This work investigates the influence of self-presentation goals on reviewers’ ratings in consumer online communities (i.e., crowd-sourced review sites such as Yelp.com). The authors contend that members of these communities are concerned about being perceived as knowledgeable by others, which influences the valence of the ratings and reviews they provide. The influence of self-presentation concerns evolves dynamically over time, leading reviewers’ ratings to become increasingly more negative. The change observed occurs because reviewers adopt different self-presentation tactics as the number of reviews they complete increases. While relatively new to a particular community, reviewers try to present themselves as knowledgeable by showing they make good product choices, which has a positive influence on ratings, controlling for other factors. After accumulating a discernible record, reviewers become relatively more concerned with demonstrating critical skills and a capacity to be discriminating, resulting in more negative ratings. The impact of self-presentation concerns and the concomitant self-presentation tactics on consumer online ratings is substantiated by a combination of laboratory experiments as well as the analysis of real-world data.

**KEYWORDS:** Impression Management, Self-Presentation, Word-of-Mouth, Consumer Ratings, Reviews.
It is a truism that the Internet allows individuals to freely express opinions uninhibited by concerns about the impressions they might make on others. Ergo, a common assumption is that an opinion regarding a product or service found in an online rating forum (e.g., Angieslist.com, Citysearch.com, CNET.com, Epinions.com, TripAdvisor.com, Yahoo Local, and Yelp.com) is unadulterated by the reviewer’s concerns about what readers might think of him or her and is therefore more reliable. The belief that opinions expressed online are unfiltered is supported by academic research. For example, Spears and Lea (1994, p. 430) highlight how academics and practitioners alike believe computer-mediated communication on the World Wide Web allows “one to express one’s true being, or authentic self, unfettered by concerns of self-presentation” (see also Baumeister 1982; Kiesler and Sproull 1992; Sussman and Sproull 1999). This research challenges this assumption directly.

One of the primary reasons reviewers are assumed to be more outspoken with their opinions online (as opposed to offline) is because they are protected by the anonymity granted by the web. Anonymity is believed to attenuate the anxiety associated with being evaluated by others, allowing an individual to be his or her true self and say what he or she truly believes (Dubrovsky, Kiesler, and Sethna 1991). Yet the most popular consumer opinion platforms have moved away from the notion of “strict” anonymity and often encourage social interaction while compelling users to build and maintain a reputation within their communities. Consider Yelp (www.yelp.com) where each reviewer (aka “Yelper”) has a public profile page on which they could, if they so desired, use their real name and an actual photograph of themselves. The profile summarizes their rating record, including designations of any reviews seen as especially “useful,” “funny,” and/or “cool” by visitors to the site. Further, Yelpers can “befriend” and “follow” other Yelpers as well as compliment them on their profile page for everyone to see. The “best” Yelpers are rewarded publicly by Yelp with status by being designated an “Elite Member.” Likewise,
reviewers on TripAdvisor (www.tripadvisor.com) have profile pages featuring their past ratings as well as displaying various “badges” they have been awarded by TripAdvisor for their contribution to the site, as a way of signaling their knowledge and expertise.

The presence of these types of reputation systems adds a social dimension to online reviewing that has been credited with stimulating participation by promoting accountability (Dellarocas 2003; Resnick et al. 2000). According to work by Wang (2010), the proportion of active contributors at Yelp in 2010 was more than 10 times the proportion of Citysearch (www.citysearch.com), the former leading reviewing platform in the U.S., which was known for guaranteeing full anonymity to its users. It therefore appears reviewers are not at all reluctant to exchange full anonymity for reputation-building activities online and any claim that self-presentation concerns are not a factor might be unwarranted. The popularity of reputation systems such as TripAdvisor’s badges implies reviewers care about the impressions they make on other visitors to the site as the only way of acquiring various badges is through the positive feedback of fellow users. We propose that the absence of absolute anonymity on opinion sites such as Yelp and TripAdvisor—caused largely by the inclusion of reputation systems—results in reviewers who are concerned with what others think about them.

