The positive effect of not following others on social media

Francesca Valsesia[*]

Davide Proserpio[†]

Joseph C. Nunes[‡§]

[*] Francesca Valsesia is Assistant Professor of Marketing at the University of Washington's Foster School of Business, 4277 NE Stevens Way, Paccar Hall 439, Seattle, WA 98185.

[†] Davide Proserpio is Assistant Professor of Marketing at the University of Southern California's Marshall School of Business, 701 Exposition Blvd., HOH 332, Los Angeles, CA 90089-1424.

[‡] Joseph C. Nunes is Professor of Marketing at the University of Southern California's Marshall School of Business, 701 Exposition Blvd., HOH 336, Los Angeles, CA 90089-1424.

[§] The authors contributed equally and are listed in reverse alphabetical order. All correspondence regarding this manuscript should be addressed to valsesia@uw.edu.

ABSTRACT

Marketers seed information through individuals believed to be influential on social media. This often involves enlisting micro influencers, users who have accumulated thousands as opposed to millions of followers—other users who have subscribed to see that individual's posts. Given an abundance of micro influencers to choose from, cues that help distinguish more versus less effective influential users on social media are of increasing interest to marketers. We identify one such cue, the number of users the prospective influencer is following. Using a combination of real-world data analysis and controlled lab experiments, we show that *following* fewer others, conditional on possessing a substantial number of *followers*, has a positive effect on a social media user's perceived influence. Further, we find greater perceived influence impacts *engagement* with the content shared in terms of other users exhibiting more favorable attitudes toward it (i.e., likes) and a greater propensity to spread it (i.e., retweets). We identify a theoretically important mechanism underlying the effect: following fewer others conveys greater autonomy, a signal of influence in the eyes of others.

Keywords: Micro Influencer, Following, Social Media, Autonomy, Opinion Leader

The commercial use of social media *influencers* is a rapidly growing phenomenon, with marketers increasingly seeding information about their brands through select individuals on social media (Libai, Muller, and Peres 2013). Motivating the market is the longstanding belief that a small subset of consumers are extraordinarily influential, those considered opinion leaders (Rogers 1962; Brown and Hayes 2008). A common social media marketing strategy thus involves identifying an initial set of influencers and incentivizing them to share specific content, thereby initiating a viral marketing campaign in which their posts generate engagement—likes, comments, and reposts—widely believed to be indicators of the audience's future consumption behavior. The development of practical approaches to identifying who is truly influential online, however, is still in its infancy (Probst, Grosswiele, and Pfleger 2013).

One widely used social media seeding strategy is selecting individuals who can reach a large number of users. This often involves enlisting celebrities with millions of *followers*. For example, singer Selena Gomez, with 167 million *followers* on Instagram and 59.9 million *followers* on Twitter (as of February 2020), has taken part in social media campaigns for Coach, Coca-Cola, Verizon, and Pantene (Friedman 2017).[5] With her massive *reach*, Gomez is a highly sought after—albeit extremely expensive—social media influencer, reportedly reaping hundreds of thousands of dollars for a single sponsored post (Heine 2016). Given this kind of reach comes at such a high cost, marketers have begun looking for more efficient ways to spread their message on social media networks.

An alternative strategy has been for marketers to direct budgets toward "micro" influencers, social media users with anywhere between 1,000 and 100,000 *followers* who charge hundreds rather

[5] Social media platforms such as Instagram, Twitter, Snapchat, and Musical.ly allow *asymmetric following*, which lets users follow an individual or account without that individual or account needing to follow them back. Other platforms, including Facebook and LinkedIn, require *symmetric following*; connected users must follow each other.

than hundreds of thousands of dollars per post (Barker 2017). Micro influencers share content daily about everything from bass fishing to bass guitars and are considered an effective way to reach a specific target market. Consider, for example, competitive bass fisherman Randy Howell (@theRandyHowell) who as of February 2020 had 33,606 *followers* on Twitter. With an audience presumably comprised of fishing aficionados, he promotes brands such as Lowrance 3D fishfinders, Power-Pole boats, and Pelican marine coolers. Another competitive bass fisherman, Mark Zona (@MarkZonaFishing), had 36,276 *followers* at the time and promotes Strike King fishing lures, Daiwal reels, and Bass Mafia bait boxes. Zona has approximately the same "reach" as Howell. What distinguishes the two?

This question is of growing concern to marketers wrangling with the question of who among an ever-expanding assortment of candidate social media influencers to sponsor (Momtaz, Aghaie, and Alizadeh 2011; Nejad, Sherrell, and Babakus 2014). Pragmatically speaking, a survey of marketing practitioners finds that selecting the "right" influencer is the biggest challenge in working with influencers online (Simpson 2016). Returning to the aforementioned fishermen, there is one obvious distinction; in February 2020, Howell was *following* 11,513 users on Twitter, while Zona was *following* a mere 155. This comparison introduces the central question addressed in this research: is there value in considering the number of other social media platform users someone is *following* (henceforth referred to simply as *following*)? In other words, does *following* tell us anything about how influential a social media user might be?

The premise of this work is that a substantial number of social media users will attend to these numbers and that this can have meaningful downstream consequences. We show how many people a user is *following* on social media affects how other users respond to the content the user shares. More specifically, we observe social media users are more *engaged* (in terms of likes and retweets) with content shared by those *following* fewer others. This is due, at least in part, to the fact that by *following* fewer others the user signals s/he is less susceptible to outside influence and thus more autonomous. Given one's own influence is often negatively correlated with one's susceptibility to the influence of others (Aral and Walker 2012; Iyengar, Van den Bulte, and Valente 2011), it stands to reason that someone *perceived* as more autonomous would also be perceived as more influential. In turn, being perceived as influential matters for engagement because, as Rogers and Cartano (1962, p. 439) pointed out long ago, people's perception of someone's influence is in large part "actually what affects behavior."

By documenting how *following* serves as a cue of influence, this research contributes to the literature in four important ways. First, theoretically, we answer the call for further research on the causal mechanisms of social influence online (Aral 2011) while adding to the literature that has examined characteristics of influentials (Rogers and Shoemaker 1971; Kopller 1984). We do so by identifying a previously unstudied characteristic driving perceptions and behavioral responses to influencers, perceptions of autonomy. We leverage on the idea of a "two-step flow of communication" (Katz and Lazarsfeld 1955) that implies communication flows from a source to opinion leaders who pass it on to others in the social system. In an age with unlimited sources of information online, we find too many sources in step one is associated with less perceived influence—and thus less engagement—in step two. Importantly, in doing so, we bridge two related yet distinct approaches used to study influence: an *individual-based* approach of identifying personal characteristics of influentials (i.e., autonomy) with a *network-based* approach of identifying sociometric measures of influence (*following*, or out-degree centrality).

Second, we contribute to an emerging literature on informative cues and inference making in digital environments (Ranganathan 2012; Berger and Barasch 2018; Li, Chan, and Kim 2019; Grewal and Stephen 2019). Digital environments often present significant ambiguity, which drives the use of contextual cues in online decision making and opinion formation (Ranganathan 2012). With social media absorbing an increasing amount of social interaction, it is important to understand how what individuals do online signals aspects of the self to others. We find *following* fewer others on a social media platform is an effective cue of an individual's autonomy, and thus influence.

A third contribution of this research is to add to a nascent body of work investigating the positive signaling effects associated with being seen as autonomous, or acting according to one's own inclinations (Bellezza, Gino, and Keinan 2014; Warren and Campbell 2014). Lacking direct access to the internal states of others, observation is the primary way in which people make inferences about the autonomy of others (Ryan and Connell 1989). *Following*, an easily observable characteristic of social media users, signals autonomy and drives important downstream consequences.[6]

Finally, this work documents the value of incorporating *following* as a useful criterion for screening influencers and micro influencers in particular. Given the latter are defined in part by their limited reach (i.e., number of followers), identifying additional indicators of influence is especially important and makes a substantive contribution to the field.[7] A review of 40 recognized influencer identification platforms revealed only three included *following* as a criterion they publicize offering prospective clients to compare micro influencers (see details in Web Appendix). If *following* is useful in assessing social media influencers (henceforth referred to as *influencer*), it appears not widely known or endorsed by practitioners.

[6] Table A1 in the Web Appendix provides a brief summary of related research on identifying influentials, informative cues in inference-making in digital environments, and the positive signal effects of autonomy.
[7] Table A2 in the Web Appendix provides a brief summary of related research on social media influencers drawn from marketing's top journals.

RELEVANT LITERATURE

Influentials

Influentials, or opinion leaders, are people who exert an extraordinary amount of influence on the attitudes and behaviors of others (Katz and Lazarsfeld 1955; Merton 1968). The level of attention directed toward studying these individuals by marketing researchers is due in large part to the belief that what they have to say affects what others purchase (Rogers and Cartano 1962) and ultimately a product's success or failure (Rogers 1962). While the literature on peer-to-peer influence and opinion leadership is vast and dates back more than a half century, two broad streams of research are of particular interest here. These include research examining: (1) characteristics of influentials, and (2) how to identify who is in fact influential.

First, an important characteristic of an influential is what type of information s/he transmits and in what domain his/her influence is exerted. Some literature suggests opinion leaders focus on specific topics, thereby being *monomorphic* (Engel, Kollat, and Blackwell 1968; Jacoby 1974). However, other literature highlights opinion leaders' influence can extend to a variety of (sometimes unrelated) topics, which is characteristic of being *polymorphic* (Marcus and Bauer 1964; King and Summers 1970; Myers and Roberston 1972). By their nature, micro influencers typically start out being monomorphic and later often evolve into being polymorphic as their popularity grows (e.g., sharing make-up advice initially and later fashion advice as well).

Besides identifying the boundaries of their influence, extensive research has looked at personal characteristics that make someone an opinion leader (see Keller and Berry 2003 for a review). Considering demographic characteristics, according to Weimann et al. (2007), influentials can be found at every social level, across the sexes, and in all professions and age groups. This is consistent with the heterogeneity one might expect for micro influencers online. Research has also shown expertise is often an antecedent of opinion leadership (Grewal, Mehta, and Kardes 2000) and that innovative consumers are more likely to be opinion leaders than consumers with conservative characteristics (Ruvio and Shoham 2007). Importantly, we show how *following* fewer others affects perceptions of autonomy, but neither perceived expertise nor innovativeness.

A second broad stream of research has focused less on understanding characteristics of an opinion leader and more on how to identify influential consumers. A number of approaches have been documented in the literature (see Weimann, et al. 2007). One popular method historically has been self-designation, which has contributed to the development of various opinion leadership scales (Rogers and Cartano 1962; King and Summers 1970; Childers 1986; Flynn, Goldsmith, and Eastman 1996). More recently, as consumers have become progressively more interconnected on social media, a different approach has increased in prominence, one that focuses on analyzing the structure of a network.

Social Networks and Influence

One way to attempt to identify influentials within a network is to apply sociometric techniques to capture relationships between members of the social system. Conventional wisdom suggests that highly connected nodes within a network should disproportionally contribute to the spread of information and thus promote product adoption. Consequently, a variety of measures of network centrality have been proposed (see Muller and Peres 2019 for a review), and research has shown influence is often more strongly associated with network centrality than commonly-used self-reports of opinion leadership (Iyengar, Van den Bulte, and Valente 2011).

Other work comparing seeding strategies based on centrality supports seeding well-connected people. For example, using controlled field experiments, Hinz and colleagues (2011) compare different seeding strategies and find seeding "hubs," individuals connected with many others (i.e., high degree centrality), is the most successful strategy, but note this is because of their extensive reach rather than because these individuals are more persuasive. What this suggests, and the opening example involving Selena Gomez illustrates, is that relying on expansive reach (i.e., number of *followers*) alone can be inefficient, and considering which well-connected people are more or less influential is important. This led us to investigate an additional, alternative cue for influence within a social network—*following*. In doing so, we distinguish between two network characteristics, the number of inbound links (*followers*, or in-degree centrality) and outbound links (*following*, or out-degree centrality) as separate indicators of influence, each informative in its own right.

Following Others as a Cue of Autonomy and Influence

In the absence of complete information, consumers often rely on signals, or cues, to make inferences that allow them to form opinions and make decisions (Huber and McCann 1982). We propose consumers use the number of other users someone is following as one such cue, one that signals autonomy. Autonomy refers to the extent to which people act in alignment with their values, unaffected by others' influence (Brehm 1993; Deci and Ryan 1985, 2000; Ng 1980; Schwartz 1992). In the marketing literature, autonomy has been defined as "a willingness to pursue one's own course irrespective of the norms, beliefs, and expectations of others" (Warren and Campbell 2014, p. 544).

Past research supports the idea individuals care a great deal about being perceived as autonomous, to the point that "people are more concerned with managing the impression of autonomy than they are with actually maintaining autonomy" (Baer et al. 1980, p. 416). Moreover, it is particularly important to influentials that they be seen as formulating their own opinions unadulterated by the influence of others (Dworkin 1988). This is consistent with findings that influence is often negatively correlated with one's susceptibility to the influence of others (Aral and Walker 2012; Iyengar et al. 2011) and may help to explain why, in the U.S., the idea of being autonomous is aspirational (Markus and Schwartz 2010) while being easily influenced by others is not (Jetten, Hornsey, and Adarves-Yorno 2006). In this work, we propose that following others on social media affects the extent to which someone is *perceived* to be autonomous, which in turn affects *perceptions* of being influential.

THE CURRENT RESEARCH

To summarize our theorizing, we propose *following* can be an important cue that helps distinguish more versus less effective influencers on a social-media platform. This is, at least in part, because individuals see those *following* fewer others as more autonomous. Given the naturally occurring negative correlation between one's own influence and one's susceptibility to the influence of others, we propose other users infer more autonomous individuals are also more influential. Importantly, we propose this inferential process has important implications for marketers since *following* fewer others can have significant downstream consequences in terms of social media engagement. By viewing certain users as more influential at the onset, consistent with theories of social influence (Salganik, Dodds, and Watts 2006), we propose other users are more likely to engage with content they post in terms of both likes and retweets (as well as click-throughs when a link is available). A schematic of our conceptual model along with how each study supports the model is illustrated in Figure 1. Table 1 features an overview of our five laboratory studies and main findings.

Insert Figure 1 here.

Insert Table 1 here

We begin our empirical process by analyzing real world social media data drawn from Twitter. In Study 1, we observe a negative correlation between the number of *likes* and *retweets* received by a particular Twitter post and the number of users the source of the post is *following* at the time, ceteris paribus. Next, we present four lab studies that seek to clarify and explain the role of *following* in how people respond to social media users. Study 2 demonstrates that *following* fewer others increases perceptions of an individual's influence, conditional on the individual having a substantial number of followers. In Study 3, we show autonomy mediates the relationship between *following* and perceptions of influence. We also show downstream consequences in terms of respondents' engagement replicating the effects observed in Study 1. In Study 4, we again replicate the effect observed in Study 1 and provide additional evidence of process through moderation, while in Study 5 we test the effect of *following* on a more consequential behavioral indicator of engagement, namely click-through.

STUDY 1

In Study 1, we use data obtained from Twitter to test the effect of *following* on how others respond to a user's post, ceteris paribus. We predict that the fewer users an individual is following, controlling for other factors, the greater the engagement his/her tweets will get in terms of both *likes* and *retweets*. Our focal independent variable is *Following* while our dependent variables include two different measures of engagement: how positive followers are toward the content (*Likes*) and how many times the content is shared (*Retweets*).

Data

The data utilized in the analysis include all tweets written in English on September 16, 2016 in a major metropolitan area on the West Coast.[8] We collected all of the data directly from *twitter.com* over a 3-day period (September 20-22, 2016). Twitter allows anyone to collect data

[8] In the Web Appendix, we show that the results replicate using alternative datasets, one comprising tweets from all over the U.S. and one comprising tweets from the Tokyo (Japan) metropolitan area.

about real-time tweets and past tweets (up to a week old) as well as user profiles through their public API. The data include 1,581,522 tweets from 784,170 distinct users as well as a wide-ranging set of features of the tweet and the user provided by the API, described next.

