

Social Judgments, Social Media, and Self-Deprecation:

Role of Information Source and Valence on Trait and Favorability Judgments

Adrienne B. Austin

Kristi A. Costabile

Lauren Smith

Iowa State University

**Manuscript accepted for publication by the *Journal of Media Psychology***

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

### **Abstract**

Two experiments examined how perceivers evaluated target individuals based on minimal information as presented in a typical social media post and whether inferences varied as a function of information source (self vs. other) and valence (positive vs. negative). Across experiments, results indicated that targets were: a) less likely to be rated with traits consistent with behavior and b) perceived less favorably when positive behavior information was self-generated than when the same information was other-generated. The inclusion of self-deprecating hashtags reduced the source effect of positive information by reducing perceived arrogance and increasing perceived sense of humor of target individuals. Together, these experiments provide greater understanding of the influence of information source, valence, and self-deprecation on trait and favorability judgments in a social media context.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

### Social Judgments, Social Media, and Self-Deprecation: Role of Information Source and Valence on Trait and Favorability Judgments

People are motivated to present themselves in ways that elicit positive impressions from others (Baumeister, 1982; Leary, 1995). Perceivers of social information, however, detect when others use self-presentational strategies (Leary, 1995; Ham & Vonk, 2011; Vonk, 1999), leading to a *source effect*, wherein individuals find *other*-generated positive information about a target to be more credible than *self*-generated positive information (Brandt, Vonk, & van Knippenberg, 2009; Rosenthal-Stott, Dicks, & Fielding, 2015). Moreover, perceivers report unfavorable impressions of individuals when it appears that the individuals were intentionally attempting to elicit favorable impressions (Lafreniere, Sedikides, Van Tongeren, & Davis, 2016). The present investigation delves into this phenomenon to examine both the antecedent conditions and downstream effects related to the source effect in a social media context. In particular, we examined how self-deprecation might attenuate the source effect.

#### **The Source Effect and Impression Formation**

Trait inferences and favorability judgments are greatly influenced by inferences of a target's motives, particularly in response to intentional behaviors (Reeder et al., 2004; Vonk, 1999). For example, inferring ulterior motives (e.g., ingratiating oneself to others) for positive behaviors (e.g., helping a friend) has been shown to reduce the likelihood that a perceiver would view the target as having the corresponding trait (e.g., helpful; Ham & Vonk, 2011). Perceivers tend to regard positive self-descriptions with suspicion and evaluate targets less favorably than when positive descriptions are made by third parties (Brandt, Vonk, & van Knippenberg, 2009; Vonk, 1999). Moreover, individuals who associate their personal characteristics with an achievement are more likely to be perceived as arrogant and in turn, less likeable (Hareli &

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Weiner, 2000). The source effect does not tend to occur for negative self-descriptions, as self-enhancement strategies are less likely to be suspected when people disclose unflattering information about themselves (Brandt et al., 2009; Jones & Davis, 1965).

**Impression formation formed via social media.** Research on impression formation processes typically presents information via an omniscient observer (e.g., “Jack stops his car and motions for the pedestrians to cross”, Costabile, 2016), or provides information in the context of a specific face-to-face social interaction (e.g., disclosure during a job interview, Brandt et al., 2009), contexts that differ from social media communication in important ways. Social media operates at a level that is public and yet personal, as what has been termed masspersonal communication (O’Sullivan & Carr, 2017) and carries its own norms and expectations for appropriate disclosures (Utz, 2015) that guide message sender behavior (French & Bazarova, 2017) as well as audience perceptions (Hancock & Dunham, 2001). Violating expected communication norms can result in negative interpersonal judgments. Boasting on social media has been found to result in especially negative judgments in contrast to other communication venues (Grant, Hodge, & Sinha, 2018). Moreover, impressions formed using computer mediated communication tend to be less detailed and more intense than those formed during face-to-face interactions (Hancock & Dunham, 2001), suggesting that information posted on social media may be particularly vulnerable to source effects. As such, a greater understanding of how perceivers form impressions of individuals described in social media posts is warranted. To better understand the processes underlying the source effect, the present investigation examined whether impressions of target individuals described in social media posts may differ as a function of information source (self versus other) and valence (positive or negative).

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

**Source effects and arrogance.** Social media presents perceivers with several cues to weigh when forming impressions of targets (e.g., photo, number of friends). The warranting principle (Walther & Parks, 2002) proposes that perceivers give less weight to social media cues that are subject to self-presentational manipulation. As such, the warranting principle would predict that perceivers would give more weight to other-generated information than to self-generated information, as other-generated information is immune to self-presentational concerns. However, it is possible that self-generated positive posts may not merely reduce weighting of the information presented, but might *also* evoke other, unfavorable inferences (i.e., arrogance) that may contribute to overall impressions. For example, Hareli and Weiner (2000) found that individuals who attribute personal achievements to their own abilities are often perceived as arrogant (see also Scott & Ravenscroft, 2017). The present investigation examines the role of source effects on both trait inference generation and favorability perceptions as doing so allows us to better understand the degree to which source effects are due to reduced behavior-consistent trait inference in accordance with the warranting principle (e.g., trait "helpful" is less likely to be inferred after target self-reports a helpful behavior) and due to the addition of a negatively-valenced trait inference (i.e., arrogance). If the source effect is driven by perceptions of target arrogance, we propose that the source effect might then be attenuated in situations in which arrogance is reduced. Thus, the present investigation examined the role of self-deprecation on impression formation.

**Impression formation and self-deprecation.** Self-deprecation can operate as a form of impression management. Individuals may engage in self-deprecation after a personal success in an effort to reduce the degree to which they worry their success may threaten the self-concept of others. In this case, self-deprecation stems from an affiliation-related motive, in which an

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

individual aims to increase their relational value by highlighting personal flaws and deflecting credit for success (Exline & Lobel, 1999). Although posters of positive content on social media are frequently viewed as braggarts, many are also perceived as having a higher status – as being happier, more competent, more attractive, and as having a better life overall (Chou & Edge, 2012; de Vries & Kühne, 2015). Individuals who are perceived to be high in status (e.g., physically attractive, high achieving, successful family) are viewed by others more positively when they engage in self-deprecation, that is, when they highlight personal flaws in intelligence, traits, morality, mental health, and physical appearance (Greengross & Miller, 2008) and downplay their own abilities in connection with their successes (Hareli & Weiner, 2000). Thus, the present investigation will examine whether self-deprecation may attenuate observed source effects found in social media contexts by reducing perceived arrogance.

### **Overview of Experiments**

The present investigation examined the influence of source on trait and favorability inferences made as a result of behavioral information posted on social media. Across two experiments, participants evaluated a target person based on positive and negative descriptions of trait-implying behavioral information presented in ostensible social media posts. Consistent with previous research (Brandt et al., 2009), we hypothesized that there would be a source effect for positive behaviors such that the target individual would be rated as more likely to possess the implied trait and be rated more favorably when the information was presented by a third party than when it was target-presented. Additionally, because boasting results in perceptions of arrogance (Hareli & Weiner, 2000), we expected that the relationship between source and favorability would be explained by perceived arrogance for positive behaviors. We did not expect a source effect when described behaviors implied a negative trait. In Experiment 2, we

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

predicted that the use of self-deprecation would attenuate the source effect by reducing perceptions of target arrogance and by enhancing perceptions of humorousness.