This research is the first to document the influence of self-presentation concerns—and the desire to appear knowledgeable (i.e., competent, skilled, and intelligent) in particular—on ratings and reviews published on quasi-anonymous user-generated review sites. We show these concerns have meaningful downstream consequences, significantly influencing the valence of ratings dynamically over time. The causal mechanism proposed and validated is as follows. Reviewers are more positive early on because, while new to a community, they are more concerned with being seen as having made good choices as a way to display their knowledgeability as reviewers. As the number of reviews in their rating history grows, reviewers become increasingly concerned
with being seen as possessing the *critical skills* necessary to discriminate between experiences, and thus become somewhat more negative later on. The result is a meaningful negative trend in ratings corresponding to the number of past ratings in a reviewer’s profile, evident even while controlling for other factors.

These findings provide a significant contribution to the literature on impression management for many reasons. First, this work identifies specific *self-presentation tactics* that individuals adopt in order to appear *knowledgeable*. Second—and more importantly—we show that an individual’s use of different self-presentation tactics evolves over time in a predictable fashion. In the context of online ratings, this implies the same individual can be more positive or more negative in his or her opinions with the exact same self-presentation goal in mind—in this case, appearing knowledgeable. By demonstrating how the opinions that reviewers express online are influenced not only by the rating behavior of *other* community members (Schlosser 2005) but also by reviewers’ *own* past ratings behavior, and that this effect is driven by self-presentation concerns, this work extends previous literature on online word of mouth (Godes and Silva 2012; Moe and Trusov 2011; Moon, Bergey, and Iacobucci 2010; Sridhar and Srinivasan 2012;). It also extends past work in this area by examining *how* people talk about an experience rather than *which* experiences they talk about. Past work has focused largely on how self-presentation concerns drive *what* topic people talk about or word of mouth *incidence* (Moe and Schweidel 2012). In contrast, we focus on *how* consumers decide to talk about (or rate) a given experience (i.e., word of mouth *evaluation*) while controlling for what experiences consumers choose to talk about (rate).

The remainder of this paper proceeds as follows. We begin by briefly discussing the concept of impression management through strategic self-presentation before surveying the literature on *self-presentation goals* and *tactics* with a focus on how the latter can evolve over
time. We then turn to past work in marketing on *word of mouth online* and, building on past work, present our four principal hypotheses regarding self-presentation in online rating forums.

Three laboratory studies and the analysis of real-world data provide compelling evidence supporting the proposed causal mechanism. In Study 1A, we show how reviewers concerned with displaying the fact they make good choices give consistently higher ratings than reviewers concerned with displaying their critical skills. In Study 1B, we replicate this effect while showing how reviewers new to an online rating community are concerned primarily with displaying their ability to make good choices. The same reviewers, after accruing a tractable record of reviews, become increasingly concerned with displaying their critical skills and an ability to discriminate between experiences. This change in motivation contributes to a declining trend in an individual’s ratings over time.

We find additional support for our explanation by moderating the effect in Study 2. In this study, we document how the effect (more positive reviews early on) manifests for reviewers who are *high* in public self-consciousness (concerned with how they present themselves to others). This is not the case for those *low* in public self-consciousness. Finally, in Study 3, we examine a real-world data set from Yelp consisting of 330,070 reviews written by 130,872 individual reviewers. As expected, the data reveal a negative trend in ratings as a function of the number of businesses rated previously, controlling for other factors. Lending additional support, the written commentary accompanying these ratings (reviews) tends to be more positive early in a reviewer’s rating history and less positive later on. We conclude by discussing the theoretical and managerial implications of our findings.

**IMPRESSION MANAGEMENT AND SELF-PRESENTATION**
Impression management is the goal-directed activity of controlling or regulating information intended to influence the impressions an audience forms of a person, object, event, or idea. When the goal is to deliberately influence others’ impressions of one’s self (as opposed to other entities), the activity is called strategic self-presentation (Schlenker 1980). Because so much time is spent in the evaluative presence of others, strategic self-presentation is considered an inescapable feature of social life (Jones and Pittman 1982). Everyone is believed to engage in strategic self-presentation; the behavior is so pervasive that it often becomes habitual and automatic (Brown 2007). Individual differences do exist as people vary in the extent to which they are concerned with the impressions they convey to others. As mentioned earlier, this difference is what is known as public self-consciousness (Brown 2007; Fenigstein, Scheier, and Buss 1975).