Tweet features. For every tweet in our dataset, we know the number of *Likes* it received and the number of *Retweets*. Moreover, we have additional information about its content, including the number of links to websites (URLs), videos, photos, financial symbols (e.g., "$", "TSLA"), user mentions, and hashtags. Further, we collected the timestamp of publication and whether the tweet is an original posting, a retweet, or a reply to someone else's tweet. Of the 1,581,522 tweets we collected, 447,793 are original tweets, 969,488 are retweets, and 164,241 are replies. Given our interest is on how others respond to original content, we focus on original tweets and remove replies and retweets from the data.[9] The 447,793 original tweets were produced by 146,444 users.

User profile features. For each tweet, we collected information regarding the user who posted the tweet, including our focal independent variable of interest, the number of fellow Twitter users s/he follows (*Following*). We also collected a number of control variables including the user's ID and screen name, a timestamp for the creation of the user's account (measuring the time the user had been on Twitter), the number of users who follow him or her (*Followers*), the total number of tweets ever written, the total number of likes ever given, and whether the account is "verified" (a verification badge assures other users an account is authentic). We also collected the length of the user's profile bio, whether it contained a URL, and whether the user has chosen to personalize his or her profile and image.

[9] In the robustness checks section, we show that our results hold even using the full sample of tweets (see column 5 of Tables 4 and 5).

Given our focus is on micro influencers, we restricted the dataset to users with at most 100,000 followers, resulting in the exclusion of 8,742 tweets by 1,325 users.[10] The final dataset includes 439,051 original tweets by 145,119 Twitter users with fewer than 100,000 followers.

Content features. We computed linguistic features of the text of each tweet using LIWC (Pennebaker et al. 2015), a program used for automated text analysis. LIWC categorizes words along several dimensions including different emotions, thinking styles, social concerns, and parts of speech. Among the standard variables in LIWC's default dictionary are social and psychological states such as positive and negative emotions, anxiety, anger, and sadness. The standard output includes the percentage of words in the text pertaining to that variable. Past literature has shown that the way individuals react to content shared online by others is often a function of identifiable linguistic features of the content (Berger and Milkman 2012). In particular, this literature has identified *Positivity*, *Anxiety*, *Anger*, *Sadness*, and *Arousal* as relevant text characteristics resulting in content being shared more often (i.e., virality).[11] Given *Retweets* is one of our dependent variables, we compute these metrics for each tweet to include as covariates. We use LIWC for the first four variables and the dictionary and word values provided by Warriner, Kuperman, and Brysbaert (2013) to compute *Arousal* scores.[12]

Descriptive Statistics

The average number of tweets per user in our dataset is 3 (SD = 10.2). The average number of *Followers* per tweet is 3,224 (SD = 9,604), and the average number of other users someone is *Following* per tweet is 1,362 (SD = 4,643). The average number of *Likes* per tweet is 2.80 (SD =

---

[10] In the robustness checks section, we show that our results hold using alternative thresholds to identify micro-influencers, as well as analyzing the full sample (see columns 1-4 of Tables 4 and 5).

[11] The Positivity index is calculated as the difference between the scores (percentages) for positive and negative emotion words and is computed on a scale from 1 to 100.

[12] Unlike with LIWC, the arousal metric is not a percentage. Every word in the dictionary is associated with an arousal value ranging from 1 to 10, and a tweet's arousal is calculated as the average arousal value for all of the words in that tweet.

36.8), and the average number of *Retweets* is 1.1 (SD = 43.8). The relative size of the standard deviations suggests the distributions of these four variables are extremely skewed.

Looking at the content of the tweets (see summary statistics in Table 2), we observe that hashtags and URLs are included more often than videos and photos, and financial symbols are rarely used. Moreover, the LIWC analysis reveals the emotional content of our tweets is, on average, relatively neutral. Finally, we observe the accounts in our dataset are approximately four years old, on average, and only two percent of tweets come from *verified* accounts. We present the correlation matrix between all of the variables in Table A3 of the Web Appendix.

Insert Table 2 here.

The Effect of Following on Likes/Retweets

To estimate the effect of *Following* on *Likes* as well as *Retweets*, we use negative binomial regression for two reasons. First, the dependent variable is a count (number of *Likes*, *Retweets*). Second, both outcome variables are over-dispersed (the variance for each is much larger than the mean). The base model takes the following form:

$$Y_{ijt} = \beta_1 \, log \, Following_{jt} + \beta_2 \, log \, Followers_{jt} + X'_{ijt}\gamma + \epsilon_{ijt} \qquad (1)$$

in which the dependent variable is either the number of *Likes* or *Retweets* received by tweet *i* written by user *j* at time *t*. The focal independent variable is log $Following_{jt}$, the (log) number of others that user *j* is following at time t. In addition to *Following*, in the model, we include log $Followers_{jt}$ the (log) number of users that follow user *j*. Thus, the coefficient of interest, β_1, measures the effect of *Following* on *Likes* holding constant *Followers*. Further, we include a number of additional covariates, $X_{jt,}$ in our regression. These are described next.

Tweet Controls

The first set of control variables included relates to features of the tweet, namely the number of user mentions, URLs, images, videos, financial symbols, and hashtags present in the tweet.[13] We also include the tweet length (word count), and measures of *Positivity*, *Anxiety*, *Anger*, *Sadness,* and *Arousal* derived from the LIWC analysis. Finally, recall the tweets in our dataset were all posted on September 16, 2016, but our data collection spanned three days from September 20-22, 2016. The difference between posting time and collection time can affect *Following*, and importantly, the number of *Likes/Retweets* a tweet receives. Thus, we include the (log) time difference (in minutes) between the time the tweet was posted and when we collected the data (i.e., difference between the two timestamps) in our regression.

User Profile Controls

We also include a set of controls related to characteristics of the Twitter user. One important factor is the user's experience within the social network, because a more experienced user might be able to write tweets that receive more *Likes* and *Retweets*. While experience is unobserved and generally difficult to measure, we utilize three separate measures as proxies. First, we include (log) *User Age*, which corresponds to the number of months between the tweet's publication and the date of the creation of the user's account. Second, we include the (log) total number of tweets written by the user prior to the focal tweet (*User Tweets*). Third, we include the total number of *Likes* the user has given (*User Total Likes*).

Another user control variable we incorporate in the model is an indicator of whether the user has a verified account (*User is Verified*). Verified Twitter accounts have a blue check mark next to the username, making verified users easily recognizable to other users. Because these accounts are associated with public or popular figures (e.g., businesses, celebrities), verified users

[13] We do not log these variables because they are relatively small and not very skewed, but results hold even if logged.

may be more likely to receive *Likes* and/or *Retweets*. Finally, we control for the length of the user bio, whether the user bio contains URLs, and whether the user has kept the default profile and default profile image versus having customized these elements.

Results

We estimate equation 1 with standard errors clustered at the user level to account for potential correlation across tweets written by the same user. We report the results for *Likes* in Table 3 and *Retweets* in Table 4. In column 1, we report the coefficients obtained with the simplest model, which controls for the number of *Followers* exclusively. As predicted, β_1, the coefficient of (log) *Following* for *Likes* is negative and statistically significant (-.24, $p < .001$). This estimate suggests a one percent increase in the number of users the person is *Following* decreases the number of *Likes* a tweet receives by approximately .24 percent. For *Retweets*, the coefficient of (log) *Following* is also negative and statistically significant (-.25, $p < .001$).

These results are in line with our predictions. Not surprisingly, we also find that β_2, the coefficient of log *Followers*, is positive and statistically significant (.63, $p < .001$ for *Likes*; .73, $p < .001$ for *Retweets*). This suggests that the number of *Followers* a user has positively affects the number of likes and retweets a tweet receives. Someone with more *Followers* is expected to have more people reading their tweets and consequently liking/retweeting their posts.

In columns 2 and 3, we report the results of the model incorporating the control variables described above. The estimates for β_1 remain negative, statistically significant, and similar in magnitude to those presented in column 1, suggesting the results are robust to the inclusion of several characteristics of the tweets and the users who posted them.

Insert Table 3 here.

Insert Table 4 here.

These results are consistent with the prediction that *Following* fewer (vs. more) others on a social media platform affects how others respond to content shared. All else being equal, tweets from Twitter users *following* fewer others are more prone to be liked and retweeted.

The Moderating Effect of Number of Followers

It stands to reason that for *following* to matter, the user must first have a substantial number of *followers*, implying a significant number of people want to hear what that person has to say. The assumption is that they accumulated many of these followers due to tweets that included interesting and valuable content, and *followers* itself serves as an important cue for influence. While identifying the exact magnitude of a "substantial number of *followers*" is beyond the scope of this research, we can use our dataset to test whether the number of *Followers* matters. An intuitive way to do so is to add the interaction $log\,Following\ \times log\,Followers$ to our main model. We present the results of this specification in column 4 of Tables 3 and 4. The interaction coefficient is negative for both dependent variables of interest and statistically significant when the dependent variable is *Likes*, suggesting a moderating effect of number of *Followers* (at least in the case of *Likes*). There are, however, a number of difficulties in interpreting the interaction between two continuous variables (Jaccard, Wan, and Turrisi 1990).

An alternative way to analyze the interaction that is easier to interpret is to estimate the model and then compute simple slopes (i.e., the slopes of the independent variable *Following* on the dependent variables *Likes/Retweets* when the moderator variable *Followers* is held constant at different values). We provide a series of plots of these slopes in Figures 2 and 3. In line with our expectations, we observe that the negative effect of *Following* on *Likes* and *Retweets* becomes easily discernible at roughly 1,000-5,000 *Followers*; it does not take tens, or hundreds of thousands of *Followers* before *Following* fewer others appears to matter.

Insert Figure 2 here.

Insert Figure 3 here.

Robustness Checks

In this section, we discuss several tests of the robustness of our results, including to different subsamples of data, a different modeling approach, and datasets collected from alternative geographical areas. First, recall our original analysis included users with 100,000 *Followers* or fewer and focused only on original tweets. In columns 1-4 of Table 5 (*Likes*) and Table 6 (*Retweets*), we show results are robust to different thresholds of *Followers* (10,000, 50,000; 150,000; and 200,000). In column 5 of Table 5 (*Likes*) and Table 6 (*Retweets*), we show they are robust to the inclusion of all users and all tweets (original, retweets, and replies).

Second, we test whether our results are robust to a different modeling approach. While our main analysis used negative binomial regression, we replicate the results using OLS regression (see Tables A4 and A5 in the Web Appendix). Additionally, because our dependent variables are correlated, we replicate the results using a bivariate negative binomial regression model (see Table A6 in the Appendix).

Finally, we replicate the results using alternative datasets scraped from different geographical areas and different points in time. These include a sample of tweets from all over the U.S. written on February 18, 2019, and a sample of tweets from the Tokyo (Japan) written between March 9 and March 11, 2019 (see Tables A7-A10 in the Web Appendix).

Discussion

Taken together, the results presented in Study 1 provide compelling correlational evidence from real world data that, ceteris paribus, content shared by a social media user *following* fewer others garners greater engagement in terms of more *likes* and *retweets*. In the lab studies that follow, we replicate these findings in a controlled setting while investigating an important contributory explanation as to why this occurs.

STUDY 2

In Study 2, we begin by testing the principal hypothesis that *following* affects perceptions of an individual's influence. We also test a proposed moderator for the effect discussed in our secondary data analysis, namely that for *following* to operate effectively as a signal of influence, the user must have a substantial number of *followers*. Simply put, a person needs to be seen as someone worth listening to (i.e., one with a significant number of *followers*) before the number of people they listen to (i.e., *following*) matters.

Method

Respondents were 276 undergraduate students (49.6% female, $M_{age} = 20.5$) who completed the study for partial course credit. The study employed a 2 (*Following*: *High* vs. *Low*) x 2 (*Followers*: *High* vs. *Low*) between-subjects design. Respondents were asked to evaluate a social media user based on the individual's profile page (see Web Appendix for all experimental stimuli). To enhance generalizability, we use Instagram as the domain in this study. The stimuli replicated the features of a real Instagram page including the user's Instagram name, picture, and number of photos posted (309).[14] What varied across conditions was the number of *Followers* (58 or 15,457) and *Following* (49 or 21,530). These numbers were selected based on the distribution of the data collected in Study 1 such that in the *Low* condition they would fall around the 10th percentile and in the *High* condition would fall between the 90th and 95th percentile.[15] The number of *Followers* was below the threshold we identified for detecting an effect of *Following* in study 1 in the *Low* condition and above such threshold in the *High* condition.

[14] The average number of Instagram posts in a sample of 20 million Instagram users (see Jang, Han, and Lee 2015).
[15] A separate study presented in the Web Appendix shows the effect is not sensitive to the choice of these specific values and the negative relationship between *following* and *perceived influence* holds across a wide range of *following*.

To assess how *Following* affects *Perceived Influence*, respondents were asked "To what extent do you think this user is influential on Instagram?" (1 = Not at All Influential, 9 = Very Influential). Further, we also asked respondents to rate the user on six domain-specific items adapted from the Opinion Leadership scale by Flynn, Goldsmith, and Eastman (1996, see Web Appendix). This measure was included to assess the robustness of our effect across a polymorphic measure of influence (the former) and a monomorphic measure, specific to the domain of travel (the latter). These DVs are highly correlated (r = .60) and present similar findings. Thus, in subsequent studies, we focus exclusively on the concept of perceived influence.

In this, and in all subsequent studies, before exiting the study, respondents were asked their age and gender as well as to recall the number of users followed by the user they just evaluated (see measures in Web Appendix).[16]

Results

A between-subject ANOVA with *Perceived Influence* as the dependent variable reveals a significant main effect of both *Following* and *Followers*. As expected, the number of *Followers* has a significant effect on perceived *Influence* ($F(1, 272) = 257.45$, $p < .001$) as does *Following* ($F(1, 272) = 8.42$, $p = .004$). More importantly, consistent with our theorizing, these effects are qualified by a significant interaction ($F_{Interaction}(1, 272) = 11.02$, $p = .001$). Simple contrasts reveal that *Following* has a significant effect on *Perceived Influence* in the *High Followers* condition ($M_{Low\ Following} = 6.09$, 95% CI = [5.67, 6.51] vs. $M_{High\ Following} = 4.80$, 95% CI = [4.40, 5.20], $F(1, 272) = 18.93$, $p < .001$) but not in the *Low Followers* condition ($M_{Low\ Following} = 2.07$, 95% CI = [1.69, 2.45] vs. $M_{High\ Following} = 2.16$, 95% CI = [1.71, 2.61], $F(1, 272) = .09$, $p = .765$). These results are plotted in Figure 4.

[16] Across studies, the number of individuals who failed to recall the number of *Following* accurately varied between 9% and 16%. For simplicity in reporting, these respondents were not excluded from any of our analyses. Importantly, the results are substantively the same if we exclude those who failed to recall the number of *Following* correctly.

Insert Figure 4 here.

As expected, a similar pattern of results is observed when *Opinion Leadership* (α = .88) is the dependent variable (see Figure 5). Both *Following* (F(1,272) = 5.03, *p* = .026) and *Followers* (F(1, 272) = 137.13, *p* < .001) predict *Opinion Leadership*. Again, these effects are qualified by a significant interaction ($F_{\text{Interaction}}(1, 272) = 4.23$, *p* = .041). Simple contrasts reveal that *Following* affects perceived *Opinion Leadership* significantly in the *High Followers* condition ($M_{\text{Low Following}}$ = 5.46, 95% CI = [5.17, 5.74] vs. $M_{\text{High Following}}$ = 4.71, 95% CI = [4.43, 5.00], *F*(1, 272) = 9.04, *p* = .003), but not in the *Low Followers* condition ($M_{\text{Low Following}}$ = 3.08, 95% CI = [2.71, 3.45] vs. $M_{\text{High Following}}$ = 3.05, 95% CI = [2.63, 3.46], *F*(1, 272) = .02, *p* = .894).