### **Experiment 1**

In this experiment, participants viewed a mock social media feed that consisted of simultaneous presentation of eight stimulus posts, each post consisted of a trait-implying behavior. We examined how the source (i.e., self or other) and valence (i.e., positive or negative) of the brief behavioral descriptions influenced trait inferences and favorability judgments of a target and whether perceptions of target arrogance would mediate the relationship between information source and favorability judgments for positive behaviors.

### **Participants**

Previous work investigating the source effect and impression formation online revealed a range of moderate to large effect sizes (e.g., Scott & Ravenscroft, 2017; Walther, Van Der Heide, Hamel, & Shulman, 2009). Conservative sample size estimates obtained via G\*Power ranged from 120 to 180 participants to detect small to medium effects, with power  $(1-\beta) = .90$  in a two-tailed, repeated measures design with within- and between-subjects factors (Faul et al., 2007). One hundred eighty-eight undergraduate students were recruited from introductory psychology and communication studies courses participated in the study. Eighty-two percent of participants identified as Caucasian, and 67% of participants identified as female, with a mean age of 18.84 years ( $SD = 1.69$ ). Two participants did not provide demographic information. Seventy-nine percent of participants indicated that they had a Twitter account currently or in the past, suggesting that participants were largely familiar with the context of the present experiment.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

### Design

A 2 (Behavior valence: Positive versus Negative) x 2 (Presentation source: Self- versus Third-party) within-participants design was used. Participants were randomly assigned to one of two sets of stimuli that varied presentation source such that if a behavior was target-generated in Set A, then the same behavior was other-generated in Set B, and vice versa.

### Materials and Procedure

Participants viewed a mock Twitter feed that consisted of simultaneous presentation of 8 posts, each provided by a different individual and then they answered questions about their judgments of the individuals described in each post. Half of the posts were self-descriptions, half contained third-party descriptions. Additionally, half of the posts described a positive behavior, and half described a negative behavior. Behavioral statements were selected based on findings from previous research that the statement implied a single, specific trait (e.g., “Left textbook at home” implied the trait “forgetful”; Costabile, 2016). Profile pictures for each post contained a photograph of two or more same-gender individuals to reduce the informative nature of the photograph, as participants would not know which individual depicted in the photograph posted the information. The same profile photograph was used for each behavior regardless of presentation source. After viewing the full Twitter feed, participants rated eight targets, one at a time, on a series of trait words and provided a rating of the overall favorability of each target.

**Trait judgments.** Participants indicated the degree to which a series of 22 trait words described the target of the post using a 7-point scale (1 = *Not at all* to 7 = *Extremely*). Analyses examined behavior-consistent traits (e.g., ratings of *clumsy* of the target who tripped walking across campus), valence-consistent traits, and perceived arrogance (single-item measure). To assess valence-consistent traits, we used the three valence-consistent traits that were *not* the



## SOCIAL JUDGMENTS AND SOCIAL MEDIA

behavior-consistent trait from the pilot study described below (e.g., if the behavior-consistent trait was *considerate*, the valence-consistent traits were *athletic*, *intelligent*, and *brave*).

**Favorability judgment.** Participants completed a one-item measure indicating their overall impression of the target of the post using a 7-point scale (1 = *Extremely negative* to 7 = *Extremely positive*).

**Trait valence pilot study.** Prior to Experiment 1, a pilot study was conducted in which participants ( $N = 24$ ) evaluated a series of trait words on a 10-point scale (1 = *Extremely negative* to 10 = *Extremely positive*). Positive trait words selected for the study included: *considerate*, *athletic*, *intelligent*, and *brave*, and were perceived as positive ( $M = 7.71$ ,  $SD = 1.10$ ),  $t(23) = 9.81$ ,  $p < .001$ , when compared to the midpoint of the scale (5.50). These traits had a Cronbach's  $\alpha$  of .706 for standardized items, with an average inter-item correlation of .375, which falls into Clark and Watson's (1995) optimal range of .2 to .4. A factor analysis using maximum likelihood model of the four items indicated loadings on a single factor, with an eigenvalue of 2.147 explaining 53.69% of the variance, and all factor loadings  $>.48$ ,  $\chi^2(2) = 2.845$ ,  $p = .241$ . Negative trait words included: *forgetful*, *clumsy*, *shy*, and *lazy* and were perceived as negative ( $M = 3.94$ ,  $SD = 1.22$ ),  $t(23) = 3.94$ ,  $p < .001$ , when compared to the midpoint of the scale (5.50). These traits had a Cronbach's  $\alpha$  of .683 for standardized items, with average inter-item correlation of .375, and factor analysis using maximum likelihood model of the four items indicated loadings on a single factor, with an eigenvalue of 2.083 explaining 52.09% of variance, and all factor loadings  $>.37$ ,  $\chi^2(2) = .612$ ,  $p = .736$ .<sup>1</sup>

---

<sup>1</sup> Results of factor analyses were consistent with other reliability indices, however, note that factor analyses were likely underpowered given the small sample size and average factor loadings below .80 (Russell, 2002).

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

**Results and Discussion**

**Behavior-consistent trait inferences.** First, to examine whether participants made behavior-consistent trait inferences to a greater degree than valence-consistent trait inferences, we subtracted ratings of valence-consistent traits from behavior-consistent traits for each target. A one-way *t*-test indicated this score was different from 0 for each target, all  $t_s(187) > 23.16$ ,  $p_s < .001$ , suggesting that trait judgments were specific and informed by the presented behavior information.

Next, a between-repeated measures mixed analysis of variance (ANOVA) of behavior valence (positive or negative) and presentation source (self or third party), with between-subjects stimuli set (Set A or Set B) on behavior-consistent trait inferences indicated small main effects of both behavior valence ( $F(1, 186) = 4.90$ ,  $p = .028$ ,  $\eta_p^2 = .03$ , 95% CI[0, .08],  $d = .29$ ) and presentation source ( $F(1, 186) = 6.89$ ,  $p = .009$ ,  $\eta_p^2 = .04$ , 95% CI[0, .10],  $d = .35$ ). These main effects were qualified by a valence-source interaction,  $F(1, 186) = 13.44$ ,  $p < .001$ ,  $\eta_p^2 = .07$ , 95% CI[.01, .14],  $d = .50$ , see Figure 1. Planned contrasts indicated that targets were more likely to be described with traits consistent with the positive behavior when information was presented by a third party than when self-presented,  $d = .74$ ; however, there was no effect of source for negative behaviors, see Table 1 for means, standard deviations, and inferential statistics.<sup>2</sup>

**Favorability judgment.** A three-way mixed measures ANOVA of behavior valence, presentation source, and stimuli set on favorability judgments revealed large main effects of behavior valence ( $F(1, 186) = 363.44$ ,  $p < .001$ ,  $\eta_p^2 = .66$ , 95% CI[.58, .72],  $d = 2.77$ ) and of source ( $F(1, 186) = 48.22$ ,  $p < .001$ ,  $\eta_p^2 = .21$ , 95% CI[.11, .30],  $d = 1.00$ ); however, these main

---

<sup>2</sup> Across studies, analyses indicated a statistically significant interaction of stimulus set with one or more independent variables, suggesting that trait inferences and favorability judgments vary as a function of the specific behavioral information. These analyses are explored more fully in supplemental materials that include analyses presented at the individual behavior level.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

effects were qualified by a large valence-source interaction effect,  $F(1, 186) = 106.49, p < .001, \eta_p^2 = .36, 95\% \text{ CI} [.26, .46], d = 1.50$ , see Figure 1. Planned contrasts indicated that targets who posted positive behaviors about themselves were viewed less favorably than when the information was posted by a third party,  $d = .99$ . However, there was no effect of source when negative behaviors were posted, see Table 1 for means, standard deviations, and inferential statistics.