What constitutes a desirable impression depends on the individual, the context, and the social circumstances. It is essential to point out that most scholars agree the top two impressions people try to convey are: (1) knowledgeability, being perceived as competent, skilled, and intelligent and (2) personability, being viewed as warm, likeable, friendly, and socially desirable (Leary et al. 1994; Leary 1995; Nezlek, Schütz, and Sellin 2007). In fact, Fiske and colleagues (Fiske et al. 1999; Fiske et al. 2002) have found that judgments of competence and warmth underlie most human perceptions in social interactions, driving both emotional and behavioral reactions to others. The specific impression one hopes to make is considered a self-presentational goal while the way one tries to convey that impression is considered a self-presentational tactic (Brown 2007).

Various taxonomies of self-presentational goals and tactics exist in the literature (e.g., Jones and Pittman 1982; Leary et al. 1994). What tactic is seen as the most effective depends on who the counterpart is, whether there are other people around during the interactions, and a variety of other contextual factors (Nezlek and Leary 2002). Jones and Pittman’s (1982) taxonomy yields the following illustrative example. With the goal of being perceived as likeable, one might employ one
or more of the following tactics: telling the person they support the same football team (conformity), complimenting the person about his or her clothes (flattery), or telling the person about an amazing vacation (bragging). Given the importance of context, the choice of self-presentation tactic can change over time, or evolve, and is informed by one’s self-presentation history as people modify their behavior based on the previous social interactions they had (Leary et al. 1994). Moreover, during a focal interaction, individuals have an expectation that their counterpart will recall the past during future interactions, which can alter their behavior in the present. Political scientist Robert Axelrod calls this effect the shadow of the future (Axelrod 1984).

In short, an effective tactic at the beginning of a relationship may become less effective as the number of interactions accrue because the efficacy of self-presentation tactics depends on whether others believe the impression sought is true or not. This implies a trade-off between: (1) beneficiality, presenting the most advantageous image possible, and (2) believability, making sure the image presented is believable (Schlenker 1980). What is both a beneficial and believable tactic at one point in time might not be both beneficial and believable later on (Tice et al. 1995). For instance, presenting oneself in only the most positive light becomes less believable as others become more familiar with the individual (Tice et al. 1995). Consequently, we argue that what constitutes a believable self-presentation tactic in an online rating community evolves over time. Our work is the first to show how an individual’s use of differing self-presentation tactics evolves over time.

SELF-PRESENTATION IN WOM COMMUNICATION

Past work suggests self-presentation matters in all interpersonal communication. A person’s attempts to project a particular image involve a wide array of decisions including which
topics to bring up in conversations as well as what is said about the topic (Berger 2014; Cialdini and Goldstein 2004; Eagly and Chaiken 1993; Schlenker 1980; Tetlock, Skitka, and Boettger 1989). Marketing is expressly concerned with when and how consumers talk about products and services, referred to as word of mouth (henceforth WOM). Stated more formally, consumer WOM involves the sharing of information about a brand, product, or service between a non-commercial communicator and one or more receivers, typically other consumers (Dichter 1966). While any type of WOM can be seen as an occasion for presenting oneself in a positive light (Berger 2014), consumers have been shown to be more likely to generate WOM in domains they deem more self-relevant, talking about products they believe communicate something about themselves to others (Chung and Darke 2006).

A reading of the literature on WOM suggests the main impression sought when making product recommendations is that of being perceived as knowledgeable. For example, Dichter's (1966) foundational investigation of WOM revealed that individuals attempt to support their self-image of superiority, connoisseurship, and expertise through WOM communications. More recently, Engel, Blackwell, and Miniard (1993) similarly argued that making product recommendations allows the reviewer to secure attention, display connoisseurship, and garner status. Not surprisingly, individuals who think of themselves as highly knowledgeable have been found to be more likely to share their opinions and do so in order to maintain a positive self-image (Feick and Price 1987; Walsh, Gwinner, and Swanson 2004). This is especially true when they have had a positive experience with a product or service (De Angelis et al. 2012; Wojnicki and Godes 2011). Therefore, based on past work, we expect reviewers to place greater importance, generally speaking, on being seen as knowledgeable as opposed to personable (Fiske, Cuddy, and Glick 2007; Gibson and Oberlander 2008; Holoien and Fiske 2013).
The literature does not accommodate a straightforward prediction regarding how a desire
to be perceived as knowledgeable would influence the valence of opinions people express (i.e.,
online ratings). On the one hand, Folkes and Sears (1977) find that more positive evaluators
(individuals expressing positive evaluations—whether about classes, movies, cities, or
politicians) are generally perceived to be more knowledgeable compared to relatively more
negative evaluators. Relatedly, De Angelis et al. (2012) report that the more uncertain an
individual is about her or his own competence, the more likely she or he is to talk about positive
product experiences as a way of bolstering her or his self-concept. In the same vein, other
research finds consumers can be reluctant to publicize negative product experiences because
doing so may give the impression they are unable to make good product choices (Wojnicki and
Godes 2011). Based on this work, one would expect consumers motivated to appear
knowledgeable to be more positive in their evaluations.