Insert Figure 5 here.

Discussion

Results from Study 2 provide initial evidence in support of our conceptual model by demonstrating how *following* fewer others can affect the perceived influence (opinion leadership) of a social media user. These results are noteworthy given practitioners' desire for cues other than *followers* to assess an influencer's potential. Moreover, we show the impact of *following* is contingent on having accumulated a substantial number of *followers*. Next, we investigate a key mechanism underlying the effect. We demonstrate *following* fewer others affects perceptions of an individual's autonomy, which in turn drives perceptions of influence.

STUDY 3 PRE-TEST

We first designed a pre-test to explore the basic proposition that *following* affects perceptions of the user's autonomy, our proposed mediator, along with two potential alternative explanations for our effect, namely that someone *following* fewer others is perceived: (1) as more of an expert, and/or (2) as more innovative in general. Previous literature has established that expertise (Grewal, Mehta, and Kardes 2000) and innovativeness (Ruvio and Shoham 2007) are both viewed as characteristics of an opinion leader. It is thus possible that *following* could serve as a cue of expertise and/or innovativeness, which could, in turn, affect perceptions of influence.

Participants included 598 Twitter users (45.7% female, M_{age} = 30.0) enlisted via Amazon mTurk in exchange for a .50 USD payment. Of these, 198 were asked to evaluate a Twitter user's autonomy, 200 were asked to evaluate a Twitter user's expertise, and 200 were asked to evaluate a Twitter user's innovativeness. As in earlier studies, respondents evaluated a Twitter user based on a snapshot of the person's profile page. The number of *Followers* (15,457), *Tweets* (7,835), and the number of *Likes* (916) remained constant across conditions. The number of *Followers* corresponded to our High Followers condition in Study 2, *Tweets* and *Likes* reflected averages in the dataset used in Study 1. We varied *Following* to be either *Low* (49) or *High* (21,530). Respondents evaluated the Twitter user either in terms of *Autonomy*, *Expertise*, or *Innovativeness* (each respondent provided only one type of evaluation to avoid common method bias). We adapted the Deci and Ryan (2001) autonomy scale to measure autonomy, the Ohanian (1990) expertise scale to measure expertise, and the Goldsmith and Hofacker (1991) consumer innovativeness scale to measure innovativeness (see Web Appendix).

Results

Autonomy (α = .89). Respondents perceived the user to be significantly more autonomous when *Following* was *Low* as opposed to *High* ($M_{\text{Low-Following}}$ = 6.07, 95% CI = [5.73, 6.41] vs. $M_{\text{High-Following}}$ = 5.35, 95% CI = [5.02, 5.68], $F(1,196) = 9.23$, $p = .003$).

Expertise (α = .97). Respondents did *not* perceive the Twitter user to be significantly more expert when *Following* was *Low* as opposed to *High* ($M_{\text{Low-Following}}$ = 5.77, 95% CI = [5.35, 6. 19] vs. $M_{\text{High-Following}}$ = 5.70, 95% CI = [5.28, 6.11], $F(1,198) = .05$, $p = .824$).

Innovativeness (α = .87). Respondents did *not* perceive the Twitter user to be significantly more innovative when *Following* was *Low* as opposed to *High* ($M_{\text{Low-Following}}$ = 6.98, 95% CI = [6.73, 7.22] vs. $M_{\text{High-Following}}$ = 7.03, 95% CI = [6.78, 7.28], $F(1,198) = .09$, $p = .766$).

Discussion

Taken together, these results suggest that while *following* fewer others on social media signals autonomy, it does not appear to signal expertise or innovativeness. Turning back to our real world data, we also observe that users who follow fewer others on Twitter are more inclined to use the first person pronoun "I" and less inclined to use second or third person pronouns ("we" "She/he") in their tweets (the analysis and results are reported in Table A11 of the Web Appendix). This suggests those who follow fewer others on Twitter may in fact be more autonomous with respect to the content they post; thus, *following* may serve as an accurate and reliable cue of autonomy.

STUDY 3

In Study 3, we directly test whether autonomy mediates the relationship between *following* fewer others and perceptions of influence. Additionally, while in study 2 respondents evaluated users in the absence of any content, we now increase external validity by including content along with the user profile. This is important as it allows us to measure respondents' engagement with the content shared (tweet) in terms of engagement (*Likes* and *Retweets*).

Method

Respondents were 315 undergraduate students (50.8% female, M_{age} = 20.4) who completed the study for partial course credit. As in earlier studies, respondents evaluated a Twitter user based on a snapshot of the person's profile page. The number of *Followers* (15,457), *Tweets* (7,835), and

Likes (916) remained constant as in previous studies. We varied *Following* to be either *Low* (49) or *High* (21,530). The profile included the latest tweet shared by the user that read: "Top ten songs of all times" accompanied by a URL link. After viewing the profile, respondents evaluated the *Perceived Influence* and *Autonomy* of the user using the same measures employed in previous studies. Importantly, in this study, they were also asked to evaluate the content of the tweet by reporting how likely they would be to *Like* and *Retweet* it (1 = Not at All Likely, 9 = Very Likely).

Results

Autonomy (α = .85). Respondents perceived the Twitter user to be significantly more autonomous when *Following* was *Low* as opposed to *High* ($M_{\text{Low-Following}}$ = 5.90, 95% CI = [5.68, 6.13] vs. $M_{\text{High-Following}}$ = 5.48, 95% CI = [5.23, 5.73], $F(1,313) = 6.01$, $p = .015$).

Influence. A between-subject ANOVA with *Influence* as the dependent variable reveals a significant main effect of *Following* such that those following fewer others were perceived as more influential ($M_{\text{Low-Following}}$ = 5.33, 95% CI = [5.04, 5.61] vs. $M_{\text{High-Following}}$ = 4.13, 95% CI = [3.85, 4.42], $F(1,313) = 34.61$ $p < .001$).

Likes. With respect to the content, respondents were more prone to *Like* the post when *Following* was *Low* as opposed to *High* ($M_{\text{Low-Following}}$ = 3.14, 95% CI = [2.81, 3.47] vs. $M_{\text{High-Following}}$ = 2.44, 95% CI = [2.16, 2.72], $F(1,313) = 10.31$, $p = .001$).

Retweets. Respondents were also more prone to *Retweet* the post when *Following* was *Low* as opposed to *High* ($M_{\text{Low-Following}}$ = 2.40, 95% CI = [2.11, 2.69] vs. $M_{\text{High-Following}}$ = 1.89, 95% CI = [1.67, 2.11], $F(1,313) = 7.72$, $p = .006$).

Mediation. We ran the sequential mediation model: *Following->Autonomy->Influence->Likes* using a bootstrap estimation approach with 5,000 bias corrected samples (Hayes 2013, model 6). In line with our conceptual model, we observe a significant indirect (b_{indirect} = -.03, 95%

CI = [-.084, -.006]) effect. Substituting *Retweets* as our dependent variable, we observe similar results ($b_{indirect}$ = -.02, 95% CI = [-.068, -.004]).

Discussion

Taken together, the results from Study 3 show *following* fewer other users on social media leads to greater perceptions of autonomy and thus influence. This leads to greater engagement manifested as a more positive attitude (*likes*) toward shared content and a greater propensity to share (*retweets*) the content. Thus, we replicate the results of Study 1 in a controlled setting in which we also test the proposed underlying mechanism, being seen as more autonomous and thus more influential.

STUDY 4

While in Study 3 we provide evidence of process through mediation, in this study, we provide additional evidence of process through moderation. We also identify an important boundary condition for the effect. Specifically, if outside information is available, that suggests a person is indeed influential and heuristic processing should no longer be a useful effort-reducing mechanism (Shah and Oppenheimer 2008); thus, *following* is less likely to serve as a useful cue. Its effect on engagement will therefore be attenuated.

Method

Respondents were 703 undergraduate students (47.8% female, M_{age} = 20.5) who completed the study for partial course credit. This study followed a 2 (*Following*: High vs. Low) by 2 (*Influence Information*: Yes vs. No) between-subjects design. As in earlier studies, respondents evaluated a Twitter user based on a snapshot of the person's profile page. The stimuli used were the same as Study 3. We varied *Following* to be either *Low* (49) or *High* (21,530). In this study, we also manipulated whether additional information about the user was provided. With respect to outside information about influence, in the *Yes* condition, respondents read a brief introduction of the Twitter user: "Robert Diaz is an influential and well respected music journalist" and were shown a screenshot of a magazine referring to Robert Diaz as one of the top influential music writers. In the *No* condition, no such information was provided, mirroring what we have done in earlier studies. This manipulation was pre-tested to ensure that providing such information would increase the perceived influence of a social media user in the absence of information regarding the user's *followers* and *following* (see Web Appendix).

Respondents subsequently reported how likely they would be to *Like* and *Retweet* the accompanying tweet (1 = Not at All Likely, 9 = Very Likely). A priori, we expected *following* fewer others would impact engagement only when respondents were not informed the user was an influential and well-respected journalist.

Results

Likes. We ran an ANOVA predicting *Likes* with *Following* and *Influence Information*. Both the main effect of *Following* ($F(1,699) = 9.81$, $p = .002$) and that of *Influence Information* ($F(1,699) = 69.09$, $p < .001$) reached statistical significance. Most importantly, as expected these main effects were qualified by a significant interaction ($F(1,699) = 6.21$, $p = .013$). When no *Influence Information* was provided, respondents were more prone to *Like* the post when *Following* was *Low* as opposed to *High* ($M_{\text{Low-Following}} = 4.30$, 95% CI = [3.95, 4.65] vs. $M_{\text{High-Following}} = 2.89$, 95% CI = [3.02, 3.61], $F(1,669) = 15.84$, $p < .001$). However, this was not the case when outside *Influence Information* was available ($M_{\text{Low-Following}} = 5.32$, 95% CI = [4.95, 5.70] vs. $M_{\text{High-Following}} = 5.21$, 95% CI = [4.85, 5.58], $F(1,669) = .20$, $p = .651$).

Retweets. Similar results were observed with *Retweet* as the dependent variable. Again, the main effects of *Following* ($F(1,699) = 4.68$, $p = .031$) and that of *Influence Information* ($F(1,699)$

= 43.75, $p < .001$) were qualified by a significant interaction ($F(1,699) = 8.88$, $p = .003$). When no *Influence Information* was provided, respondents were more prone to *Retweet* the post when *Following* was *Low* as opposed to *High* ($M_{\text{Low-Following}} = 3.61$, 95% CI = [3.28, 3.95] vs. $M_{\text{High-Following}} = 2.71$, 95% CI = [2.44, 2.97], $F(1,699) = 13.25$, $p < .001$). This was not the case when *Influence Information* was available ($M_{\text{Low-Following}} = 4.26$, 95% CI = [3.86, 4.65] vs. $M_{\text{High-Following}} = 4.40$, 95% CI = [4.02, 4.78], $F(1,699) = .33$, $p = .564$).

Discussion

The results of Study 4 provide evidence that engagement is higher for social media users who follow fewer others, but only when *following* serves as a cue for influence. They are also consistent with the opinion of Rogers and Cartano (1962) who point out that the perception of being influential can affects others' behavioral responses, thus resulting in actual influence.

STUDY 5

We designed Study 5 to test the effect of *following* on other users' engagement in a more behaviorally consequential manner. In this study, we thus focus on a different measure of engagement, namely click-through. We give respondents the opportunity to actually click on a link posted by a social media user and spend time exploring a list of personally recommended restaurants. We test whether click-through rates vary as a function of *following*.

Method

Respondents were 256 undergraduate students (47.7% female, $M_{age} = 20.0$) who completed the study for partial course credit. As in earlier studies, respondents evaluated a Twitter user based on a snapshot of the person's profile page. The number of *Followers* (15,457), *Tweets* (7,835), and *Likes* (916) remained constant across conditions. We varied *Following* to be either *Low* (49) or

High (21,530). The profile included the following content, displayed as the latest tweet shared by the user: "This is my list of 10 new restaurants to try in [university city]" together with a URL link. This content was expected to have the potential to elicit participants' interest because it was particularly relevant to respondents who were expected to be relatively new to the city in which the study was run.

Respondents first rated the perceived *Influence* of the Twitter user. Next, they were asked whether they wanted to *Click* on the link provided in the tweet to review the 10 restaurants recommended by the user or move to an unrelated task. Those who clicked the link were redirected to a list of 10 (real) new restaurants drawn from a popular food magazine. We expected respondents in the *Low* condition to show a greater interest in the restaurant recommendation made by the Twitter user (perceived as more influential) and therefore be more likely to click on the link.

Results

Influence. A between-subject ANOVA with perceived *Influence* as the dependent variable reveals a significant main effect of *Following* ($M_{\text{Low-Following}} = 5.77$, 95% CI = [5.46, 6.06] vs. $M_{\text{High-Following}} = 4.53$, 95% CI = [4.23, 4.84], $F(1,254) = 31.56$, $p < .001$).

Click. As anticipated, significantly more respondents chose to click on the link in the *Low* following condition than in the *High* following condition (53.9% vs. 40.6%, respectively, $\chi2$ (1, 255) = 4.53, $p = .033$).

Mediation. We ran the mediation model *Following-> Influence->Click* using a bootstrap estimation approach with 5,000 bias corrected samples (Hayes 2013, model 4) and observed a significant indirect effect ($b_{\text{indirect}} = -.23$, 95% CI = [-.475, -.049]).

Discussion

Study 5 replicates previous findings that *following* on social media can affect how others perceive someone and, in turn, impact engagement with the content they share. Importantly, this study reinforces the external validity of our findings and extends Study 1 by showing in a well-controlled laboratory setting that this simple cue of influence can affect consequential behavior, namely clicking on a link to view additional content suggested by the focal user. Intuitively, one might expect someone following more users to be able to draw information from more sources, leading them to be seen as more knowledgeable and making it more worthwhile to attend to their posts. Study 5 suggests this is not necessarily the case.

GENERAL DISCUSSION

Marketers are increasingly seeding information about their products and brands through individuals deemed to be influential on social media (i.e. influencers). Hence, practitioners and researchers alike have expressed interest in identifying what makes a social media user more (or less) influential (Van den Bulte 2010). This question is especially relevant as marketers shift from relying on internationally known celebrities, those with millions of *followers*, to micro influencers, online personalities with fewer, but presumably more loyal, followers (Maim 2017). Choosing among the myriad of micro influencers available is not easy, and a key question remains: What makes an influencer more or less likely to affect the behavior of his/her *followers*? We find that, holding a users' number of followers constant, an important indicator of this individual's online influence is the number of individuals this person chooses to follow.

Worth noting is that, in general, there is a correlation between the number of people a user follows and the number who follow that user. Reciprocity is a well-worn method of accumulating *followers*, so much so that Twitter has identified the practice of users *following* others solely to be followed in return and then employing algorithms to unfollow those *followers*. This practice offers anecdotal support for the central idea here that following fewer others is seen more positively by others. Otherwise, why go to such lengths to reduce the number one is *following*? Twitter has put a halt to the practice of bulk *following*; hence, *following* few others is a more reliable (less corrupted) signal on that platform once again.

There are limitations to this work worth mentioning as they might provide interesting directions for future research. First, we varied *followers* and *following* without offering precise guidance regarding what constitutes a substantial enough number of *followers* for *following* to be impactful. In our studies, we tried to use numbers that reflect actual numbers associated with micro influencers we studied online. The "right" numbers, we believe, are context dependent (i.e., depend on platform, topic category, influencer) and difficult to specify a priori with any degree of certainty. One can imagine, however, researchers in information systems attempting to address the question of the "right" numbers in various contexts using big data.