**Perceived arrogance as mediator of judgments.** A mediation analysis was conducted in MEMORE for SPSS, which has the ability to estimate direct, indirect, and total effects for two-condition, within-participant designs (Montoya & Hayes, 2017). We examined whether perceived arrogance mediated the relationship between source (self and other) and favorability for positive behaviors. Results indicated that there was a total effect of source on favorability,  $b = 1.01, t(188) = 6.95, p < .001, 95\% \text{ CI} [0.73, 1.30]$ . Source was a significant predictor of arrogance,  $b = -1.41, t(188) = -9.06, p < .001, 95\% \text{ CI} [-1.72, -1.11]$ , and arrogance was a significant predictor of favorability,  $b = -0.75, t(186) = -17.50, p < .001, 95\% \text{ CI} [-0.83, -0.66]$ . After controlling for arrogance, source was no longer a significant predictor of favorability,  $b = -0.04, t(186) = -0.38, p = .70, 95\% \text{ CI} [-0.25, 0.17]$ . The variables combined to explain 63% of the variance in favorability,  $F(2, 186) = 155.55, p < .001$ . A percentile bootstrap confidence interval for the indirect effect was also constructed in MEMORE using 10,000 bootstrap samples. The indirect effect of source on favorability through arrogance was significant,  $b = 1.05, 95\% \text{ CI} [0.82, 1.30]$ , illustrating that arrogance mediated the relationship between source and favorability, see Figure 3x in supplemental materials.

Results of Experiment 1 indicated that participants made behavior-consistent trait inferences to a greater degree than valence-consistent trait inferences, suggesting that the trait

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

judgments were informed by the brief behavioral descriptions presented on the mock social media feed. Moreover, when positive information was presented by a third party, targets were more likely to be assigned the trait consistent with the posted behavior and were evaluated more favorably overall than when positive information was target-generated. Additionally, targets who posted positive information about themselves were perceived to be more arrogant and in turn, were viewed less favorably than when positive information was posted by a third party. Causal conclusions regarding the role of perceived arrogance in this study are limited due to the use of mediation analyses, which do not allow causal conclusions to be drawn (Grice et al., 2015). Thus, Experiment 2 examined the role of perceived arrogance using an experimental approach. Specifically, we examined whether a factor that we hypothesized would reduce perceptions of arrogance (i.e., self-deprecation) would similarly reduce the source effect observed in Experiment 1.

### **Experiment 2**

Experiment 2 sought to examine whether use of self-deprecation might attenuate the source effect observed in Experiment 1. Self-deprecation is an impression management strategy that occurs when individuals highlight personal flaws and downplay their own abilities in connection with personal achievements or positive events (Greengross & Miller, 2008; Hareli & Weiner, 2000). Previous research suggests that individuals who engage in self-deprecation are less likely to be perceived as arrogant (Hareli & Weiner, 2000). Accordingly, Experiment 2 experimentally manipulated the presence of self-deprecation in target-generated positive behavioral descriptions to examine whether the presence of self-deprecation would attenuate the source effect observed in Experiment 1.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Given that self-deprecation is often used as a form of conversational humor (Dyrel, 2009), exploratory analyses were performed to examine the role of humor as an additional mediator of the relationship between use of self-deprecation in positive posts and favorability judgments. The addition of humor to the disclosure of negative information has been shown to reduce perceived veracity of the negative information and in turn, increase favorable impressions of a target (Bitterly & Schweitzer, 2019), suggesting that humor might offset the negative inferences made regarding perceived arrogance. Experiment 2 also included a measure to assess the degree to which participants reported using various social media cues (e.g., text, source, names) when forming impressions of the targets to better understand perceptions of social media cue weighting processes.

Our laboratory group previously conducted a smaller in-person study examining the role of self-deprecating hashtags on the source effect. Experiment 2 was designed as a high-powered, pre-registered replication and extension of this previous investigation. This extension includes a greater variety of stimulus images and names, and measures assessing humor and cue use. Due to space constraints, full details of the previous study are presented in supplemental materials.

### **Participants**

Given our predictions for within-subjects interactions, we used PANGEA to estimate the appropriate sample size. The smallest interaction effect size obtained in the smaller previous study<sup>3</sup> was  $d = .50$ . Effect size for the current study was estimated to be  $d = .25$  to allow for a smaller effect size due to online data collection as required by Covid-19 lab protocols. With a sample size of 250, and error variance estimates of .49, power was determined to be .803.

According to Fritz and MacKinnon's (2007) guidelines for mediation sample sizes and power, a

---

<sup>3</sup>We previously conducted an in-person study using the same behavioral statements and hashtags as in Experiment 2 with an underpowered sample ( $N = 87$ ). Full details of this study are presented in supplemental materials.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

sample of 250 participants would have the power to detect mediation if the alpha and beta paths are mid-sized (.26). We planned to collect data from 300 participants to allow for potential attention check failures.

Three hundred and three participants were recruited from the online subject recruitment platform Prolific ([www.prolific.co](http://www.prolific.co)) and compensated \$2.00 for their participation. The number of participants surpassed our goal of 300 participants due to outstanding recruitment slots. Study recruitment was restricted to native English speakers between the ages of 18 and 30 and individuals currently living in the Midwest of the United States to best match the demographics of participants used during pilot testing of materials in Experiment 1. To confirm screening procedures, participants were asked to self-report first language, age, and current U.S. state of residence at the end of the study. These self-report items indicated that 9 individuals indicated a language other than English as their first language and 13 individuals indicated that they lived outside of the Midwestern U.S. Three individuals failed at least one of our two direct attention check items (items are described in full below), leaving a final sample of 278. Of this final sample, 80% of participants identified as Caucasian, and 57% identified as female, with a mean age of 24.00 years ( $SD = 3.51$ ). Seventy-eight percent of participants indicated that they had a Twitter account currently or in the past, suggesting that most participants were familiar with the context of the experiment.

### **Design, Materials, and Procedure**

Design, materials, procedure, and measures were similar to those used in Experiment 1 with some changes. Unlike Experiment 1, recruitment and data collection occurred online, and social media posts were presented one at a time.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Participants viewed six posts, three of these posts described a behavior implying a positive trait and three described a behavior implying a negative trait. Each trait-implying behavior was presented with a neutral hashtag, a self-deprecating hashtag, or by a third party with no hashtag. Hashtags are text set off with a pound sign (#) that are used to tag content of a social media post with a keyword or phrase, full materials are presented in supplemental documents. Participants were randomly assigned to one of three presentation orders to ensure that participants viewed one post of each behavior-source combination (e.g., positive behavior with self-deprecating hashtag, positive behavior by third party, negative behavior with neutral hashtag). Because inferences are drawn from features in social media photographs (e.g., environment, Segalin et al., 2017), each behavior-source combination was paired with one of six possible photographs/names.<sup>4</sup> Photographs were collected from Pexels (www.pexels.com) and, as in Experiment 1, included at least two same-gender individuals to reduce the informative nature of the photograph.