On the other hand, conflicting research supports the opposite prediction: the goal of
appearing knowledgeable leads to more negative evaluations. Work by Amabile (1983) argues
that, despite being liked less, negative evaluators are perceived to be more knowledgeable than
positive evaluators. She also finds consumers who are insecure about their competence—or find
themselves in a low status position—are more likely to produce negative evaluations of other
people’s work as a means of boosting their self-esteem (Amabile and Glazebrook 1980). In the
context of online evaluations, Moe and Schweidel (2012) present evidence that frequent
contributors to online rating communities (i.e., those who are highly involved and therefore likely
more concerned with displaying knowledgeable) post lower ratings compared to less frequent
contributors. Schlosser (2005) argues after being exposed to another reviewer’s negative online
review, a reviewer’s desire to look smarter causes them to be more negative. Based on this body
of work, one would expect consumers motivated to appear *knowledgeable* to be more negative in their evaluations.

The apparent conflict about how to appear knowledgeable, we argue, reflects two alternative self-presentation tactics individuals can adopt. One tactic involves displaying one’s ability to *make good product choices*, signified by being relatively more positive in one’s ratings and reviews. The alternative involves displaying *critical skills and a capacity to discriminate* by being relatively more negative in one’s ratings and reviews. We show how a reviewer’s rating record within a given rating community influences which of the two tactics s/he adopts and hence the valence of her or his next rating. The proposed explanatory mechanism gives rise to four formal hypotheses detailed in the next section.

**THE DYNAMIC INFLUENCE OF SELF-PRESENTATION CONCERNS ON RATINGS**

In online rating communities, a reviewer’s past ratings are publicly available for others to see. A new rating, therefore, is comparable to a new interaction in the offline world in its ability to provide incremental information about the individual. We argue the self-presentation tactic reviewers adopt when they are new to a community differs from the tactic reviewers adopt when they have been active members for some time. Our hypotheses are based on the following logic.

New members of an online community will be concerned about other members attributing negative ratings to incompetence (i.e., an inability to make good product choices). Hence, to appear knowledgeable, they are better off publicly displaying satisfaction with their choices rather than highlighting the negative. This parallels behavior documented offline in the terrestrial world, whereby being more boastful has been shown to be a more beneficial (and believable)
tactic when a counterpart doesn’t have past information about the boastful person (Tice et al. 1995). The result is an expectation of more positive online ratings early on.

Once a reviewer has established a reputation (for making good choices), s/he is less likely to run the risk of more negative ratings being attributable to poor product choices. Further, consistently positive ratings run the risk of being perceived as either indiscriminant and/or Pollyannaish. Therefore, s/he has an incentive to switch to a different self-presentation tactic, that of displaying critical skills and ability to discriminate. The change in tactics has downstream consequences in terms of the valence of ratings. For the reviewer growing increasingly concerned with displaying the critical skills necessary to discriminate among distinct experiences, positivity in one’s earlier ratings needs to be counterbalanced by a more critical approach and thus greater negativity in later ratings.