By limiting our focus to *following*, we intentionally did not consider other potential drivers of perceptions of an influencer online. For instance, we did not explore the nature of the connections between users on a social network. We do not study whether the accounts one follows and the accounts that follow a user matter at the time of inferring one's influence, or whether disclosing the identity of these accounts would affect one's perceptions. While two social media users may have the same number of *followers*, who those *followers* are, or who those users are choosing to follow (i.e., their network structures) may differ. This information is not transparent on a user's profile (it requires clicking on additional pages and scrolling) and therefore less salient to others compared to simply the number of *following* and *followers*. Yet, we believe this information might also affect how a social media user is perceived, and could be interesting to study in further research.

Future research may set out to test these as well as other drivers of perceived influence. Randy Howell from our opening example is a veteran of the Bassmaster Tour, the 2014 Bassmaster Classic Champion, and specializes in shallow-water fishing. Mark Zona is co-host of the

Bassmaster Elite Series on ESPN2. Each presumably has his own writing style and other idiosyncratic characteristics that differentiate him. How much weight do followers give to their accolades and other information? This seems context specific (bass fishing) while our goal was to instead try to identify a cue that applies more generally. But a richer model using language analysis and machine learning may one day incorporate all of this information.

Looking at the downstream consequence of influence perceptions studied in this work, we should also stress that perceived influence is only one factor expected to affect engagement. While we find *following* affects perceived autonomy and in turn perceived influence, we consider the mechanism we unveiled as contributing to this perception and not necessarily wholly explaining it. It is possible following can work as a cue for other user characteristics not ruled out in this work, which might be worth exploring in future research.

Abstracting away from *following* and *followers*, one might reframe our investigation in terms of numbers of sources and receivers of information. One can judge a person by the number of receivers who subscribe to the information they provide as well as by the number of sources they rely on for their information. How the number of sources and number of receivers affects perceptions of the source would seem to be important outside as well as inside of a social media setting. We believe more work could be done examining the effect of number of sources on receivers' responses outside of social media per se. The broader question of how many sources an information provider should draw from—and the impact that number has on the perceptions of those who receive information from the provider—is much broader than what we have done here and worthy of more investigation. While one might argue more sources are better, our results suggest that this intuition does not always hold true.

REFERENCES

Aral, Sinan (2011), “Identifying Social Influence: A Comment on Opinion Leadership and Social Contagion in New Product Diffusion,” *Marketing Science*, 30 (2), 217–223.

Aral, Sinan and Dylan Walker (2012), "Identifying Influential and Susceptible Members of Social Networks," *Science*, 337 (6092), 337-341.

Baer Robert, Steve Hinckle, Karen Smith, and Marsha Fenton (1980), “Reactance as a Function of Actual versus Projected Autonomy,” *Journal of Personality and Social Psychology*, 38 (3), 416-422.

Barker, Shane (2017), “Using Micro-Influencers to Successfully Promote Your Brand,” *Forbes*, Sept. 29, accessed online July 2, 2019 at https://www.forbes.com/sites/forbescoachescouncil/2017/09/29/using-micro-influencers-to-successfully-promote-your-brand/#4f0aa2971763

Bellezza, Silvia, Francesca Gino, and Anat Keinan (2014), “The Red Sneakers Effect: Inferring Status and Competence from Signals of Nonconformity,” *Journal of Consumer Research*, 41 (1), 35-54.

Berger, Jonah and Alixandra Barasch, (2018), “A Candid Advantage? The Social Benefits of Candid Photos,” *Social Psychological and Personality Science*, 9 (8), 1010-1016.

Berger, Jonah and Katherine L. Milkman (2012), “What Makes Online Content Viral?” *Journal of Marketing Research*, 49 (2), 192-205.

Bogost, Ian (2018), “All Followers Are Fake Followers,” *The Atlantic*, Jan. 30, Technology. Accessed on Sept. 25, 2018 at https://www.theatlantic.com/technology/archive/2018/01/all-followers-are-fake-followers/551789/

Brehm, Jack W. (1993), *Control Motivation and Social Cognition*. New York, NY: Springer.

Brown, Duncan and Nick Hayes (2008), *Influencer Marketing: Who Really Influences Your Customers*. Oxford, UK: Butterworth-Heinemann.

Childers, Terry L, (1986), “Assessment of the Psychometric Properties of an Opinion Leadership Scale,” *Journal of Marketing Research*, 23 (2), 184-188.

Deci Edward L. and Richard M. Ryan (1985), *Intrinsic Motivation and Self-Determination in Human Behavior.* New York: Plenum.

Deci Edward L. and Richard M. Ryan (2000), “Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being,” *American Psychologist*, 55 (1), 68-78.

Deci, Edward L. and Richard M. Ryan (2001), “Questionnaires: Basic Psychological Needs Scales,” Retrieved from the World Wide Web: http://selfdeterminationtheory.org/ questionnaires on July 30, 2019.

Dworkin, Gerald (1988), *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press.

Engel, James F., David T. Kollat, and Roger D. Blackwell (1968), *Consumer Behavior*. New York, NY: Holt, Rinehart and Winston, Inc., 391-3.

Flynn, Leisa Reinecke, Ronald E. Goldsmith, and Jacqueline K. Eastman (1996), "Opinion Leaders and Opinion Seekers: Two New Measurement Scales," *Journal of the Academy of Marketing Science*, 24 (2), 137.

Friedman, Vanessa (2017), "Influencers Run into Consequences," *Fashion & Style, The New York Times*, May 10, page D2, New York Edition. Published online at https://www.nytimes.com/2017/05/10/fashion/kendall-jenner-fyre-festival-pepsi-bella-hadid-influencers.html

Goldsmith, Ronald E., and Charles F. Hofacker. "Measuring Consumer Innovativeness," *Journal of the Academy of Marketing Science*, 19 (3), 209-221.

Grewal, Lauren, and Andrew T. Stephen (2019), "In Mobile We Trust: The Effects of Mobile Versus Nonmobile Reviews on Consumer Purchase Intentions," *Journal of Marketing Research*: 0022243719834514.

Grewal, Rajdeep, Raj Mehta, and Frank R. Kardes (2000), "The Role of the Social-Identity Function of Attitudes in Consumer Innovativeness and Opinion Leadership," *Journal of Economic Psychology*, 21 (3), 233-252.

Hayes, Andrew F. (2013), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: Guilford Press.

Heine, Christopher (2016), "Selena Gomez's Social Media Posts Are Evidently Worth $550,000 Apiece," *Adweek, Digital*. Accessed 9-20-17 at http://www.adweek.com/digital/selena-gomezs-social-media-posts-are-evidently-worth-550000-apiece-172552/

Hinz, Oliver, Bernd Skiera, Christian Barrot, and Jan U. Becker (2011), "Seeding Strategies for Viral Marketing: An Empirical Comparison," *Journal of Marketing*, 75 (6), 55–71.

Iyengar, Raghuram, Christophe Van den Bulte, and Thomas W. Valente (2011), "Opinion Leadership and Social Contagion in New Product Diffusion," *Marketing Science*, 30 (2), 195-212.

Jaccard, James, Choi C. Wan, and Robert Turrisi (1990), "The Detection and Interpretation of Interaction Effects Between Continuous Variables in Multiple Regression," *Multivariate Behavioral Research*, 25 (4), 467-78.

Jacoby, Jacob (1974), "The Construct Validity of Opinion Leadership," *The Public Opinion Quarterly*, 38 (1), 81-89.

Jang, Jin Yea, Kyungsik Han, and Dongwon Lee (2015), "No Reciprocity in Liking Photos: Analyzing Like Activities in Instagram," In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, 273-282.

Jetten, Jolanda, Matthew J. Hornsey, and Inmaculada Adarves-Yorno (2006), "When Group Members Admit to Being Conformist: The Role of Relative Intragroup Status in Conformity Self-Reports," *Personality and Social Psychology Bulletin*, 32 (2), 162–73.

Katz, Elihu and Paul F. Lazarsfeld (1955), *Personal Influence*. Glencoe, Ill: The FreePress.

Keller, Edward and Jonathan Berry (2003), *The Influentials: One American in Ten Tells the Other Nine How to Vote, Where to Eat, and What to Buy*. Simon and Schuster.

King, Charles W. and John O. Summers (1970), "Overlap of Opinion Leadership Across Consumer Product Categories," *Journal of Marketing Research*, 7 (1), 43-50.

Kopller, Karlfritz (1984), *Opinion Leaders.* Hamburg: Heinrich Bauer Verlag.

Li, Xueni, Kimmy Wa Chan, and Sara Kim (2018), "Service with Emoticons: How Customers Interpret Employee Use of Emoticons in Online Service Encounters," *Journal of Consumer Research*, 45 (5), 973-987.

Libai, Barak, Eitan Muller and Renana Peres (2013), "Decomposing the Value of Word-of-Mouth Seeding Programs: Acceleration Versus Expansion," *Journal of Marketing Research*, 50 (2), 161–176.

Main, Sami (2017), "Micro-Influencers Are More Effective with Marketing Campaigns than Highly Popular Accounts," *Adweek*, published March 30. Accessed online on November 9, 2017 at http://www.adweek.com/digital/micro-influencers-are-more-effective-with-marketing-campaigns-than-highly-popular-accounts/

Marcus, Alan S. and Raymond A. Bauer (1964), "Yes: There Are Generalized Opinion Leaders," *The Public Opinion Quarterly*, 28 (4), 628-632.

Markus, Hazel Rose and Barry Schwartz (2010), "Does Choice Mean Freedom and Well-Being?" *Journal of Consumer Research*, 37 (2), 344-55.

Merton Robert K. (1968), "Patterns of Influence: Local and Cosmopolitan Influentials," in *Social Theory and Social Structure*, ed. Robert K. Merton, New York: Free Press, 441–74.

Momtaz, Niyoosha Jafari, Abdollah Aghaie, and Somayeh Alizadeh (2011), "Identifying Opinion Leaders for Marketing by Analyzing Online Social Networks," *International Journal of Virtual Communities and Social Networking*, 3 (1), 43-59.

Muller, Eitan and Renana Peres (2019), “The Effect of Social Networks Structure on Innovation Performance: A Review and Directions for Research,” *International Journal of Research in Marketing,* 36 (1). 3-19.

Myers, James H. and Thomas S. Robertson (1972), “Dimensions of Opinion Leadership,” *Journal of Marketing Research*, 9 (1), 41-46.

Nejad, Mohammad G., Daniel L. Sherrell, and Emin Babkus (2014), “Influential and Influence Mechanisms in New Product Diffusion: An Integrative Review,” *Journal of Marketing Theory and Practice*, 22 (2), 185-208.

Ng, Sik Hung (1980), *The Social Psychology of Power (Vol. 21)*. Academic Press.

Ohanian, Roobina (1990), “Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness,” *Journal of Advertising*, 19 (3), 39-52.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn (2015), *The Development and Psychometric Properties of LIWC2015*. Austin, TX: University of Texas at Austin.

Probst, Florian, Laura Grosswiele, and Regina Pfleger (2013), “Who Will Lead and Who Will Follow: Identifying Influential Users in Online Social Networks: A Critical Review and Future Research Directions,” *Business & Information Systems Engineering*, 5 (3), 179–193.

Ranganathan, C. (2012). “The Role of Extrinsic Cues in Consumer Decision Process in Online Shopping Environments,” *Journal of Electronic Commerce in Organizations*, 10 (1), 52-71.

Rogers, Everett M. (1962), *Diffusion of Innovations,* Free Press of Glencoe.

Rogers, Everett M., and David G. Cartano (1962), “Methods of Measuring Opinion Leadership,” *Public Opinion Quarterly*, 26 (3), 435-441.

Rogers, Everett M., and F. Floyd Shoemaker (1971), *Communication of Innovations: A Cross-Cultural Approach*, Free Press.

Ruvio, Ayalla and Aviv Shoham (2007), “Innovativeness, Exploratory Behavior, Market Mavenship, and Opinion Leadership: An Empirical Examination in the Asian Context,” *Psychology and Marketing*, 24 (8), 703-722.

Ryan, Richard M. and James P. Connell (1989), “Perceived Locus of Causality and Internalization: Examining Reasons for Acting in Two Domains,” *Attitudes and Social Cognition*, 57 (5) 749-761.

Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts (2006), “Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market,” *Science,* 311 (5762), 854-856.

Schwartz, Shalom H. (1992), “Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries,” *Advances in Experimental Social Psychology*, 25, 1-65.

Shah, Anuj K. and Daniel M. Oppenheimer (2008), "Heuristics Made Easy: An Effort-Reduction Framework," *Psychological Bulletin*, 134 (2), 207.

Simpson, Jack (2016) "Eight Influencer Marketing Stats for Fashion and Beauty Brands," *Econsultancy, Xeim: Excellency in Marketing*. Accessed 7-9-19 at https://econsultancy.com/eight-influencer-marketing-stats-for-fashion-beauty-brands/

Van den Bulte, Christophe (2010), "Opportunities and Challenges in Studying Consumer Networks," in *The Connected Customer*, ed. S. Wuyts, M.G. Dekimpe, E. Gijsbrechts, and R. Pieters, London: Routledge, 7-35.

Warren, Caleb and Margaret C. Campbell (2014), "What Makes Things Cool? How Autonomy Influences Perceptions of Coolness," *Journal of Consumer Research*, 41 (2), 543–63.

Warriner, Amy Beth, Victor Kuperman, and Marc Brysbaert (2013), "Norms of Valence, Arousal, and Dominance for 13,915 English Lemmas," *Behavior Research Methods*, 45 (4), 1191-207.

Watts, Duncan J. and Peter Sheridan Dodds (2007), "Influentials, Networks, and Public Opinion Formation," *Journal of Consumer Research*, 34 (4), 441-58.

Weimann, Gabriel, Deon Tustin, Daan van Vuuren, and Pierre Joubert (2007), "Looking for Opinion Leaders: Traditional vs. Modern Measures in Traditional Societies," *Journal of Public Opinion Research,* 19 (2), 173-190.

TABLE 1

SUMMARY OF STUDIES

Study 2 (N = 276, 49.6% female, M_{age} = 20.5)

	High Followers			**Low Followers**		
	High Following (N = 70)	**Low Following** (N = 65)		**High Following** (N = 70)	**Low Following** (N = 71)	
Influence	4.80 [4.40, 5.20]	6.09 [5.67, 6.51]	$p < .001$	2.16 [1.71, 2.61]	2.07 [1.69, 2.45]	$p = .765$
Opinion Leadership	4.71 [4.43, 5.00]	5.46[5.17, 5.74]	$p = .003$	3.05 [2.63, 3.46]	3.08 [2.71, 3.45]	$p = .894$

Main Findings: *following* fewer others increases perceptions of an individual's influence, conditional on the individual having a substantial number of followers.

Study 3 (N = 315, 50.8% female, M_{age} = 20.4)

	High Following (N = 159)	**Low Following** (N = 156)	
Autonomy	5.48 [5.23, 5.73]	5.90 [5.68, 5.13]	$p = .015$
Influence	4.13 [3.85, 4.42]	5.33 [5.04, 5.61]	$p < .001$
Likes	3.14 [2.81, 3.47]	2.44 [2.16, 2.72]	$p = .001$
Retweets	2.40 [2.11, 2.69]	1.89 [1.67, 2.11]	$p = .006$

Main Findings: perceived autonomy mediates the relationship between *following* and perceptions of influence; perceptions of influence, in turn, affect engagement intentions.

Study 4 (N = 703, 47.8% female, M_{age} = 20.5)

	Control			**Influence Info**		
	High Following (N = 176)	**Low Following** (N = 176)		**High Following** (N = 175)	**Low Following** (N = 176)	
Likes	4.30 [3.95, 4.65]	2.89 [3.02, 3.61]	$p < .001$	5.32 [4.95, 5.70]	5.21 [4.85, 5.58]	$p = .651$
Retweets	3.61 [3.28, 3.95]	2.71 [2.44, 2.97]	$p < .001$	4.26 [3.86, 4.65]	4.40 [4.02, 4.78]	$p = .564$

Main Findings: *following* drives engagement intentions only when it serves as a cue of influence.