Participants indicated the degree to which a series of 13 trait words described the target of the post using a 7-point scale (1 = *Not at all* to 7 = *Extremely*). Analyses examined behavior-consistent traits, perceived arrogance (single-item measure), and perceived sense of humor (single-item measure). Participants then completed a one-item measure indicating their overall impression of the target of the post using a 7-point scale (1 = *Extremely negative* to 7 = *Extremely positive*). To detect careless responding that is commonly observed in online research (Huang, Liu, & Bowling, 2015), we also included two direct attention check items among the

---

<sup>4</sup> It was possible that behaviors used may be interpreted differently as a function of target gender. To examine this possibility, half of the stimuli sets used the implied gender consistent with those used in Experiment 1 and half used the opposite. This factor did not reveal any statistically significant main effects or interactions (all  $ps > .14$ ), thus results are reported collapsing across this factor.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

trait ratings for different targets (e.g., "It is important that you read the items in this survey. Please click "1 Not at all").

To assess self-reported attention to social media cues, participants were asked to report the degree to which they used different cues as they formed the impression of the target on a 7-point scale (1 = *Not at all* to 7 = *Very much*) at the end of the study. Participants were asked to rate the degree to which they used: a) Text of the post, b) Photograph, c) Target name, d) Who posted the information (self or other), and e) Hashtags presented.

### Results and Discussion

**Behavior-consistent trait inferences.** Results of a 3 (between subjects - stimulus set) X 2 (repeated measures - valence) x 3 (repeated-measures source: self-deprecating hashtags, self-neutral hashtags, posts by another person) mixed ANOVA on behavior-consistent trait inferences indicated a valence-by-source interaction,  $F(2, 274) = 71.01, p < .001, \eta_p^2 = .34, 95\% \text{ CI} [.25, .42], d = 1.42$ , see Figure 2. Planned contrasts using Tukey's HSD test indicated that consistent with Experiment 1 targets were more likely to be described with traits consistent with the positive behavior when that information was presented by a third party when compared with self-generated posts with a self-deprecating hashtag,  $p = .040, d = .15$ , and a self-neutral hashtag,  $p = .057, d = .14$ . There were no differences in trait ratings for positive self-deprecating posts and positive self-neutral posts,  $p = .900$ . This pattern was reversed for trait ratings for negative behaviors, such that targets were less likely to be described with traits consistent with the negative behavior when that information was presented by a third party when compared with self-generated posts presented with a neutral hashtag,  $p < .001, d = .92$ , and presented with a self-deprecating hashtag,  $p < .001, d = .84$ . There were no differences in trait ratings for negative behaviors when comparing self-generated posts,  $p = .148$ , see Table 2.



## SOCIAL JUDGMENTS AND SOCIAL MEDIA

**Favorability judgment.** Results of a 3 (between subjects - stimulus set) X 2 (repeated measures - valence) x 3 (repeated-measures - source: self-deprecating hashtags, self-neutral hashtags, posts by another person) mixed ANOVA indicated a significant valence-by-source interaction when examining self-deprecating hashtags, self-neutral hashtags, and posts by another person by stimulus set,  $F(2, 274) = 65.02$   $p < .001$ ,  $\eta_p^2 = .32$ , 95% CI[.23, .40],  $d = 1.36$ , see Figure 2. Planned contrasts using Tukey's HSD test indicated significant differences in favorability ratings when comparing sources for positive posts,  $d = 1.53$ , such that targets were rated more favorably when positive behavior information was presented by a third party when compared with information that was self-generated and presented with a self-deprecating hashtag,  $p < .001$ ,  $d = .63$ , and presented with a self-neutral hashtag,  $p < .001$ ,  $d = .99$ . When comparing self-generated positive posts, targets were rated more favorably when their post included a self-deprecating hashtag than a self-neutral hashtag,  $p < .001$ ,  $d = .35$ . This pattern was reversed for favorability ratings for negative behaviors,  $d = .48$ , such that targets were rated more favorably when negative behavior information was self-generated and presented with a self-deprecating hashtag when compared with self-neutral hashtags,  $p = .030$ ,  $d = .15$ , and posts by a third party,  $p < .001$ ,  $d = .31$ . Target-generated negative posts with a neutral hashtag were rated more favorably than negative posts by a third party,  $p = .024$ ,  $d = .16$ , see Table 2 for descriptive and inferential statistics.

**Perceived arrogance and humor as mediators of judgments.** MEMORE software was used to examine whether perceived arrogance and perceived humor each mediated the relationship between hashtag type (neutral versus self-deprecating) and favorability judgments for positively-valenced posts.<sup>5</sup> Hashtag type was a predictor of arrogance,  $b = -0.50$ ,  $t(277) = -$

---

<sup>5</sup> Results without humor included are presented in supplemental materials.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

3.91,  $p < .001$ , 95% CI[-0.75, -0.25] and humor,  $b = 0.77$ ,  $t(277) = 8.02$ ,  $p < .001$ , 95% CI[0.58, 0.99]. Favorability was predicted by arrogance,  $b = -0.48$ ,  $t(273) = -12.52$ ,  $p < .001$ , 95% CI[-0.55, -0.40] and humor,  $b = 0.24$ ,  $t(273) = 4.67$ ,  $p < .001$ , 95% CI[0.14, 0.35]. After controlling for arrogance and humor, hashtag type no longer influenced favorability,  $b = 0.02$ ,  $t(273) = 0.29$ ,  $p = .774$ , 95% CI[-0.14, 0.19]. Results indicated that variables combined to explain 47% of the variance in favorability,  $F(4, 273) = 60.58$ ,  $p < .001$ . A percentile bootstrap confidence interval for the indirect effect was also constructed in MEMORE using 10,000 bootstrap samples. There were indirect effects of hashtag type on favorability through arrogance,  $b = 0.23$ , 95% CI[0.11, 0.38] and through humor,  $b = 0.19$ , 95% CI[0.11, 0.28], and arrogance and humor were mediators of the relationship between hashtag type and favorability,<sup>6</sup> see Figure 3.