Our first two formal hypotheses center on the association between self-presentation tactics and rating valence as well as the evolution in self-presentation tactics in online rating forums:

**H1:** *A reviewer concerned with displaying an ability to make good choices provides more positive ratings than a reviewer concerned with displaying critical skills and a capacity to discriminate between experiences.*

**H2:** *As a reviewer accumulates ratings, the relative importance of displaying an ability to make good choices decreases while the relative importance of displaying critical skills and a capacity to discriminate between experiences increases.*

If reviewers, on average, are more concerned with displaying an ability to make good choices early on, the result would be a negative trend in ratings across individuals’ rating history. This trend is anticipated as follows:
**H3:** *A reviewer’s rating record (number of previous ratings posted) influences the valence of subsequent ratings: specifically, the higher (lower) the number of previous ratings, the lower (higher) the next rating.*

Two alternative explanations for a negative trend in individual ratings come to mind. First, reviewers may rate their “favorite” products and services first when they initially join a ratings community, resulting in more positive ratings out of the gate. This explanation cannot apply to our experiments in which respondents cannot choose which products to rate and is also addressed in our analysis of secondary data in Study 3. Second, while new to a rating community, reviewers may provide ratings for only those products and services for which they had a positive experience, a form of self-censorship (i.e., an effect of rating record on WOM *incidence*). This alternative explanation is addressed directly in Study 2 in which we collect both respondents’ likelihood to rate and rating measures.

In order to devise an alternative test for our mechanism, recall that individuals differ with respect to how concerned they are with the impressions they convey to others (Brown 2007). Fenigstein, Scheier, and Buss (1975) have developed a public self-consciousness scale to assess the degree to which people focus on the public, observable aspects of themselves. Those who are low in public self-consciousness should be less likely to exhibit behavior consistent with hypothesis 3. Hence, hypothesis 4 makes the following prediction (tested explicitly in Study 2):

**H4:** *The less a reviewer cares about the impressions conveyed to others (low public self-consciousness), the less likely it is that the number of his or her previous ratings (rating record) will impact the valence of subsequent ratings.*

*EMPIRICAL EVIDENCE*
We begin our empirical analysis with a pre-test. Respondents were given different self-presentation goals to test two fundamental assumptions motivating this research: (1) different self-presentation goals differentially influence the valence of ratings consumers contribute to user-generated review sites and (2) the default self-presentation goal customarily pursued in the context of online rating communities is that of being perceived as knowledgeable.

**PRE-TEST**

Respondents were given either the self-presentational goal of appearing knowledgeable or of appearing personable—or no goal at all (a control condition). Next, they were asked to rate a series of three contemporary paintings as members of an online community that publishes crowdsourced ratings and reviews of art called “ArtsyArt.com.” A priori we expected reviewers concerned about looking *personable* to be generally positive in their ratings based on literature demonstrating how positive (negative) individuals are perceived to be more (less) likeable (Bell 1978; Folkes and Sears 1977). In contrast, we expected reviewers concerned about looking *knowledgeable* to want to show discrimination in their evaluations and hence be more negative.

Note that respondents in this study were asked to evaluate a product they had *not* chosen personally and therefore were *not* expected to want to display they had made good choices. Instead, a desire to appear knowledgeable should motivate them to display their critical skills and an ability to discriminate. By comparison, having the goal of appearing knowledgeable should result in lower ratings, on average, compared to those whose goal was appearing personable. Moreover, we expected a desire to be perceived as knowledgeable to be the default goal, so we did not expect rating behavior to differ between those assigned the goal of being perceived as knowledgeable and the control group not assigned a goal.
Method

Participants were recruited through Amazon Mechanical Turk (mTurk). Eligible respondents were restricted to U.S. residents with a 95% or higher approval rate and a completion rate of at least 50 hits. Participants were compensated .50 USD. Only one response per IP address was allowed (Goodman, Cryder, and Cheema 2013). Two hundred and ten participants completed the survey. The study was scenario based, and participants first read about an online rating community of art lovers called ArtsyArt.com. They were then instructed to imagine that they were members of this community and asked to rate three paintings. They were told in no uncertain terms that their ratings would be made public as part of their online profile (similar to Yelp, TripAdvisor, and other such sites).

The study employed a single factor between subject design. Self-Presentation Goal was manipulated by asking respondents to rate each artwork with the objective of being perceived as either Knowledgeable (described as someone who is “competent, skilled, and intelligent”) or Personable (described as someone who is “likeable, friendly, and socially desirable”) by other community members. Respondents in the Control condition were not given any self-presentation goal. Participants were then shown the first painting, accompanied by the following (fictitious) information: artist, title, year, and materials. They were asked to rate the painting on a 10-point scale (1 = very bad, 10 = very good). They rated the second and third paintings in an equivalent fashion. The order in which the paintings were presented was counterbalanced.