Study 5 (N = 256, 47.7% female, M_{age} = 20.0)

	High Following (N = 128)	**Low Following** (N = 128)	
Influence	4.53 [4.23, 4.84]	5.77 [5.46, 6.06]	$p < .001$
Click	40.6 %	53.9%	$p = .033$

Main Findings: *following* drives actual engagement by affecting perceptions of influence.

TABLE 2

SUMMARY STATISTICS

Variable	Mean	SD
User Following	1,362.30	4,643.62
User Followers	3,224.21	9,604.13
Tweet Likes	2.80	36.83
Tweet Retweets	1.10	42.79
Tweet Controls		
User Mentions	.21	.62
Hashtags	.55	1.25
URLs	.54	.52
Photos	.15	.35
Videos	.01	.08
Financial Symbols	.00	.13
Length	15.22	6.70
Tone (Positivity)	45.37	37.22
Arousal	1.24	.69
Anger	.98	4.17
Anxiety	.23	1.89
Sadness	.39	2.49
Scrape Timestamp - Tweet Timestamp (mins)	7,658.48	899.67
User Controls		
User Age (months)	51.08	30.23
User Tweets	36,241.92	104,226.96
Total Likes	5,793.32	15,563.68
Is Verified	.02	.15
Bio Length	75.73	53.62
Bio Has URL	.53	.50
Default User Profile	.33	.47
Default User Image	.01	.08

Note: Statistics computed at the tweet level

TABLE 3

THE EFFECT OF FOLLOWING ON LIKES

	(1)	(2)	(3)	(4)
log Following	−.237***	−.238***	−.240***	−.066
	(.028)	(.020)	(.019)	(.070)
log Followers	.634***	.784***	.777***	.916***
	(.015)	(.021)	(.018)	(.069)
log Followers × log Following				−.023*
				(.010)
Tweet User Mentions		.148***	.117***	.115***
		(.014)	(.014)	(.014)
Tweet Hashtags		−.082***	−.094***	−.093***
		(.021)	(.018)	(.018)
Tweet URLs		−.693***	−.794***	−.800***
		(.042)	(.033)	(.032)
Tweet Photos		.662***	.604***	.607***
		(.076)	(.073)	(.073)
Tweet Videos		1.557***	1.532***	1.527***
		(.220)	(.233)	(.233)
Tweet Financial Symbols		−.256**	−.289***	−.284***
		(.091)	(.087)	(.083)
lag Scrape Timestamp - Tweet Timestamp (mins)		−.062	.001	−.003
		(.112)	(.097)	(.097)
log User Age (months)		−.206	−.180	−.189*
		(.126)	(.093)	(.096)
log User Tweets		−.304***	−.317***	−.321***
		(.050)	(.038)	(.036)
log User Total Likes		.157***	.174***	.170***
		(.030)	(.023)	(.024)
User Is Verified		.784***	.789***	.820***
		(.076)	(.075)	(.074)
log User Bio Length		.017	−.014	−.016
		(.029)	(.021)	(.020)
User Bio Has URL		−.353***	−.343***	−.335***
		(.089)	(.071)	(.067)
Default User Profile		.149**	.154***	.149***
		(.047)	(.042)	(.040)
Default User Image		−1.125***	−1.214***	−1.148***
		(.237)	(.223)	(.216)
log Tweet Length			.542***	.541***
			(.072)	(.071)
Tweet Positivity			.014	.015
			(.021)	(.020)
Tweet Arousal			−.005	−.005
			(.011)	(.011)
Tweet Anger			.063	.060
			(.038)	(.036)
Tweet Anxiety			−.008	−.007
			(.005)	(.004)
Tweet Sadness			.076	.073
			(.041)	(.039)
N	439051	439051	439051	439051
Pseudo R^2	.066	.12	.13	.13

Note: The dependent variable is the number of likes of tweet *i* of user *j* at time *t.* Cluster-robust standard errors at the individual user level are shown in parentheses.
Significance levels: * p<.05, ** p<.01, *** p<.001.

TABLE 4

THE EFFECT OF FOLLOWING ON RETWEETS

	(1)	(2)	(3)	(4)
log Following	−.250***	−.214***	−.217***	−.178*
	(.042)	(.030)	(.029)	(.077)
log Followers	.730***	.891***	.880***	.911***
	(.029)	(.034)	(.031)	(.080)
log Followers × log Following				−.005
				(.010)
Tweet User Mentions		.216***	.196***	.195***
		(.022)	(.022)	(.022)
Tweet Hashtags		−.048	−.052	−.052
		(.030)	(.029)	(.028)
Tweet URLs		−.450***	−.531***	−.532***
		(.064)	(.054)	(.053)
Tweet Photos		.752***	.703***	.704***
		(.137)	(.142)	(.142)
Tweet Videos		1.920***	1.912***	1.909***
		(.235)	(.252)	(.252)
Tweet Financial Symbols		−.295	−.284	−.282
		(.171)	(.172)	(.171)
lag Scrape Timestamp - Tweet Timestamp (mins)		.030	.096	.094
		(.221)	(.200)	(.200)
log User Age (months)		−.371**	−.333***	−.335***
		(.119)	(.092)	(.096)
log User Tweets		−.163***	−.177***	−.177***
		(.044)	(.037)	(.036)
log User Total Likes		.081*	.097***	.096**
		(.033)	(.029)	(.030)
User Is Verified		.483***	.473***	.480***
		(.100)	(.099)	(.099)
log User Bio Length		−.012	−.045*	−.046*
		(.031)	(.023)	(.022)
User Bio Has URL		−.523***	−.500***	−.498***
		(.105)	(.093)	(.092)
Default User Profile		.218**	.219**	.218***
		(.074)	(.067)	(.066)
Default User Image		−1.054***	−1.109***	−1.093***
		(.289)	(.285)	(.281)
log Tweet Length			.435***	.435***
			(.080)	(.080)
Tweet Positivity			.026	.026
			(.038)	(.038)
Tweet Arousal			.024	.024
			(.017)	(.017)
Tweet Anger			.082*	.082*
			(.035)	(.034)
Tweet Anxiety			.003	.003
			(.007)	(.007)
Tweet Sadness			.078*	.077*
			(.035)	(.034)
N	439051	439051	439051	439051
Pseudo R^2	.072	.10	.11	.11

Note: The dependent variable is the number of retweets of tweet *i* of user *j* at time *t.* Cluster-robust standard errors at the individual user level are shown in parentheses.
Significance levels: * p<.05, ** p<.01, *** p<.001.

TABLE 5

ROBUSTNESS CHECKS: ALTERNATIVE DATA SAMPLE (DV = LIKES)

	(1) Followers ≤ 10k	(2) Followers ≤ 50k	(3) Followers ≤ 150k	(4) Followers ≤ 200k	(5) All Users and Tweets
log Following	−.233***	−.239***	−.229***	−.230***	−.216***
	(.023)	(.020)	(.020)	(.019)	(.016)
log Followers	.760***	.770***	.772***	.768***	.723***
	(.032)	(.021)	(.018)	(.018)	(.015)
Controls	Yes	Yes	Yes	Yes	Yes
N	406006	433414	441735	443629	1581522
Pseudo R^2	.098	.12	.14	.14	.32

Note: The dependent variable is the number of likes of tweet i of user j at time t. Cluster-robust standard errors at the individual user level are shown in parentheses. In column 1, we report the results for a sample of original tweets written by users with 10,000 followers or fewer; in column 2, we report the results for a sample of original tweets written by users with 50,000 followers or fewer; in column 3, we report the results for a sample of original tweets written by users with 150,000 followers or fewer; in column 4, we report the results for a sample of original tweets written by users with 200,000 followers or fewer; in column 5, we use the full sample, i.e., all users and tweets.
Significance levels: * p<.05, ** p<.01, *** p<.001.

TABLE 6

ROBUSTNESS CHECKS: ALTERNATIVE DATA SAMPLE (DV = RETWEETS)

	(1) Followers ≤ 10k	(2) Followers ≤ 50k	(3) Followers ≤ 150k	(4) Followers ≤ 200k	(5) All Users and Tweets
log Following	−.224***	−.215***	−.208***	−.205***	−.170***
	(.036)	(.030)	(.029)	(.029)	(.017)
log Followers	.894***	.870***	.872***	.866***	.576***
	(.046)	(.034)	(.031)	(.031)	(.021)
Controls	Yes	Yes	Yes	Yes	Yes
N	406006	433414	441735	443629	1581522
Pseudo R^2	.075	.093	.11	.12	.083

Note: The dependent variable is the number of retweets of tweet i of user j at time t. Cluster-robust standard errors at the individual user level are shown in parentheses. In column 1, we report the results for a sample of original tweets written by users with 10,000 followers or fewer; in column 2, we report the results for a sample of original tweets written by users with 50,000 followers or fewer; in column 3, we report the results for a sample of original tweets written by users with 150,000 followers or fewer; in column 4, we report the results for a sample of original tweets written by users with 200,000 followers or fewer; in column 5, we use the full sample, i.e., all users and tweets.
Significance levels: * p<.05, ** p<.01, *** p<.001.

Figure 1

Conceptual Model Along with Associated Studies and Context

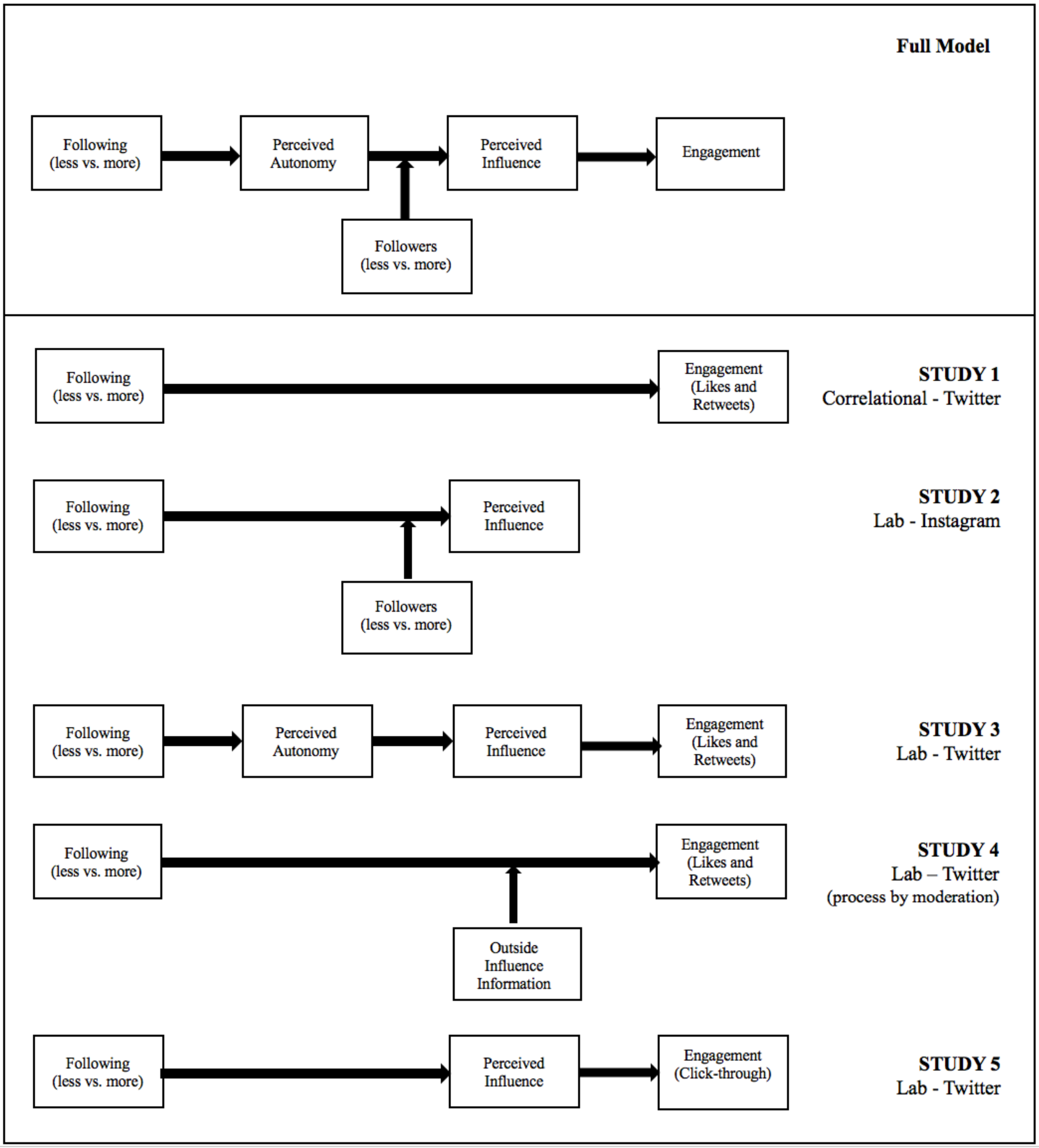


Figure 2

The Influence of Followers on Following (Likes)

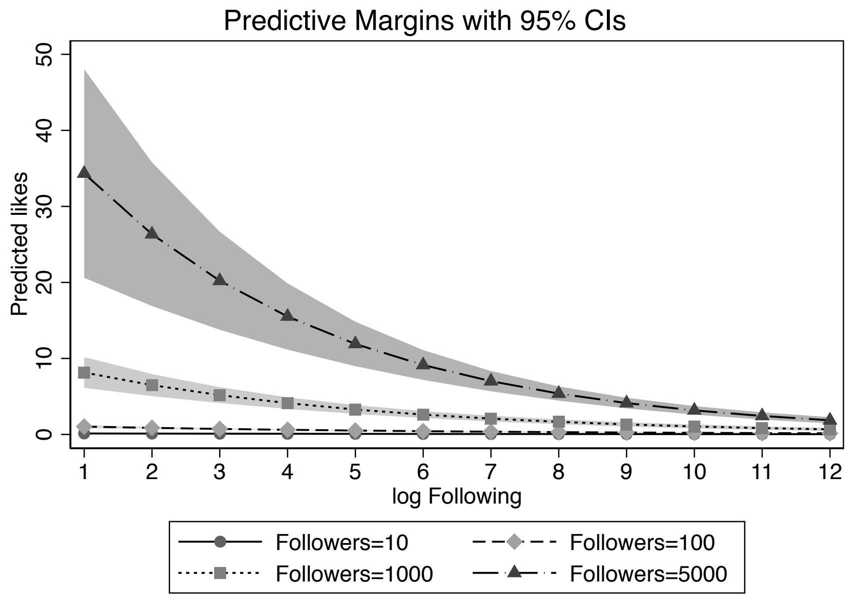


Figure 3

The Influence of Followers on Following (Retweets)

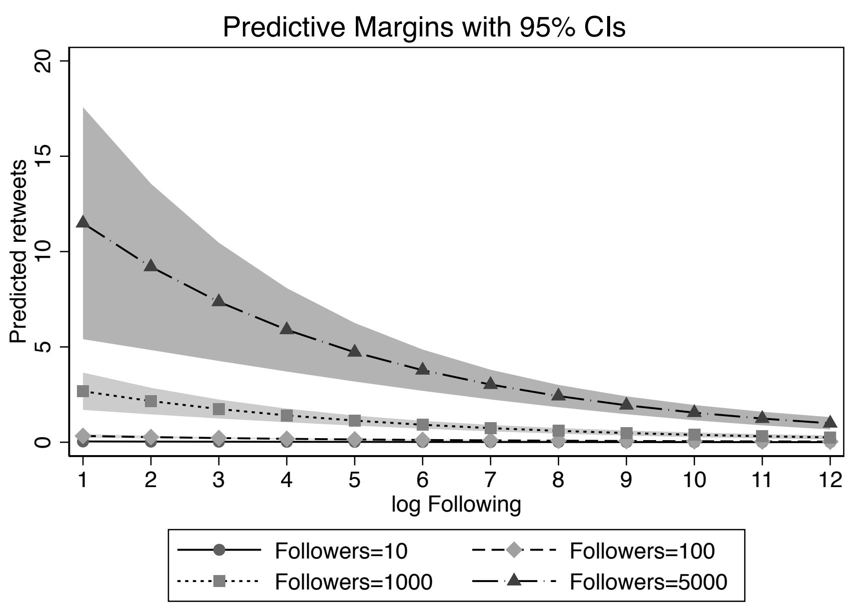


Figure 4

The Moderating Effect of Followers (Influence)