**Cue use.** Participants were asked to self-report the cues they used when forming impressions of the target. A repeated-measure ANOVA indicated a main effect of the degree to which participants reported each of five cue types (text of post, photograph, name, source, or hashtag),  $F(4, 274) = 483.36$ ,  $p < .001$ ,  $\eta_p^2 = .64$ , 95% CI[.57, .69],  $d = 2.63$ . Participants indicated that they were more likely to use the text of post ( $M = 6.46$ ,  $SD = 1.01$ ) when forming impressions of the targets than all the other cues,  $F_s > 300.00$ ,  $p_s < .001$ ,  $d_s = 2.13-5.84$ . Participants were equally likely to indicate that they used hashtags ( $M = 4.18$ ,  $SD = 1.83$ ) and source ( $M = 3.99$ ,  $SD = 2.05$ ) when forming impressions of the targets,  $F(4, 274) = 1.67$ ,  $p = .197$ ,  $\eta_p^2 = .02$ , 95% CI[0, .06],  $d = .20$ , and they were more likely to indicate use of hashtags or source than they were to indicate use of profile photograph ( $M = 1.96$ ,  $SD = 1.41$ ) or target name ( $M = 1.83$ ,  $SD = 1.25$ ),  $F_s > 300.00$ ,  $p_s < .01$ ,  $d_s = 1.15-1.50$ .

---

<sup>6</sup> Results of the previous study indicated the same general pattern of results of those reported here; however, unlike the present results the previous study found that self-deprecating hashtags attenuated the source effect for behavior-consistent trait inferences. Full details are presented in supplemental materials.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Experiment 2 results indicated that, consistent with Experiment 1, targets were evaluated more favorably overall when positive information was posted by a third party than when the same information was self-generated. Moreover, results indicated that targets were evaluated more favorably for self-generated positive posts when they were presented with a self-deprecating hashtag than when they were presented with a neutral hashtag. Indeed, targets who posted positive information about themselves with a self-deprecating hashtag were perceived to be less arrogant and more humorous, and in turn, were viewed more favorably than targets who posted positive information with a neutral hashtag.

The use of self-deprecation did not appear to influence trait inferences. Self-deprecation cues involve downplaying the role of personal characteristics in positive events; thus, it is likely that these cues reduced the strength of the association between the described behavior and trait-consistent inferences. Additionally, participants indicated that they used multiple sources of information when forming impressions of the target individuals, giving particular weight to the content of the post and, to a lesser degree, the source of the post and the use of hashtags. Participants indicated little weight was given to profile photographs or target names.

### **General Discussion**

The present investigation contributes to the literature on impression formation by delving into both the antecedent conditions and downstream effects related to the source effect in a social media context. We examined how information source (presented by oneself or by a third party) and information valence (positive or negative) of brief behavioral descriptions influence trait inferences and favorability judgments made for a target person. In Experiment 1, we found that participants made behavior-consistent trait inferences to a greater degree than valence-consistent trait inferences, suggesting that the brief behavioral information provided in a social media feed

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

do inform trait judgments (see also Levordashka & Utz, 2017). Moreover, targets were more likely to be described with traits consistent with posted behavior and evaluated more favorably when positive information was presented by a third-party than when the same information was target-generated, an effect mediated by perceived target arrogance. Results of Experiment 2 demonstrated that an experimental manipulation of self-deprecation attenuated the source effect for favorability judgments, but not for behavior-consistent trait ratings. Targets who used self-deprecating hashtags were perceived to be less arrogant and more humorous than those who used neutral hashtags and in turn, were viewed more favorably when describing their own positive behavior. This work is informative for identifying how aspects of self-presentation in a digital environment influence social perceptions.

Previous research on correspondent inferences has shown an effect of source on impression formation, in that individuals were more likely to believe positive trait information when provided by a third party (i.e., a previous employer) than when provided by a job applicant him/herself (Brandt et al., 2009) as such information is unlikely to be biased by self-presentational concerns. Our work extends this body of research to a social media context by revealing that such effects are evident with mere behavioral statements, with source effects occurring at the trait inference level. Importantly, when self-deprecating information was added to a self-presented description of a positive behavior, targets were less likely to be perceived as arrogant and more likely to be perceived as humorous, leading to more favorable overall evaluations. Accordingly, the present results indicate that self-deprecation can be used as an effective self-presentation strategy for individuals engaging in a positive masspersonal disclosure.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

By illustrating the role of perceived arrogance and humor, the present investigation highlights that the source effect is not entirely due to underweighting trait inferences as would be suggested by the warranting principle (Walther & Parks, 2002), but rather the effect is also bound to the evocation of other valenced traits. When target individuals engaged in boasting, or describing their own positive behavior, the information was not merely given less weight in the perceiver's overall impression. Rather, the target individuals were perceived to be arrogant, and the perception of arrogance was associated with unfavorable evaluation of the targets. The present experiments identified that this inference process can be disrupted through the use of self-deprecation. By downplaying their positive attributes, targets were perceived to be less arrogant and more humorous, illustrating how self-deprecation can serve to buffer targets from negative evaluations related to their boasting behavior. Additionally, perceivers reported that when forming judgments, they weighted in equal measure source and self-generated hashtags, providing more evidence that perceivers do consider self-presented information a useful cue in the impression formation process.

**Size of effects on trait inferences versus favorability judgments.** Our findings highlight that trait inferences and favorability judgments are distinct forms of social inferences. Across experiments, results consistently indicated that source effects were larger for favorability judgments than they were for trait judgments, an illustration of how favorability judgments are based on a constellation of inferences (trait-consistent inferences, as well as other traits evoked by inferred motives, Vonk, 1999). Our results suggest that the mere act of posting self-favorable information on social media evokes negative trait inferences (e.g., arrogance), inferences that can be reconsidered when the target engages in self-deprecation. Our results are consistent with previous research that finds that individuals attend to and use a combination of behavior and trait

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

information when forming global impressions of others (e.g., Brambilla, Sacchi, Rusconi, Cherubini, & Yzerbyt, 2011).

The different effect sizes across dependent measures may have also been due in part to the prevalence and remarkable efficiency of implicit trait inference generation (Todorov & Uleman, 2003). It is possible that trait inferences may have been bound by ceiling effects preventing larger differences between source conditions from being obtained in the present experiments.

### **Limitations and Future Directions**

There are a number of limitations to the current investigation. First, it is important to note that we used a single behavior for each trait-related inference, which limits the generalizability of the present investigation. Along these lines, it is important to note that although the style of presented posts were designed to mimic those used by the actual social media source, the presented content of described typical behaviors may not fully reflect the norms of disclosure on social media sites.

Additionally, the observed effects may not be applicable to all social interactions (e.g., when straightforward self-presentation strategies are expected, such as job interviews) or even to all social media platforms. Different norms exist for different social media platforms. More work is needed to examine the processes underlying source effects and self-deprecation in other settings.

The effects of self-deprecation on impression formation observed in the present investigation stand in contrast to research on humblebragging, a self-presentation strategy in which individuals mask their boasting in a complaint (e.g., “It’s exhausting being the only one people come to for advice.”) or humility (e.g., “I won first place in the marathon, but I’m not

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

even a fast runner.”). Perceivers typically rate humblebraggers as insincere, resulting in even more negative perceptions than straightforward braggarts (Sezer, Gino, & Norton, 2018). The use of self-deprecating humor, in contrast to humblebragging, may illustrate a degree of self-awareness and sincerity, possibilities that should be explored in future research.

Additionally, previous work has indicated that individuals who engaged in self-enhancement were viewed as having poor social skills (Colvin, Block, & Funder, 1995), and that utilizing humor can increase social attractiveness (Wanzer, Booth-Butterfield, & Booth-Butterfield, 2009). This work suggests that in addition to arrogance and humor, perceived social competence could be an additional potential mechanism to explore in future research. Indeed, given the extensive limitations of mediation analyses, a thorough investigation of potential alternative paths would be necessary to draw firm causal conclusions (Bullock, Green, & Ha, 2010).

