyone was asked to be funny, and it could be that men would spontaneously regard more occasions as appropriate for humor, because they feel it is expected of them, to impress women, or because they are less cautious about hurt feelings. It could also be that our caption contest required a style of humor at which women are relatively adept, and that other domains would produce different results. Arguing against these points, however, we did find dramatic difference between the male and female caption authors in their predictions of success. Male confidence, in this domain at least, does seem to outstrip male competence.

References

Bayen, U. J., Nakamura, G. V., Dupuis, S. E., & Yang, C. (2000). The use of schematic

knowledge about sources in source monitoring. *Memory & Cognition, 28,* 480-500.

Bressler, E. R., Martin, R., & Balshine, S. (2006). Production and appreciation of humor as

sexually selected traits. *Evolution & Human Behavior, 27,* 121-130.

Coser, R. L. (1960). Laughter among colleagues. A study of the social functions of humor among

the staff of a mental hospital. *Psychiatry, 23,* 81-95.

Greer, G. (2009, March 2). Beaten to the punchline. *The Guardian*, Retrieved from

http://www.guardian.co.uk/stage/2009/mar/02/germaine-greer-comedy-women.

Hicks, J., & Cockman, D. (2003). The effect of general knowledge on source memory and

decision processes. *Journal of Memory and Language, 48*, 489-501.

Hitchens, C. (2007, January). Why women aren't funny. *Vanity Fair*, p. 54. Retrieved from

http://www.vanityfair.com/culture/features/2007/01/hitchens200701.

Kaplan R. M., & Pascoe, G. C. (1977). Humorous lectures and humorous examples: Some

effects upon comprehension and retention. *Journal of Educational Psychology, 69,* 61–

65.

Kothoff, H. (2006). Gender and humor: the state of the art. *Journal of Pragmatics, 38,* 4-25.

Lewis, J. (2000, June 8). Aspen's U.S. Comedy Arts Festival. Aspen, CO.

Li, N. P., Griskevicius, V., Durante, K. M., Jonason, P. K., Pasisz, D. J., & Aumer, K. (2009).

An evolutionary perspective on humor: Sexual selection or interest indication?

Personality and Social Psychology Bulletin, 35, 923-936.

Long, D. L. & Graesser A. C. (1988). Wit and humor in discourse processing, *Discourse Processes, 11,* 35–60.

Marsh, R. L., Cook, G. I., & Hicks, J. L. (2006). Gender and orientation stereotypes bias source-monitoring attributions. *Memory, 14,* 148-160.

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology, 16,* 437-458.

McGraw, A. P. & Warren, C. (2010). Benign violations: Making immoral behavior funny. *Psychological Science*, *21*, 1141-1149.

Mickes, L., Hoffman, D. E., Mankoff, R., Jones, J. J., & Wixted, J. T. *I remember him (but not her) being funny.* Poster presented at the Psychonomic Society 50th Annual Convention, Boston, MA, November 19-22, 2009).

Miller, G. F. (2000). *The Mating Mind*. New York: Random House.

Murnane, K. & Bayen, U. J. (1996). An evaluation of empirical measures of source identification. *Memory and Cognition, 24,* 417–42.

Provine, R. R. (2000). *Laughter: A Scientific Investigation.* New York: Viking.

Robinson, D. T. & Smith-Lovin, L. (2001). Getting a laugh: Gender, status, and humor in task discussions. *Social Forces, 80,* 123-158.

Schmidt, S. R. (1994). Effects of humor on sentence memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 953-967.

Schmidt, S. R. (2002). The humour effect: Differential processing and privileged retrieval. *Memory*, *10,* 127-138.

Schmidt, S. R., & Williams, A. R. (2001). Memory for humorous cartoons. *Memory &*

   Cognition, 29, 305-311.

Takahashi, M., & Inoue, T. (2009). The effects of humor on memory for non-sensical pictures.

   Acta Psychologica, 132, 80-84.

Figures

Figure 1. An example of the tournament style rating system where two captions were pitted against one another. Participants chose the funnier of the two. (In this case, the top caption was the overall winner, and the bottom one was most often eliminated in the first round.)

Figure 2. Conditional Source Identification Measure (CSIM) scores by humor type (with standard errors). Higher CSIM scores reflect greater attribution of authorship.