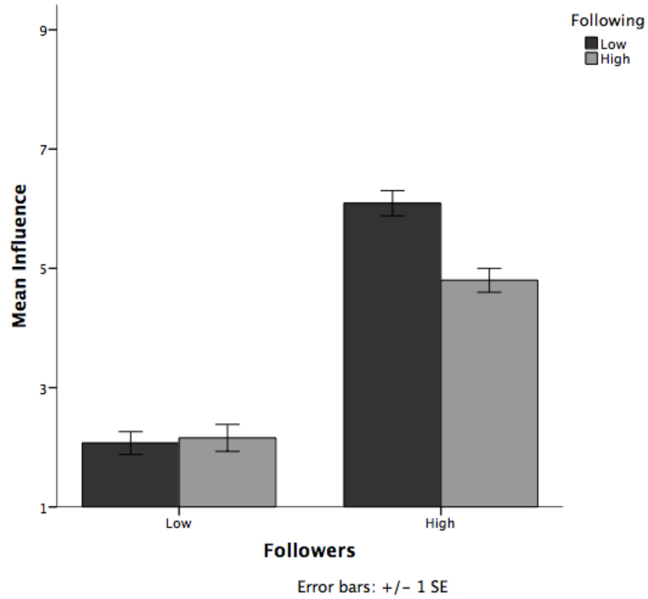
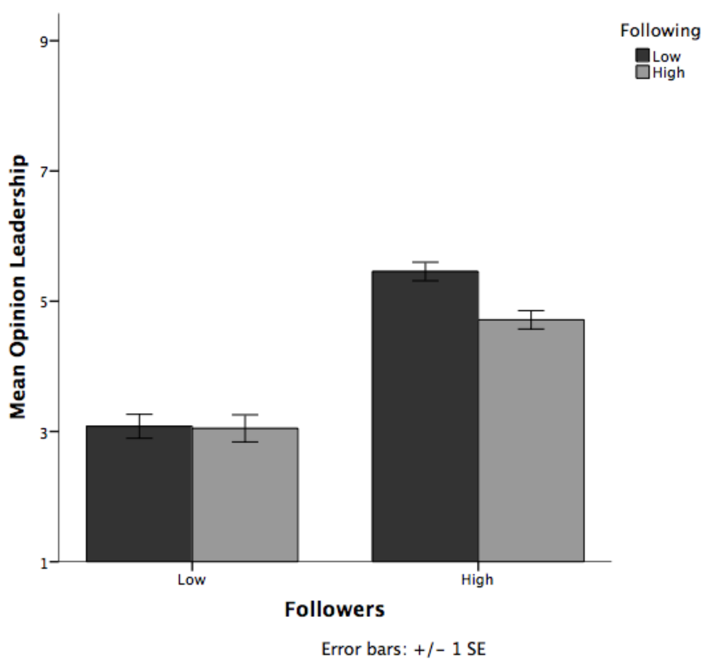


Figure 5

The Moderating Effect of Followers (Opinion Leadership)



WEB APPENDIX

1. INFLUENCER IDENTIFICATION PLATFORMS

2. LITERATURE ON INFLUENTIALS AND SOCIAL MEDIA INFLUENCERS
   a. SUMMARY OF OUR CONTRIBUTION AND RELATED LITERATURE
   b. PAPERS ON SOCIAL MEDIA INFLUENCERS IN MARKETING TOP JOURNALS

3. STUDY 1
   a. CORRELATION MATRIX
   b. ROBUSTNESS CHECKS
      i. OLS REGRESSION MODELS
      ii. ADDITIONAL DATASETS
   c. PERSONAL PRONOUNS USAGE ANALYSIS

4. STUDY 2
   a. STIMULI
   b. OPINION LEADERSHIP SCALE (also used in Study 3)
   c. OTHER MEASURES (also collected for all lab studies)

5. STUDY 3
   a. PRE-TEST STIMULI
   b. PRE-TEST MEASURES (authenticity measure used also in Study 3)
   c. STIMULI (also used in Study 4)

6. STUDY 4 PRE-TEST

7. STUDY 5 STIMULI

8. STUDY TESTING FOLLOWING ACROSS A RANGE OF VALUES

1. INFLUENCER IDENTIFICATION PLATFORMS

Brand24
Buzzstream
Buzzsumo
Collabor 8
Command For Instagram
Deep Social
Famebit
Fohr
Followerwonk
Grin
Grouphigh
Hypetap
Hypr
Ifluenz
Influanza
Influence.Co
Influencerdb
Insightpool
Inzpire.Me
Keyhole
Klear
Lefty
Markerly
Mavrck
Meltwater
Neoreach
Ninja Outreach
Open Influence
Peoplemap
Pixlee
Reech
Revfluence
Scrunch
Speakr
Tapinfluence
Traackr
Tribe
Upfluence
Webfluential
Zine

2. LITERATURE ON INFLUENTIALS AND SOCIAL MEDIA INFLUENCERS

a. TABLE A1: SUMMARY OF OUR CONTRIBUTION AND RELATED LITERATURE

Our theoretical contribution	Related literature		Examples of related literature	
Identify a new mechanism of social influence online Bridge individual-based and network-based approaches to identify influentials online	Characteristics of influentials (individual-based approaches)		Myers and Roberston (1972)	Opinion leaders are innovators, interested and knowledgeable. Opinion leadership overlaps across topic areas.
			Ruvio and Shoham (2007)	Opinion leaders are innovators.
			Grewal, Mehta and Kardes (2000)	Opinion leaders are innovators, experts and highly involved
	Identifying influentials	Self-designation (individual-based approaches)	Rogers and Cartano (1962)	Opinion leadership scale development.
			King and Summers (1970)	Opinion leadership scale development.
			Childers (1986)	Opinion leadership scale development.
			Flynn, Goldsmith and Eastman (1996)	Opinion leadership scale development.
		Socio-metric techniques (network-based approaches)	Iyengar, Van den Bulte and Valente (2011)	Network centrality scores are strongly associated with social influence.
			Hinz and his colleagues (2011)	Seeding 'hubs' (high network centrality) is most effective seeding strategy.
Identify following as a cue for autonomy and influence	Informative cues and inference-making in digital environments		Ranganathan (2012)	Web interface cues, transaction cues, and vendor image cues are predict online purchase intentions.
			Berger and Barasch (2018)	Type of pictures posted are used as a cue to evaluate other social media users.
			Li, Chan and Kim (2019)	Use of emojis is used as a cue to assess service personnel personality.
			Grewal and Stephen (2019)	Device type used is a cue to assess online review credibility.
Identify novel downstream consequences of autonomy perceptions (perceptions of influence and engagement)	Positive signaling effects of autonomy		Bellezza, Gino, Keinan (2014)	Autonomy perceptions drive perceptions of status and competence.
			Warren and Campbell (2014)	Autonomy perceptions drive perceptions of "coolness."

b. TABLE A2: PAPERS ON SOCIAL MEDIA INFLUENCERS IN MARKETING TOP JOURNALS (*Journal of Marketing Research*, *Journal of Marketing*, *Journal of Consumer Research*, *Marketing Science*)

AUTHORS	FOCUS	CONCLUSIONS
Ansari, A., F. Stahl, M. Heitmann, and L. Bremer (2018), "Building a Social Network for Success," *Journal of Marketing Research*, 55(3), 321-338.	Model how musical artists can enhance their social networking presence and stimulate relationships between fans to achieve long-term benefits in terms of music plays on a European online social networking site.	Artists can influence the structure of their ego network (a central actor, the friends of the actor, and all of their friends) and drive song plays over the long run by actively sending friend requests or comments to fans.
Goldenberg, J., G. Oestreicher-Singer, and S. Reichman (2012), "The Quest for Content: How User-Generated Links Can Facilitate Online Exploration," *Journal of Marketing Research*, 49(4), 452-468.	Investigate the role of dual network structure and user generated links in facilitating content exploration.	User-generated links improve exploration efficiency by leading consumers to find better content more quickly and improve exploration effectiveness by increasing overall consumer satisfaction.
Gong, S., J. Zhang, P. Zhao and X. Jiang (2017), "Tweeting as a Marketing Tool: A Field Experiment in the TV Industry," *Journal of Marketing Research*, 54(6) 833-850.	Explore whether and how tweeting affects product demand in the domain of TV shows.	Company tweets increase viewership and influential tweets (from a Weibo user who has many followers, tweets actively, and is retweeted actively by followers) increase viewing and company followers.
Hinz, O., B. Skiera, C. Barrot, and J. U. Becker (2011), "Seeding Strategies for Viral Marketing: An Empirical Comparison," *Journal of Marketing*, 75(6), 55-71.	Compare four seeding strategies: those targeting "hubs," people with a high number of connections; "fringes," people poorly connected; "bridges," people who connect two otherwise unconnected parts of the network; and random people.	The best strategies target the message to hubs (high-degree seeding) or bridges (high-betweenness seeding).
Katona, Z., P. Pal Zubcsek, and M. Savary (2011), "Network Effects and Personal Influences: The Diffusion of an Online Social Network," *Journal of Marketing Research*, 48(3), 425-443.	Uncover the effects of differences in individuals' connection patterns within a social network on the diffusion process (network adoption). The authors look at the structure of connection patterns, individual characteristics of prior adopters and characteristics of potential adopters.	The number and interconnectedness of already adopted friends has a positive effect on the probability of an individual's adoption. People with many friends have a lower average influence than those with fewer friends. Certain demographic variables also play a role.
Kumar V., V. Bhaskaran, R. Mirchandani and M. Shah (2013), "Practice Prize Winner—Creating a Measurable Social Media Marketing Strategy: Increasing the Value and ROI of Intangibles and Tangibles for Hokey Pokey," *Marketing Science*, 32(2), 191-363.	Creation of a unique metric to measure the net influence wielded by a user in a social network, customer influence effect (CIE), and predicting the user's ability to generate the spread of viral information.	Development and validation of CIE (an extension of extend Hubbell's influence measure based on tracking the spread of a message) and CIV (calculated by iteratively summing the CLV of all the people influenced by the Individual) metrics.
Lambrecht, A., C. Tucker and C. Wiertz (2018), "Advertising to Early Trend Propagators," *Marketing Science*, 37(2), 177-331.	Examine the effectiveness of promoted tweets (i.e., advertising messages sent to Twitter users) in engaging early trend propagators (i.e., Twitter users who post on a trend the day it emerges).	Early trend propagators are significantly less likely to respond positively to a targeted ad than users who post on the trend during the following days.
Lanz, A., J. Goldenberg, D. Shapira, and F. Stahl (2019), "Climb or Jump: Status-Based Seeding in User-Generated Content Networks," *Journal of Marketing Research*, 56(3), 361-378.	Investigate what measures help a creator of content to build and increase his or her follower base on a user-generated content network.	High-status seeding targets (high in degree) are associated with very low responsiveness and the return is higher with low status (i.e., ordinary individuals) targets. A creator who has secured followers reallocates outbound activities from low- to high-status seeding targets.
Park, E., R. Rishika, R. Janakiraman, M. B. Houston, and B. Yoo (2018), "Social Dollars in Online Communities: The Effect of Product, User, and Network Characteristics," *Journal of Marketing*, 82(1), 93-114.	Examine the impact of gamers' social networks on their purchase behavior.	Social interactions between users of online communities influence repeat purchase behavior of users. Gamers who have friends who made purchases spend more on purchases themselves. Gamer experience is an important moderator.
Trusov, M., A. V. Bodapati, and R. E. Bucklin (2010), "Determining Influential Users in Internet Social Networks," *Journal of Marketing Research*, 47(4), 643-658.	Develop an approach to determine which users have significant effects on the activities of others based on daily log-in activities on a major social networking site.	Develop a nonstandard Bayesian shrinkage approach to calculating influence scores. The method extracts, with limited data, the strong links from a large overt network that has mostly weak links.

3. STUDY 1

a. CORRELATION MATRIX

TABLE A3

CORRELATION MATRIX

log Following	1.000																						
log Followers	0.657***	1.000																					
Tweet User Mentions	0.104***	0.088***	1.000																				
Tweet Hashtags	0.011***	0.038***	0.044***	1.000																			
Tweet URLs	0.031***	0.067***	0.035***	0.189***	1.000																		
Tweet Photos	0.047***	0.074***	0.049***	0.179***	−0.045***	1.000																	
Tweet Videos	0.005**	0.023***	0.039***	0.027***	−0.059***	−0.032***	1.000																
Tweet Financial Symbols	−0.028***	−0.021***	−0.005**	0.042***	0.014***	0.000	−0.002	1.000															
lag Scrape Timestamp - Tweet Timestamp (mins)	−0.014***	−0.038***	−0.039***	−0.052***	−0.067***	−0.037***	−0.004*	−0.002	1.000														
log User Age (months)	0.310***	0.410***	0.042***	−0.009***	0.030***	−0.053***	−0.006***	−0.028***	−0.043***	1.000													
log User Tweets	0.291***	0.544***	−0.050***	−0.042***	0.007***	−0.058***	−0.020***	−0.007***	0.048***	0.429***	1.000												
log User Total Likes	0.347***	0.306***	0.041***	−0.236***	−0.358***	−0.083***	0.018***	−0.026***	0.046***	0.189***	0.275***	1.000											
User Is Verified	0.080***	0.279***	0.064***	−0.006***	0.023***	0.041***	0.030***	−0.000	−0.036***	0.108***	0.055***	0.069***	1.000										
log User Bio Length	0.174***	0.255***	0.094***	0.166***	0.203***	0.094***	0.005***	0.003	−0.071***	0.095***	0.087***	−0.096***	0.085***	1.000									
User Bio has URL	0.105***	0.249***	0.095***	0.155***	0.238***	0.088***	0.009***	−0.016***	−0.092***	0.181***	0.024***	−0.132***	0.116***	0.355***	1.000								
Default User Profile	−0.211***	−0.304***	−0.033***	−0.053***	−0.051***	−0.013***	0.002	0.020***	0.042***	−0.441***	−0.245***	−0.072***	−0.073***	−0.184***	−0.243***	1.000							
Default User Image	−0.103***	−0.096***	−0.007***	−0.010***	0.033***	−0.006***	−0.006***	−0.001	0.005***	−0.021***	−0.004**	−0.083***	−0.013***	−0.092***	−0.056***	0.092***	1.000						
log Tweet Length	0.031***	0.081***	0.129***	0.222***	0.375***	0.201***	0.017***	0.017***	−0.084***	0.061***	−0.049***	−0.266***	0.055***	0.244***	0.235***	−0.088***	0.028***	1.000					
Tweet Positivity	0.039***	0.019***	0.095***	0.049***	0.033***	0.037***	0.005**	0.002	−0.018***	0.019***	−0.064***	−0.021***	0.014***	0.046***	0.057***	−0.023***	−0.002	0.115***	1.000				
Tweet Arousal	−0.008***	−0.015***	−0.093***	−0.014***	−0.069***	−0.028***	−0.012***	−0.014***	0.005***	−0.003	0.017***	−0.040***	−0.018***	−0.018***	−0.023***	0.009***	0.013***	−0.087***	0.053***	1.000			
Tweet Anger	0.001	−0.017***	−0.044***	−0.070***	−0.131***	−0.064***	−0.008***	−0.005***	0.026***	−0.012***	0.030***	0.098***	−0.016***	−0.077***	−0.085***	0.027***	−0.007***	−0.165***	−0.235***	0.217***	1.000		
Tweet Anxiety	−0.003*	−0.006***	−0.018***	−0.025***	−0.047***	−0.026***	−0.005**	−0.002	0.012***	0.003	0.007***	0.025***	−0.004**	−0.021***	−0.024***	−0.000	−0.002	−0.050***	−0.120***	0.053***	0.006***	1.000	
Tweet Sadness	−0.002	−0.011***	−0.023***	−0.036***	−0.078***	−0.034***	−0.005**	0.001	0.024***	−0.007***	0.009***	0.049***	−0.009***	−0.039***	−0.047***	0.009***	−0.006***	−0.090***	−0.152***	0.039***	0.003	0.040***	1.000

Significance levels: * p<.05, ** p<.01, *** p<.001.

b. ROBUSTNESS CHECKS

i. OLS REGRESSION MODELS

TABLE A4

ROBUSTNESS CHECKS: OLS (DV=LIKES)

	(1)	(2)
log Following	−.093***	−.163***
	(.005)	(.008)
log Followers	.204***	.150***
	(.007)	(.010)
log Followers x log Following		.010***
		(.001)
Controls	Yes	Yes
N	439051	439051
Pseudo R^2	.30	.30

Note: The dependent variable is the logarithm of the number of likes of tweet i of user j at time t. Cluster-robust standard errors at the individual user level are shown in parentheses. Significance levels: * p<.05, ** p<.01, *** p<.001.