### **Concluding Comments**

The results of these experiments highlight the importance of source on immediate trait inferences and favorability judgments in a social media context, and identify how self-deprecation contributes to online impression management.

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

## Author Disclosure Statement

Materials, data, and supplemental analyses are presented on the Open Science Framework at: [https://osf.io/ew2dp/?view\\_only=f0d43f535c3f4ffe8b79e0faf67cbc10](https://osf.io/ew2dp/?view_only=f0d43f535c3f4ffe8b79e0faf67cbc10) .



## SOCIAL JUDGMENTS AND SOCIAL MEDIA

## Preregistration Statement

The hypotheses and analysis plan for Experiment 2 of this project were preregistered at [https://osf.io/8s7kx/?view\\_only=b5980129b905494697d50ba4b623e749](https://osf.io/8s7kx/?view_only=b5980129b905494697d50ba4b623e749) .

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

## References

- Baumeister, R. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*, 3-26. <http://dx.doi.org/10.1037/0033-2909.91.1.3>
- Bitterly, T. B., & Schweitzer, M. E. (2019). The impression management benefits of humorous self-disclosures: How humor influences perceptions of veracity. *Organizational Behavior and Human Decision Processes*, *151*, 73-89. <https://doi.org/10.1016/j.obhdp.2019.01.005>
- Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology*, *51*(1), 149-166. <https://doi.org/10.1111/j.2044-8309.2010.02011.x>
- Brandt, A. C., Vonk, R., & Knippenberg, A. V. (2009). The source effect: Person descriptions by self versus others have differential effects on impression formation. *Personality and Social Psychology Bulletin*, *35*, 965-977. <https://doi.org/10.1177/0146167209335056>
- Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*(4), 550-558. <https://doi.org/10.1037/a0018933>
- Chou, H-T. G., & Edge, N. (2012). "They are happier and having better lives than I am": The impact of using Facebook on perceptions of others' lives. *Cyberpsychology, Behavior, and Social Networking*, *15*(2), 117-121. <https://doi.org/10.1089/cyber.2011.0324>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309-319. <https://doi.org/10.1037/14805-012>

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

- Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68*(6), 1152–1162. <https://doi.org/10.1037/0022-3514.68.6.1152>
- Costabile, K. A. (2016). Narrative construction, social perceptions, and the situation model. *Personality and Social Psychology Bulletin, 42*(5), 589-602.  
<https://doi.org/10.1177%2F0146167216636627>
- Deters, F. G., & Mehl, M. R. (2013). Does posting Facebook status updates increase or decrease loneliness? An online social networking experiment. *Social Psychological and Personality Science, 4*(5), 579-586.
- de Vries, D. A., & Kühne, R. (2015). Facebook and self-perception: Individual susceptibility to negative social comparison. *Personality and Individual Differences, 86*, 217–221.  
<https://doi.org/10.1016/j.paid.2015.05.029>
- Dynel, M. (2009). Beyond a joke: Types of conversational humour. *Language and Linguistics Compass, 3*(5), 1284-1299. <https://doi.org/10.1111/j.1749-818X.2009.00152.x>
- Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook “friends:” Social capital and college students’ use of online social network sites. *Journal of Computer-Mediated Communication, 12*(4), 1143-1168. doi:10.1111/j.1083-6101.2007.00367.x
- Exline, J. J., & Lobel, M. (1999). The perils of outperformance: Sensitivity about being the target of a threatening upward comparison. *Psychological Bulletin, 125*(3), 307-337. doi:  
[10.1037/0033-2909.125.3.307](https://doi.org/10.1037/0033-2909.125.3.307)
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191. <https://doi.org/10.3758/BF03193146>

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

- French, M., & Bazarova, N. N. (2017). Is anybody out there?: Understanding masspersonal communication through expectations for response across social media platforms. *Journal of Computer-Mediated Communication*, 22(6), 303-319.  
<https://doi.org/10.1111/jcc4.12197>
- Gonzales, A. L., & Hancock, J. T. (2011). Mirror, mirror on my Facebook wall: Effects of exposure to Facebook on self-esteem. *Cyberpsychology, Behavior, and Social Networking*, 14(1-2), 79-83. <https://doi.org/10.1089/cyber.2009.0411>
- Greengross, G., & Miller, G. F. (2008). Dissing oneself versus dissing rivals: Effects of status, personality, and sex on the short-term and long-term attractiveness of self-deprecating and other-deprecating humor. *Evolutionary Psychology*, 6(3), 393-408.  
<https://doi.org/10.1177%2F147470490800600303>
- Grice, J. W., Cohn, A., Ramsey, R. R., & Chaney, J. M. (2015). On muddled reasoning and mediation modeling. *Basic and Applied Social Psychology*, 37(4), 214-225.  
<https://doi.org/10.1080/01973533.2015.1049350>
- Ham, J., & Vonk, R. (2011). Impressions of impression management: Evidence of spontaneous suspicion of ulterior motivation. *Journal of Experimental Social Psychology*, 47(2), 466-471. <https://doi.org/10.1016/j.jesp.2010.12.008>
- Hancock, J. T., & Dunham, P. J. (2001). Impression formation in computer-mediated communication revisited: An analysis of the breadth and intensity of impressions. *Communication Research*, 28(3), 325-347. <https://doi.org/10.1177/009365001028003004>
- Hareli, S., & Weiner, B. (2000). Accounts for success as determinants of perceived arrogance and modesty. *Motivation and Emotion*, 24(3), 215-236.  
<https://doi.org/10.1023/A:100566621>

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

- Huang, J. L., Liu, M., & Bowling, N. A. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology, 100*(3), 828–845.  
<https://doi.org/10.1037/a0038510>
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219-266). New York: Academic Press.
- Lafrenière, M. A. K., Sedikides, C., Van Tongeren, D. R., & Davis, J. (2016). On the perceived intentionality of self-enhancement. *The Journal of Social Psychology, 156*(1), 28-42.  
<https://doi.org/10.1080/00224545.2015.1041447>
- Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behavior*. Madison, WI: Brown & Benchmark.
- Levordashka, A., & Utz, S. (2017). Spontaneous trait inferences on social media. *Social Psychological and Personality Science, 8*(1), 93-101.  
<https://doi.org/10.1177%2F1948550616663803>
- Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods, 22*(1), 6-27.  
<https://psycnet.apa.org/doi/10.1037/met0000086>
- Nadkarni, A., & Hofmann, S. G. (2012). Why do people use Facebook? *Personality and Individual Differences, 52*(3), 243-249. <https://doi.org/10.1016/j.paid.2011.11.007>
- O’Sullivan, P. B., & Carr, C. T. (2018). Masspersonal communication: A model bridging the mass-interpersonal divide. *New Media & Society, 20*(3), 1161-1180.  
<https://doi.org/10.1177%2F1461444816686104>

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Reeder, G. D., Vonk, R., Ronk, M. J., Ham, J., & Lawrence, M. (2004). Dispositional attribution:

Multiple inferences about motive-related traits. *Journal of Personality and Social Psychology*, *86*, 530–544. <http://psycnet.apa.org/doi/10.1037/0022-3514.86.4.530>

Rosenthal-Stott, H. E. S., Dicks, R. E., & Fielding, L. S. (2015). The valence of self-generated (status updates) and other-generated (wall-posts) information determines impression formation on Facebook. *PLoS ONE*, *10*(6), e0125064.