The primary dependent variable of interest was respondents’ average rating for the three paintings. If respondents with the goal of appearing Personable provided consistently more positive ratings than those with the goal of appearing Knowledgeable, this would be reflected in a higher average rating. Further, if appearing knowledgeable is the default goal for members of
online rating communities, there should be no significant difference in average ratings between the Knowledgeable and Control conditions.

Results

An ANOVA with average rating as the dependent variable reveals that Self-Presentation Goal is a significant predictor ($F(2, 207) = 7.03, p < .01$). In line with expectations, planned contrasts reveal that the average rating of respondents in the Personable condition ($M = 6.84$) is significantly higher than the average rating of respondents in the Knowledgeable condition ($M = 6.30, t(1, 207) = 2.09, p = .04$) as well as the Control condition ($M = 5.92, t(1, 207) = 3.75, p < .01$). As expected, there is no difference in average rating between the Knowledgeable condition and the Control condition ($t(1, 207) = 1.55, p = .12$). The results are equivalent if the data are analyzed using a mixed ANOVA where Self-Presentation Goal is a between subject factor and Painting is a within subject factor.

Discussion

These results provide empirical evidence that self-presentation goals can influence what would otherwise be expected to be unadulterated consumer ratings and that the default goal pursued by reviewers in online communities is likely to be that of being perceived as knowledgeable. The pre-test, however, is limited in the following two ways. First, it analyzes the effects of self-presentation goals on online ratings in a somewhat static fashion (while each reviewer provided three ratings, they did so in a very brief time span), while our theorizing predicts self-presentation concerns influence reviewers’ online ratings differentially as a function of their ratings record. Typically, a reviewer’s ratings accumulate over many discrete rating occasions. Second, respondents were asked to evaluate three paintings that they had not chosen for themselves. This is unlikely to be the case in online rating communities; consumers
characteristically evaluate products or services they personally chose to experience (restaurants, movies, songs). These limitations are addressed in the subsequent studies.

**STUDY 1A: GOOD CHOICES VS. DISCRIMINATION**

Study 1A was designed to demonstrate that consumers who want to show their ability to make good choices exhibit rating behavior that differs systematically from consumers who want to show they possess the critical skills that provide the ability to be discriminating. We expect the former to rate the exact same product experience more positively than the latter. In this study, respondents were asked to listen to and then rate a song they had supposedly purchased. They were asked to do so with either the goal of showing their ability to make good product choices or displaying critical skills and an ability to discriminate.

**Method**

One hundred and twenty-six undergraduate business students from a major West Coast university participated in the study for partial course credit. Respondents read about a new (fictitious) online rating community named MyUniversityMusic.com (*University here is used to mask the acronym of the school*). The site was described as a university-based community of music lovers. Once again, we manipulated *Self-Presentation Goal*. In the Good Choice condition, respondents were told they should assume they wanted to appear as “someone who makes good music choices,” while in the Critical condition they wanted to appear “critical and discriminating.” They were also provided their username (chosen to remind them of the directive), “GoodChoice” [“Critical1”] and asked to make note of it as they would need to use it to log into the review site.
Next, respondents listened to a 40-second excerpt of one of two songs, selected at random. The songs were “Starz in Their Eyes” by Just Jack, and “Worst Case Scenario” by The Hoosiers, pretested to be equally liked by members of the same population (as expected and consistent with the pre-test, we do not find an effect of song in either Study 1A or Study 1B; hence, data are collapsed across songs). After logging into the online community website with the username provided, respondents rated the song. The rating was provided on a scale ranging from 1 to 10 stars. Respondents could also provide a written review if they liked.

Results

The data were analyzed using ANOVA with Rating as the dependent variable and Self-Presentation Goal (Good Choice vs. Critical) as the independent variable. Consistent with our first hypothesis, we expected respondents in the Good Choice condition to provide more positive ratings, on average, compared to those in the Critical condition. This is indeed what we observed ($M_{\text{GoodChoice}} = 5.68$ vs. $M_{\text{Critical}} = 3.98$, $F(1,124) = 19.83, p < .01$).