TABLE A5

ROBUSTNESS CHECKS: OSL (DV = RETWEETS)

	(1)	(2)
log Following	−.042***	−.118***
	(.004)	(.005)
log Followers	.124***	.065***
	(.004)	(.007)
log Followers x log Following		.011***
		(.001)
Controls	Yes	Yes
N	439051	439051
Pseudo R^2	.18	.19

Note: The dependent variable is the logarithm of the number of retweets of tweet i of user j at time t. Cluster-robust standard errors at the individual user level are shown in parentheses. Significance levels: * p<.05, ** p<.01, *** p<.001.

TABLE A6: ROBUSTNESS CHECK: BIVARIATE NEGATIVE BINOMIAL ESTIMATION

	Univariate		Bivariate
	Likes	Retweets	
			Likes
log Following	−.240***	−.217***	−.210***
	(.019)	(.029)	(.011)
log Followers	.777***	.880***	.658***
	(.018)	(.031)	(.008)
			Retweets
log Following			−.147***
			(.014)
log Followers			.770***
			(.012)
Controls	Yes	Yes	Yes
N	439,051	439,051	439,051
Pseudo R^2	.13	.11	-

Note: In column 1, the dependent variable is the number of likes of tweet *i* of user *j* at time *t*; in column 2, the dependent variable is the number of retweets of tweet *i* of user *j* at time *t*; in column 3, we employ a bivariate model that takes into account the correlation between likes and retweets. Clustered robust standard errors at the individual level are shown in parentheses.
Significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$.

ii. ADDITIONAL DATASETS

TABLE A7

THE EFFECT OF FOLLOWING ON LIKES (TOKYO DATASET)

	(1)	(2)	(3)	(4)
log Following	−.407***	−.459***	−.443***	−.327***
	(.037)	(.035)	(.035)	(.056)
log Followers	.813***	1.090***	1.062***	1.156***
	(.031)	(.035)	(.037)	(.061)
log Followers × log Following				−.017+
				(.009)
Tweet User Mentions		−.000	−.115***	−.120***
		(.036)	(.034)	(.034)
Tweet Hashtags		.030*	−.033**	−.030**
		(.012)	(.012)	(.012)
Tweet URLs		−.163	−.479***	−.483***
		(.120)	(.091)	(.089)
Tweet Photos		1.340***	1.122***	1.113***
		(.061)	(.057)	(.055)
Tweet Videos		2.331***	2.078***	2.061***
		(.261)	(.236)	(.227)
Tweet Financial Symbols		−.206	−.295	−.302
		(.192)	(.198)	(.198)
lag Scrape Timestamp - Tweet Timestamp (mins)		−.215	−.198	−.198
		(.202)	(.179)	(.176)
log User Age (months)		.063**	.037+	.028
		(.022)	(.019)	(.019)
log User Tweets		−.525***	−.496***	−.498***
		(.053)	(.050)	(.048)
log User Total Likes		.194***	.203***	.199***
		(.028)	(.025)	(.026)
User Is Verified		.460	.564*	.648*
		(.288)	(.260)	(.257)
log User Bio Length		.067	.010	.009
		(.044)	(.038)	(.037)
User Bio Has URL		−.219**	−.187*	−.186*
		(.082)	(.074)	(.072)
Default User Profile		.252*	.276**	.268**
		(.101)	(.092)	(.088)
Default User Image		−.649***	−.773***	−.730***
		(.171)	(.157)	(.146)
log Tweet Length			.618***	.620***
			(.032)	(.032)
N	464798	464798	464798	464798
Pseudo R^2	.055	.13	.14	.14

Note: The dependent variable is the number of likes of tweet *i* of user *j* at time *t*. Cluster-robust standard errors at the individual user level are shown in parentheses.
Significance levels: * p<.05, ** p<.01, *** p<.001.

TABLE A8

THE EFFECT OF FOLLOWING ON RETWEETS (TOKYO DATASET)

	(1)	(2)	(3)	(4)
log Following	−.510***	−.469***	−.468***	−.411***
	(.053)	(.046)	(.045)	(.076)
log Followers	.948***	1.126***	1.095***	1.144***
	(.048)	(.048)	(.048)	(.083)
log Followers × log Following				−.009
				(.012)
Tweet User Mentions		.052	−.160***	−.162***
		(.048)	(.044)	(.044)
Tweet Hashtags		.131***	−.012	−.011
		(.024)	(.023)	(.022)
Tweet URLs		.671***	.092	.092
		(.148)	(.116)	(.116)
Tweet Photos		1.633***	1.429***	1.425***
		(.099)	(.088)	(.087)
Tweet Videos		2.953***	2.708***	2.700***
		(.294)	(.243)	(.241)
Tweet Financial Symbols		−.010	−.242	−.243
		(.117)	(.188)	(.185)
lag Scrape Timestamp - Tweet Timestamp (mins)		−.345	−.319	−.320
		(.281)	(.249)	(.248)
log User Age (months)		.145**	.121**	.117**
		(.053)	(.042)	(.042)
log User Tweets		−.399***	−.353***	−.354***
		(.056)	(.054)	(.053)
log User Total Likes		.066	.086**	.085**
		(.035)	(.030)	(.030)
User Is Verified		.211	.251	.294
		(.436)	(.303)	(.295)
log User Bio Length		.057	−.051	−.052
		(.072)	(.063)	(.062)
User Bio Has URL		−.319**	−.243*	−.242*
		(.117)	(.106)	(.105)
Default User Profile		.102	.168	.167
		(.133)	(.119)	(.118)
Default User Image		−.706**	−.931***	−.906***
		(.246)	(.237)	(.228)
log Tweet Length			1.380***	1.380***
			(.059)	(.059)
N	464798	464798	464798	464798
Pseudo R^2	.061	.10	.13	.13

Note: The dependent variable is the number of likes of tweet *i* of user *j* at time *t*. Cluster-robust standard errors at the individual user level are shown in parentheses.
Significance levels: * p<.05, ** p<.01, *** p<.001.

TABLE A9

THE EFFECT OF FOLLOWING ON LIKES (ALL US DATASET)

	(1)	(2)	(3)	(4)
log Following	−.329***	−.219***	−.223***	−.091
	(.023)	(.031)	(.029)	(.047)
log Followers	.724***	.985***	.970***	1.097***
	(.041)	(.052)	(.046)	(.052)
log Followers × log Following				−.021***
				(.006)
Tweet User Mentions		.068***	.019	.019
		(.018)	(.020)	(.020)
Tweet Hashtags		−.074***	−.095***	−.094***
		(.010)	(.010)	(.010)
Tweet URLs		−.665***	−.707***	−.701***
		(.078)	(.071)	(.070)
Tweet Photos		.485***	.446***	.448***
		(.055)	(.060)	(.059)
Tweet Videos		1.949***	2.124***	2.111***
		(.289)	(.347)	(.340)
Tweet Financial Symbols		−.082***	−.107***	−.097***
		(.023)	(.025)	(.023)
lag Scrape Timestamp - Tweet Timestamp (mins)		−.927	−1.092	−1.070
		(1.035)	(1.009)	(1.005)
log User Age (months)		−.081***	−.099***	−.115***
		(.023)	(.021)	(.021)
log User Tweets		−.372***	−.345***	−.346***
		(.042)	(.041)	(.041)
log User Total Likes		.018	.025	.022
		(.033)	(.030)	(.030)
User Is Verified		.094	.061	.107
		(.266)	(.236)	(.236)
log User Bio Length		−.043	−.078**	−.078**
		(.026)	(.025)	(.025)
User Bio Has URL		−.143	−.194*	−.196*
		(.100)	(.089)	(.089)
Default User Profile		.233*	.241*	.242*
		(.101)	(.099)	(.098)
Default User Image		−.443*	−.515**	−.439*
		(.172)	(.186)	(.199)
log Tweet Length			.425***	.427***
			(.066)	(.065)
Tweet Positivity			.146**	.148**
			(.052)	(.052)
Tweet Arousal			−.050*	−.049*
			(.024)	(.024)
Tweet Anger			.023	.023
			(.017)	(.017)
Tweet Anxiety			−.032***	−.032***
			(.008)	(.008)
Tweet Sadness			.023	.023
			(.024)	(.023)
N	297324	297324	297324	297324
Pseudo R^2	.053	.088	.092	.092

Note: The dependent variable is the number of likes of tweet *i* of user *j* at time *t.* Cluster-robust standard errors at the individual user level are shown in parentheses.
Significance levels: * p<.05, ** p<.01, *** p<.001.

TABLE A10

THE EFFECT OF FOLLOWING ON RETWEETS (ALL US DATASET)

	(1)	(2)	(3)	(4)
log Following	−.251***	−.182***	−.183***	−.222**
	(.076)	(.051)	(.043)	(.085)
log Followers	.701***	.993***	.994***	.958***
	(.061)	(.059)	(.051)	(.077)
log Followers × log Following				.006
				(.010)
Tweet User Mentions		.120***	063**	.063**
		(.026)	(.023)	(.023)
Tweet Hashtags		−.009	−.024*	−.025*
		(.012)	(.011)	(.011)
Tweet URLs		−.785***	−.795***	−.796***
		(.103)	(.091)	(.091)
Tweet Photos		.158	.104	.105
		(.113)	(.092)	(.092)
Tweet Videos		1.950***	2.122***	2.126***
		(.383)	(.445)	(.445)
Tweet Financial Symbols		−.039	−.064*	−.066*
		(.024)	(.025)	(.026)
lag Scrape Timestamp - Tweet Timestamp (mins)		−5.242*	−4.651*	−4.639*
		(2.180)	(1.863)	(1.851)
log User Age (months)		−.167***	−.205***	−.199***
		(.046)	(.039)	(.036)
log User Tweets		−.281***	−.251***	−.251***
		(.041)	(.040)	(.040)
log User Total Likes		.010	−.004	−.003
		(.033)	(.030)	(.030)
User Is Verified		.184	.061	.050
		(.295)	(.242)	(.243)
log User Bio Length		−.141*	−.152**	−.152**
		(.062)	(.052)	(.052)
User Bio Has URL		−.101	−.159	−.160
		(.178)	(.158)	(.157)
Default User Profile		.502**	.465**	.464**
		(.180)	(.157)	(.157)
Default User Image		−.431	−.369	−.392
		(.303)	(.343)	(.336)
log Tweet Length			.344***	.343***
			(.080)	(.081)
Tweet Positivity			−.018	−.018
			(.055)	(.055)
Tweet Arousal			−.273***	−.273***
			(.054)	(.054)
Tweet Anger			−.027	−.026
			(.024)	(.024)
Tweet Anxiety			−.072***	−.072***
			(.014)	(.015)
Tweet Sadness			−.049**	−.049**
			(.018)	(.018)
N	297324	297324	297324	297324
Pseudo R^2	.042	.070	.076	.076

Note: The dependent variable is the number of retweets of tweet *i* of user *j* at time *t*. Cluster-robust standard errors at the individual user level are shown in parentheses.
Significance levels: * p<.05, ** p<.01, *** p<.001.

c. PERSONAL PRONOUNS ANALYSIS

TABLE A11
PERSONAL PRONOUNS USAGE

	(1) I	(2) You	(3) She/he	(4) We	(5) They
log Following	−.079*	.055***	.021***	−.002	.009***
	(.035)	(.016)	(.004)	(.005)	(.002)
N	439051	439051	439051	439051	439051
Pseudo R^2	.00037	.00053	.00027	.0000016	.00011

Note: In column 1, the dependent variable is the LIWC variable "I" that measures first person singular pronouns; in column 2, the dependent variable is the LIWC variable "you" that measures second person pronouns (singular and plural); in column 3, the dependent variable is the LIWC variable "she/he" that measures third person singular pronouns; in column 4, the dependent variable is the LIWC variable "they" that measures third person plural pronouns. Cluster-robust standard errors at the individual user level are shown in parentheses.
Significance levels: * p<.05, ** p<.01, *** p<.001.

Description of the analysis: We tested whether *Following* is negatively correlated with the presence of "first person pronouns" in a tweet and positively correlated with "third person pronouns."
We estimated the following regressions:

$$LIWC_{it} = \beta_1 \log Following_{it} + \epsilon_{it},$$

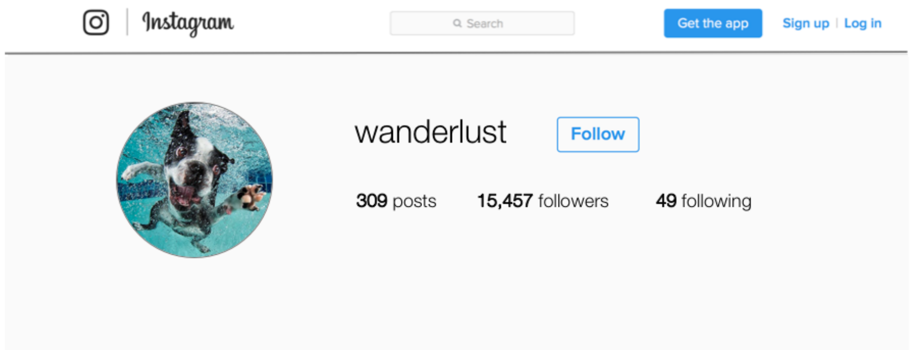
where LIWC takes on the presence (count) of various personal pronouns in a tweet, and $\log Following_{it}$ is the number of *following* of user *i* at time *t.*
Results show that as *Following* goes up, the use of the first person pronouns including "I" goes down, while the use of second and third person pronouns ("You" "She/he" and "They") goes up.

Note that all results reported in the manuscript hold even including personal pronouns as additional controls.