<https://doi.org/10.1371/journal.pone.0125064>

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin*, *28*(12), 1629-1646. <https://doi.org/10.1177/014616702237645>

Scott, G. G., & Ravenscroft, K. (2017). Bragging on Facebook: The interaction of content source

and focus in online impression formation. *Cyberpsychology, Behavior, and Social Networking*, *20*(1), 58-63. <https://doi.org/10.1089/cyber.2016.0311>

<https://doi.org/10.1089/cyber.2016.0311>

Segalin, C., Celli, F., Polonio, L., Kosinski, M., Stillwell, D., Sebe, N., ... & Lepri, B. (2017).

What your Facebook profile picture reveals about your personality. In *Proceedings of the 25th ACM International Conference on Multimedia*, 460-468.

<https://doi.org/10.1145/3123266.3123331>

Sezer, O., Gino, F., & Norton, M. I. (2018). Humblebragging: A distinct—and ineffective—self-presentation strategy. *Journal of Personality and Social Psychology*, *114*(1), 52-74.

<http://dx.doi.org/10.1037/pspi0000108>

Shane-Simpson, C., Manago, A., Gaggi, N., Gillespie-Lynch, K. (2018). Why do college

students prefer Facebook, Twitter, or Instagram? Site affordances, tensions between

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

- privacy and self-expression, and implications for social capital. *Computers in Human Behavior*, 86, 276-288. <https://doi.org/10.1016/j.chb.2018.04.041>
- Todorov, A. & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*, 39, 549-562.  
doi:10.1016/S0022-1031(03)00059-3
- Utz, S. (2015). The function of self-disclosure on social network sites: Not only intimate, but also positive and entertaining self-disclosures increase the feeling of connection. *Computers in Human Behavior*, 45, 1–10. <https://doi.org/10.1016/j.chb.2014.11.076>
- Vonk, R. (1999). Impression formation and impression management: Motives, traits, and likeability inferred from self-promoting and self-deprecating behavior. *Social Cognition*, 17, 390-412. <https://doi.org/10.1521/soco.1999.17.4.390>
- Walther, J. B., & Parks, M. R. (2002). Cues filtered out, cues filtered in. *Handbook of Interpersonal Communication*, 3, 529-563.
- Walther, J. B., Van Der Heide, B., Hamel, L. M., & Shulman, H. C. (2009). Self-generated versus other-generated statements and impressions in computer-mediated communication. *Communication Research*, 36(2), 229-253.  
<https://doi.org/10.1177%2F0093650208330251>
- Wanzer, M. B., Booth-Butterfield, M., & Booth-Butterfield, S. (1996). Are funny people popular? An examination of humor orientation, loneliness, and social attraction. *Communication Quarterly*, 44(1), 42-52. <https://doi.org/10.1080/01463379609369999>

## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Table 1

*Experiment 1 Contrast of Self and Other for Behavior-Consistent Traits, Favorability, and Arrogance*

	<u>Self</u>		<u>Other</u>		<i>F</i> (1, 187)	<i>p</i>	$\eta_p^2$	$\eta_p^2$ 95% CI	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				<i>LL</i>	<i>UL</i>
Positive B-C Traits	6.11	.92	6.48	.71	26.54	<.001	.12	.05	.21
Negative B-C Traits	6.23	.92	6.13	1.14	1.23	.27	.01	0	.05
Positive Favorability	4.52	1.34	5.52	1.10	46.86	<.001	.20	.11	.30
Negative Favorability	3.81	.85	3.64	.83	5.11	.025	.03	.00	.09
Positive Arrogance	4.05	1.89	2.65	1.57	80.45	<.001	.30	.20	.40
Negative Arrogance	2.85	1.34	2.84	1.18	.01	.92	.00	0	.01

*Note.* Positive and Negative = positive and negative behavior posts, respectively; B-C = behavior-consistent; CI = confidence interval; *LL* = lower limit; *UL* = upper limit. Full analyses examining arrogance presented in supplemental materials.



## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Table 2

*Experiment 2 Contrast of Self-Deprecating Hashtag, Self-Neutral Hashtag, and Post by Other for Behavior-Consistent Traits, Favorability, Arrogance, and Humor*

	Self-Deprecating		Self-Neutral		Post by Other		<i>F</i> (2, 274)	$\eta_p^2$	$\eta_p^2$ 95% CI	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			<i>LL</i>	<i>UL</i>
Positive B-C Traits	5.55	1.28	5.56	1.42	5.77	1.60	2.43	.02	0	.05
Negative B-C Traits	5.81	1.12	5.92	1.07	4.47	1.95	72.06**	.34	.26	.42
Positive Favorability	4.62	1.32	4.15	1.40	5.39	1.10	82.53**	.38	.29	.45
Negative Favorability	3.97	1.12	3.80	1.17	3.60	1.27	8.83**	.06	.01	.12
Positive Arrogance	3.86	1.80	4.38	1.77	2.25	1.25	164.80**	.55	.47	.61
Negative Arrogance	3.02	1.47	2.97	1.50	3.29	1.70	3.88*	.03	.00	.07
Positive Humor	3.47	1.62	2.72	1.25	3.07	1.35	34.76**	.20	.12	.28
Negative Humor	4.23	1.48	3.76	1.65	3.27	1.53	41.53**	.23	.15	.31

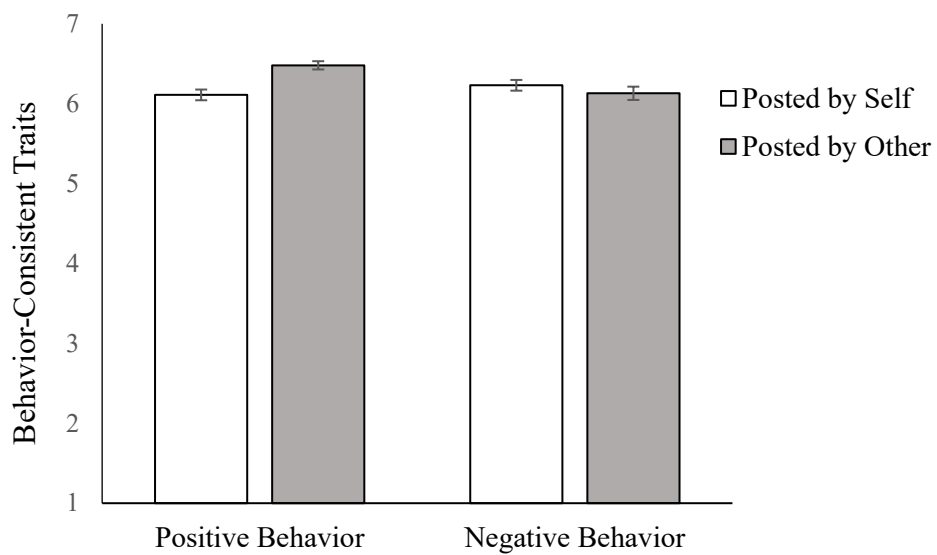
*Note.* Positive and Negative = positive and negative behavior posts, respectively; B-C = behavior-consistent; CI = confidence interval; *LL* = lower limit; *UL* = upper limit. Full analyses examining arrogance and humor presented in supplemental materials.