Discussion

Study 1A provides compelling evidence that online community members concerned with exhibiting an ability to make good choices provide higher product ratings compared to community members concerned with displaying their critical skills. This pattern of results was predicted by our first hypothesis (H1). What Study 1A does not show, however, is that individuals who want to appear knowledgeable (the self-presentation goal customarily pursued by online community members) change which of these two tactics they employ over time (H2), and as a result, exhibit a declining trend in ratings (H3). Study 1B was designed to address H2.
and H3 by testing the relationship between the choice of self-presentation tactics and the corresponding trend in ratings.

**STUDY 1B: THE EVOLUTION OF SELF-PRESENTATION TACTICS**

In this study, as in study 1A, we asked each respondent to rate songs as though they were part of an online rating community. At the onset, they were instructed to imagine they would be rating their first song for the site. After a significant delay, they were asked to rate a second song as if it was the 81st rating they were posting online. Simultaneously, we provided respondents one of three different goals (one general, and two specific): (1) the general goal of displaying knowledgeability, (2) the specific tactic of displaying one’s ability of making good music choices, or (3) the specific tactic of showing critical skills and discrimination.

A priori, we expected to see a declining trend in respondents’ ratings only when they were given the general goal of displaying knowledgeability, as this would lead respondents to change tactics between their first and 81st rating. If hypotheses 2 and 3 are true, we would expect ratings to be more positive for the first rating but only for those assigned the goal of displaying knowledgeability. We would also expect the first ratings to be comparable to ratings from those assigned the tactic of displaying their ability of making good music choices and different from those assigned the tactic of showing discrimination. Conversely, we also expect the 81st rating to be more negative. It should also be comparable between those assigned the goal of displaying knowledgeability and those assigned the tactic of showing discrimination, as well as different from those assigned the tactic of displaying that they make good choices.

*Method*
Ninety undergraduate business students at a major West Coast university participated in the study for partial course credit. Respondents read about a new (fictitious) university-based online rating community of music lovers. They were asked to imagine they had recently joined the community and cared about the impression they would make on other members. In particular, we manipulated Self-Presentation Goal by telling them they cared about being perceived as “knowledgeable (competent, skilled, and intelligent)” in the Knowledgeable condition, “someone who makes good music choices” in the Good Choice condition, and “critical and discriminating” in the Critical condition. As a reminder, they were told they had chosen the username of “Smart1,” “GoodChoice,” or “Critical1,” respectively, and were asked to take note of this username and use it when logging into the system.

Previous Ratings was manipulated within-subject. Respondents were asked to imagine they had recently purchased a song and were about to rate it on the website. This was their first rating. They listened to a 40-second excerpt of one of the same two songs used in Study 1A, selected at random, before logging in with their username in order to rate the song. The rating was provided on a scale ranging from 1 to 10 stars. Respondents next moved on to an unrelated study that lasted approximately 15 minutes before coming to the second part of our study. They were asked to imagine a few months had gone by and that they had been active members on the site during that time. They were reminded of their self-presentation goal and asked to imagine they had recently purchased another song and were about to rate it. This was going to be their 81st rating. They listened to the second excerpt (the song they did not hear in the first part of the study), logged into their profile page with the username and rated the song.

Results

The data were analyzed using ANOVA with Rating as the dependent variable, Self-
Presentation Goal (Knowledgeable vs. Good Choice vs. Critical) as a between-subject independent variable and Previous Ratings (0 vs. 80) as a within-subject independent variable. Consistent with our third hypothesis (H3), a priori we expected to observe a declining trend in ratings for those in the Knowledgeable condition but not for ratings in either the Good Choice or Critical conditions. This is exactly what we observed. In the Knowledgeable condition, the song’s rating is significantly higher when respondents provided their first rating ($M = 6.07$) compared to their 81st rating ($M = 4.97$, $F(1,87) = 4.61$, $p = .03$) while ratings do not differ significantly in either the Good Choice ($M_0 = 5.73$ vs. $M_{80} = 6.00$, $F(1,87) = .26$, $p = .61$) or the Critical condition ($M_0 = 4.66$ vs. $M_{80} = 4.24$, $F(1,87) = .61$, $p = .44$).