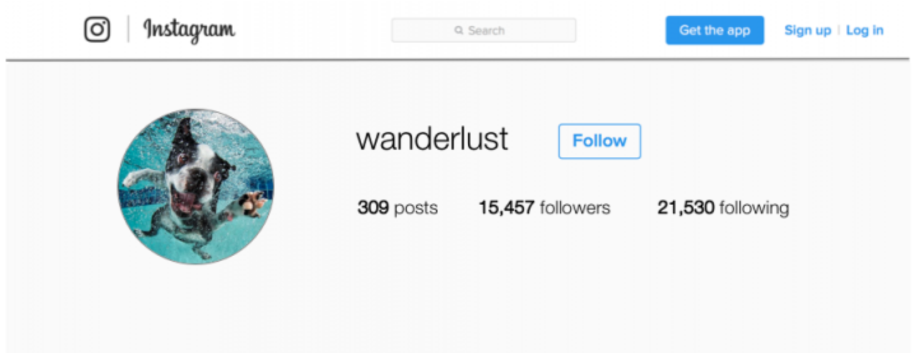
4. STUDY 2

a. STIMULI

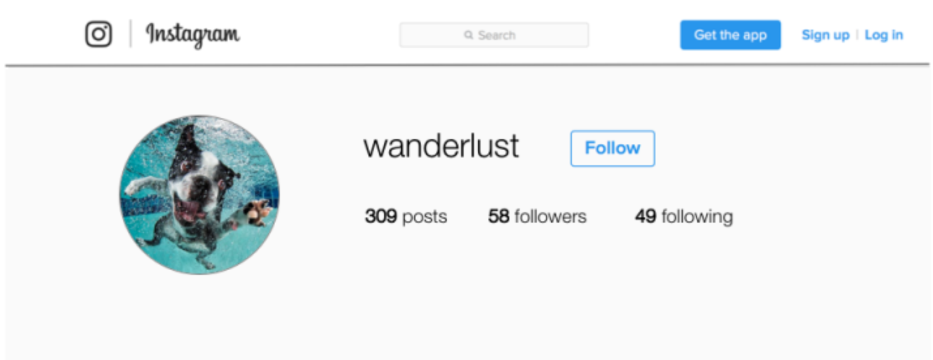
<u>HIGH FOLLOWERS, LOW FOLLOWING</u>



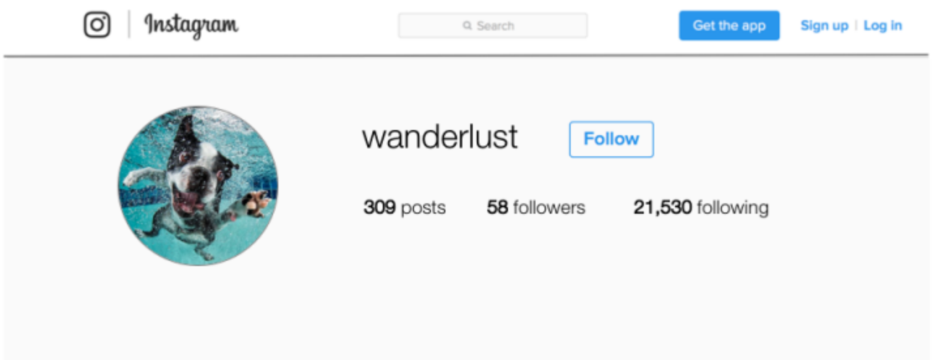
<u>HIGH FOLLOWERS, HIGH FOLLOWING</u>



<u>LOW FOLLOWERS, LOW FOLLOWING</u>



LOW FOLLOWERS, HIGH FOLLOWING



b. OPINION LEADERSHIP SCALE

Please rate your agreement with the following statements (1 = Strongly Disagree, 9 = Strongly Agree):

- Wanderlust's opinion on travel seems not to count with other people (R)
- When they make travel decisions, other people do not turn to Wanderlust for advice (R)
- Other people rarely come to Wanderlust for advice about travel (R)
- People make travel decisions based on what Wanderlust posts
- Wanderlust often persuades other people to buy items
- Wanderlust often influences other people's opinions about travel

c. OTHER MEASURES (also collected for all lab studies)

ATTENTION CHECK

- We would now like to assess how much you can remember of the profile you just evaluated. How many other users does this Instagram user follow? (49; 998; 21,530; I don't recall)

DEMOGRAPHICS

- What is your age?
- What is your gender? (Male, Female)

5. STUDY 3

a. PRE-TEST STIMULI

HIGH FOLLOWING

LOW FOLLOWING

b. PRE-TEST MEASURES (authenticity measure used also in Study 3)

AUTONOMY

Please rate your agreement with the following statements. This Twitter user: (1 = Strongly Disagree, 9 = Strongly Agree):

- Decides what to post without the influence of others
- Doesn't feel pressured with regards to what to post
- Generally expresses his ideas and opinions in his posts
- Has the opportunity to post what he feels like
- Is himself in his posts
- Frequently posts without being influenced by anyone.

EXPERTISE

To what extent do you find this Twitter user to be:

- 1 = Expert, 9 = Not an expert (R)
- 1 = Experienced, 9 = Inexperienced (R)
- 1 = Knowledgeable, 9 = Unknowledgeable (R)
- 1 = Qualified, 9 = Unqualified (R)
- 1 = Skilled, 9 = Unskilled (R)

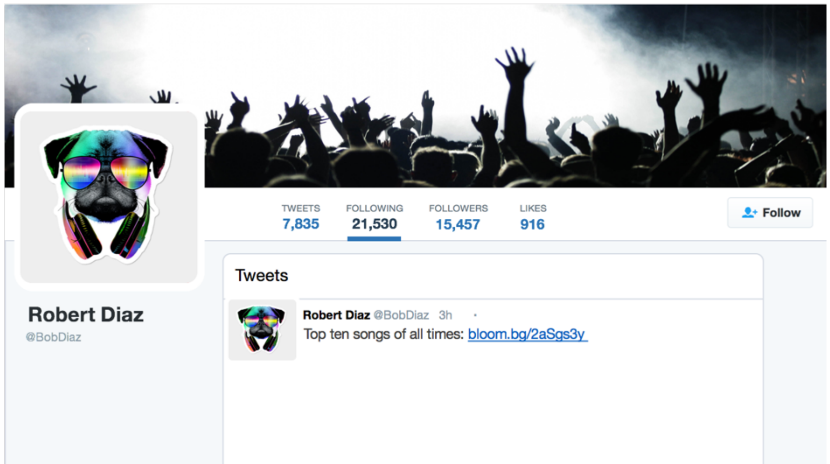
INNOVATIVENESS

Please rate your agreement with the following statements. This Twitter user: (1 = Strongly Disagree, 9 = Strongly Agree):

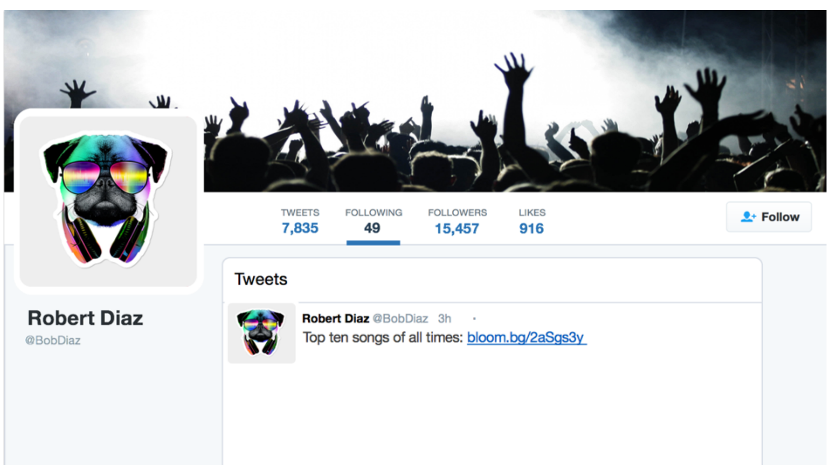
- Has tried fewer restaurants compared to most people. (R)
- In general, is the LAST among his group of friends to know when a restaurant opens in town. (R)
- In general, is the FIRST among his group of friends to try a new restaurant once it opens.
- If he heard of a new food, he would be interested in trying it.
- Will try a new food even if he hasn't heard anything about it yet.
- Knows about new foods before other people do.

c. STIMULI (also used in Study 4)

HIGH FOLLOWING



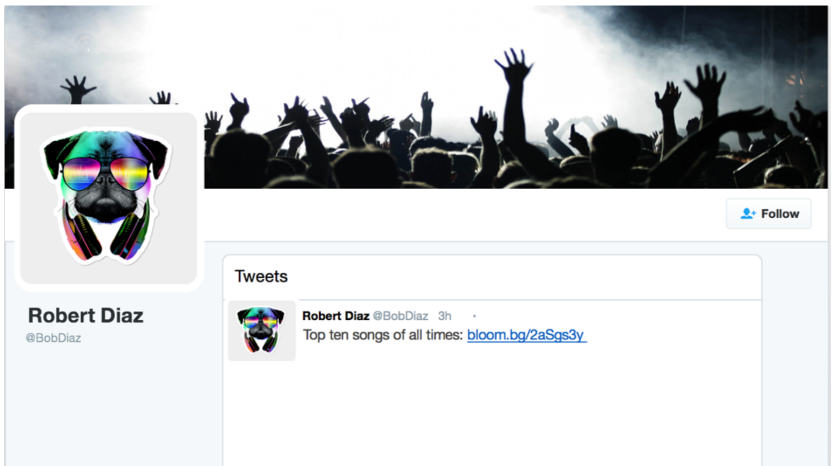
LOW FOLLOWING



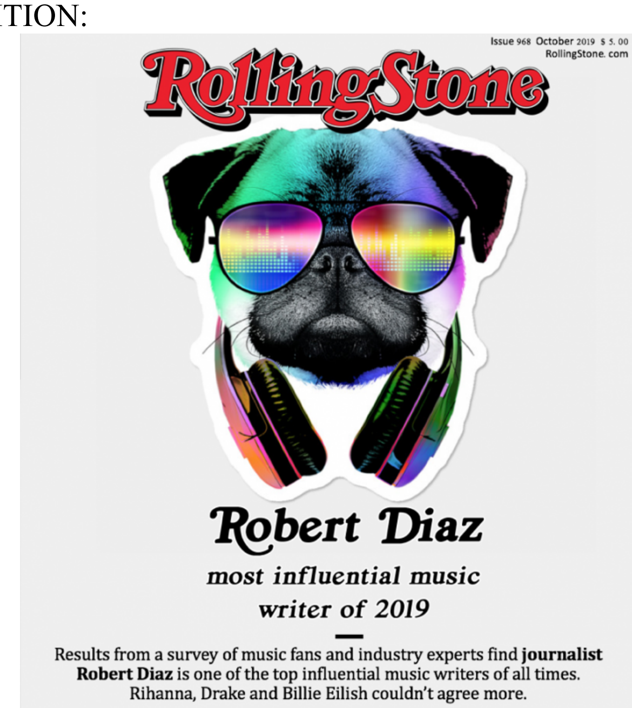
6. STUDY 4 PRE-TEST

N = 203 (mTurk Twitter users)

Stimuli
ALL:



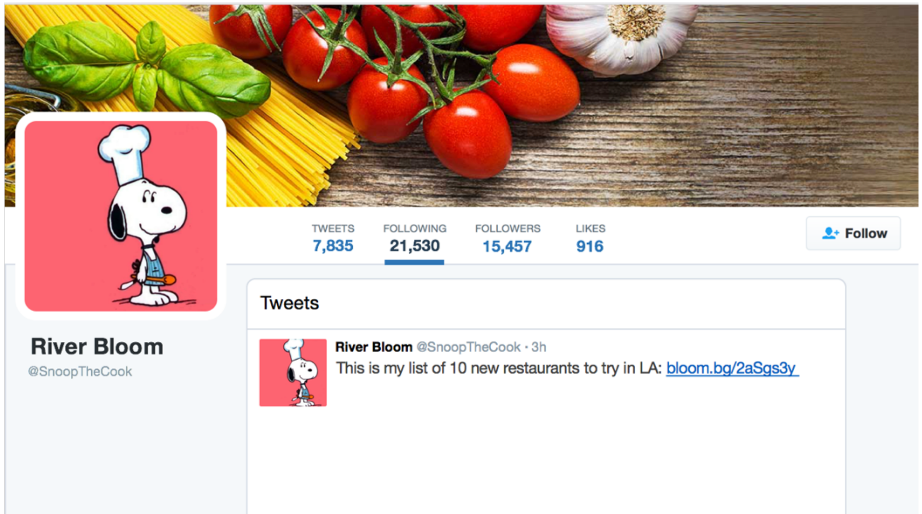
ONLY INFO CONDITION:



Perceived Influence: $M_{Info} = 7.38$, $SD = 1.81$ vs $M_{NoInfo} = 3.85$, $SD = 2.27$, $F(1, 201) = 151.27$, $p < .001$

7. STUDY 5 STIMULI

HIGH FOLLOWING



LOW FOLLOWING

The landing page of the link contains information that could reveal the city in which the study was run. It will be therefore be made available in this web appendix at the end of the review process.

8. STUDY TESTING FOLLOWING ACROSS A RANGE OF VALUES

This study was intended as an exploratory test of the number of users someone is *following* on social media and its impact on the perceived influence of that user.

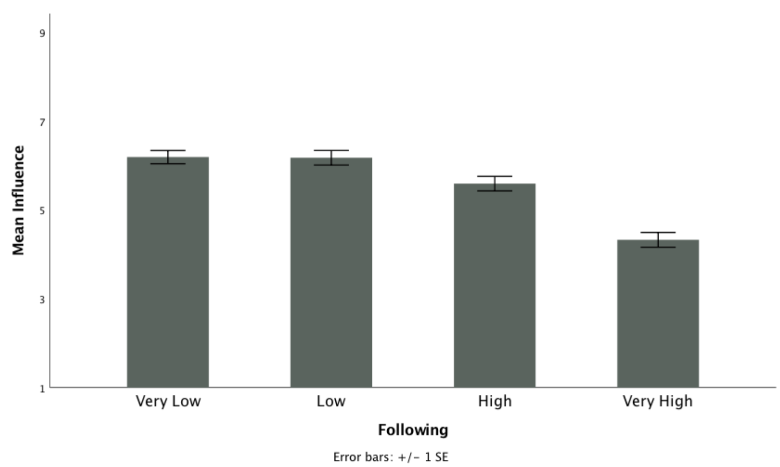
Method

Respondents were 407 undergraduate students from a major university who participated for partial course credit. Respondents' task was to evaluate a Twitter user based on the individual's profile page. The stimuli replicated the features of a real Twitter page, including a fictional user name and picture as well as a predetermined number of *followers* (31,647), tweets (7,835), and total *likes* given (916). The number of *followers* aligns with the micro-influencers described in the opening example while the number of *tweets* and *likes* reflects the averages in the dataset of real Twitter users employed in Study 1. This information remains constant across conditions.

The study employed a single factor, between-subjects design. We varied the number of Twitter users the individual was *Following* based on the real Twitter user data from Study 1, with endpoints at the extremes of the distribution and chosen numbers an order of magnitude apart. The levels include *91* (*Very Low*), *901* (*Low*), *9,001* (*High*), and *90,001* (*Very High*). To assess how the level of *Following* affected perceived influence, respondents were asked "To what extent do you think this user is influential on Twitter?" (1 = Not at All Influential, 9 = Very Influential).

Results

A between-subject ANOVA with *Influence* as the dependent variable reveals a significant main effect of *Following* ($F(3,403) = 29.29$, $p < .001$). Average scores for each condition are displayed in the figure below.



Planned contrasts reveal that the Twitter user in the *Very Low* ($M_{VeryLow} = 6.19$, 95% CI = [5.89, 6.49]) condition was perceived as significantly more influential compared to the user in both the *High* ($M_{High} = 5.59$, 95% CI = [5.27, 5.92], $F(1,403) = 6.86$, $p = .009$) and *Very High* ($M_{VeryHigh} = 4.43$, 95% CI = [4.00, 4.66], $F(1,403) = 66.00$, $p < .001$) conditions but not to the user in the *Low* condition ($M_{Low} = 6.17$, 95% CI = [5.85, 6.50], $F(1,403) = .01$, $p = .947$). The Twitter user in the *Low* condition was perceived as significantly more influential compared to the user in both the *High* ($F(1,403) = 6.61$, $p = .010$) and *Very High* condition ($F(1,403) = 65.89$, $p < .001$). The Twitter user in the *High* condition was perceived as significantly more influential compared to the user in the *Very High* condition ($F(1,403) = 30.89$, $p < .001$).

Discussion

The results affirm the basic premise that the fewer others a social media user is *following*, the more influential s/he is perceived to be, ceteris paribus. We deliberately chose numbers that reflected reality and that spanned the range of possibilities in the real world. Interestingly, we observe no difference between someone following a mere 91 others and someone following 901 others, evidence of a diminishing effect of reducing the number of users someone is *Following*. Detecting differences when following is quite low may require a more sensitive test.