\*  $p < .05$ ; \*\*  $p < .001$

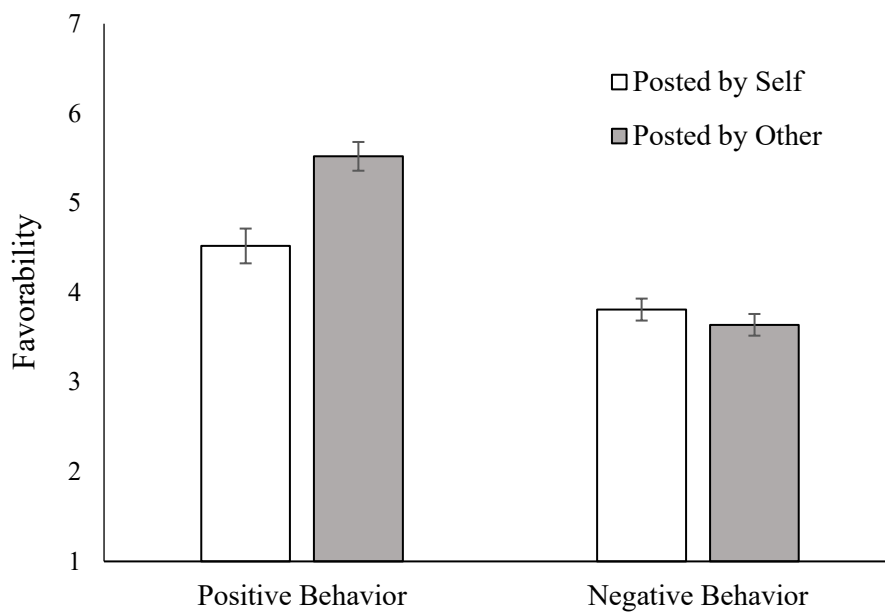
## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Figure 1. Interactions of behavior valence and presentation source on dependent measures in Experiment 1. Error bars represent standard error of the mean.

A.



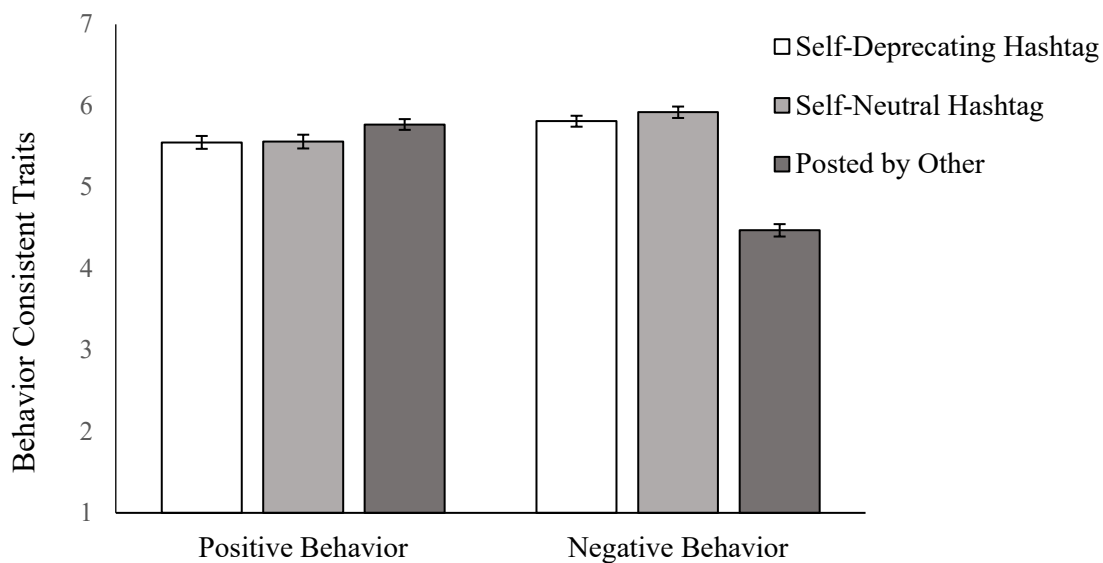
B.



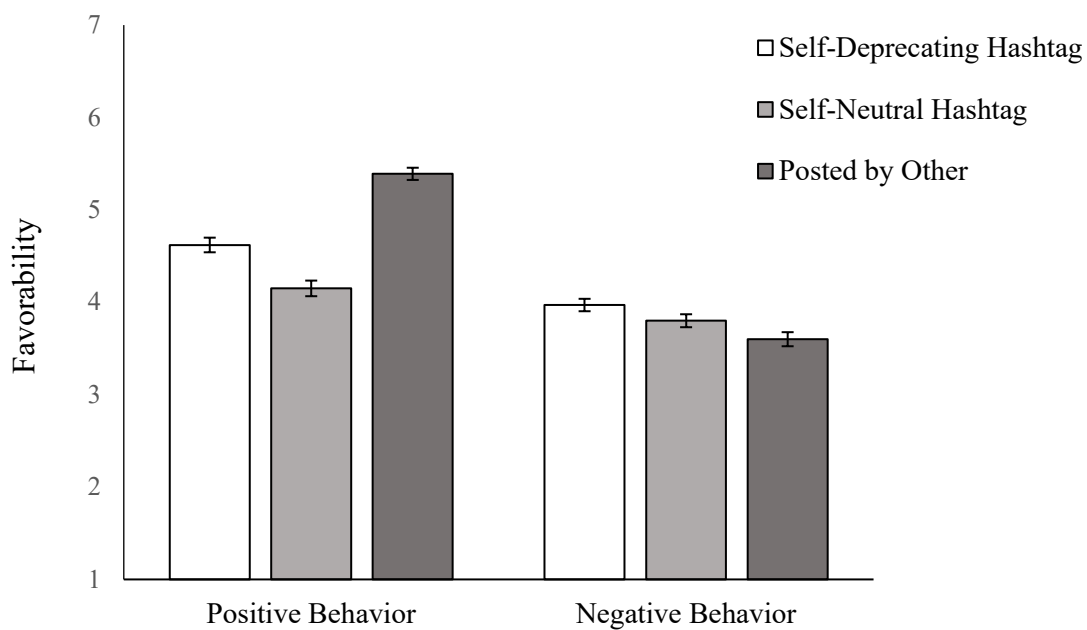
## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Figure 2. Interaction of behavior valence and presentation source on dependent measures in Experiment 2. Error bars represent standard error of the mean.

A.



B.



## SOCIAL JUDGMENTS AND SOCIAL MEDIA

Figure 3. Mediation model with hashtag type (neutral coded as 0, self-deprecating coded as 1) for positive posts as the independent variable, arrogance and humor as mediators, and favorability as the dependent variable in Experiment 2. The model presents the unstandardized regression coefficients. The inclusion of the mediator in the equation is reflected in the parentheses.