Also, as expected and consistent with our second hypothesis (H2), for respondents giving their first rating, those assigned the goal of showing knowledgeability behave no differently from those who want to show they make good choices ($M_{Knowledgeable} = 6.07$ vs. $M_{GoodChoice} = 5.73$, $F(2,87) = .65$, $p = .52$) but differently from those charged with displaying their critical skills ($M_{Critical} = 4.66$, $F(2,87) = 5.11$, $p < .01$). Likewise, as expected, the opposite occurred when giving their 81st rating. Respondents behaved no differently than those who want to display their critical skills ($M_{Knowledgeable} = 4.97$ vs. $M_{Critical} = 4.24$, $F(2,87) = 1.45$, $p = .24$) but differently from those who want to show they make good choices ($M_{GoodChoice} = 6.00$, $F(2,87) = 5.55$, $p < .01$).

Discussion

Study 1B demonstrates how consumers become more negative in their ratings when they want to appear knowledgeable to other community members. The pattern of results is consistent with our process explanation and supports our hypotheses: online reviewers provide more positive reviews early on because they are concerned with displaying their ability to make good product choices while later on they are more negative, relatively speaking, because they are
concerned with showing they possess critical skills and are able to discriminate between experiences. Taken together, this results in a negative trend across individual ratings records.

STUDY 2: THE ROLE OF PUBLIC SELF-CONSCIOUSNESS

In Study 2, we take an alternative route toward testing our proposed process explanation by exploring whether and how public self-consciousness, an individual difference reflecting individuals’ concern for their public appearance, impacts the relationship between online reviewers’ rating records and the valence of subsequent ratings. Failure to exhibit a negative trend when self-presentation concerns are not present (low public self-consciousness) would provide confirmatory evidence that self-presentation concerns contribute to the pattern of results observed in Study 1 as well as offer direct support for hypothesis 4. Moreover, any attenuation of the effect would highlight an important boundary condition to the generalizability of our results; self-presentation tactics should not matter for reviewers unconcerned with how they appear to others; we should observe no discernible effect on their ratings.

In this study, we investigate rating behavior in a new domain. Rather than rating art or music, respondents were asked to rate a university professor within a school-based online rating community. The study was scenario-based, and respondents were told either that they were new to the community or that they had been a member for some time and already had provided a considerable number of ratings. We expected the number of previous ratings to influence the valence of the focal rating (more positive for those with fewer ratings) and that this effect would diminish along with the respondent’s idiosyncratic level of Public Self-Consciousness.

Method
One hundred and sixty-three undergraduate business students from a major West Coast university participated in the study for partial course credit. Trait Public Self-Consciousness was measured using the 5-item Public Self-Consciousness scale developed by Fenigstein, Scheier, and Buss (1975). This measure was taken in a pre-screening session at the beginning of the semester (approximately one month before the study took place).

Respondents read about an online rating community associated with their university described as “the best source of your school’s class and professor reviews and ratings based on student feedback.” Respondents imagined they were members of the community and were told “this website works very much like Yelp: students create a public profile, give ratings, and write reviews, etc. Yet, rather than rating restaurants, users rate their professors. Other members of the community can in turn rate your reviews and decide to ‘follow you’ if they find your reviews to be interesting and insightful.” Students were also reassured that only fellow students would be able to access the community. This was stated explicitly to eliminate any potential concerns regarding their professors reading what they had to say about them. They were then asked to imagine they had chosen to attend a consumer psychology class and read a detailed description of the class and the behavior of the professor (Professor Smith, see description Appendix A).

Respondents rated the professor based on the description of their experience. The number of previous ratings (Rating Record) was varied by telling respondents to imagine they were about to give either their first rating or their 31st rating (Previous Ratings: 0 vs. 30). Each respondent rated the professor in terms of clarity, usefulness, and easiness (all positive characteristics among undergraduate students). All ratings were provided on 9-point scales. These measures were combined to create an index of overall “quality” just as on an actual faculty rating site, Ratemyprofessor.com. Finally, respondents were asked how likely they would be to go online to rate Professor Smith, also on a 9-point scale (1 = not at all likely; 9 = very likely). The likelihood
of providing a rating is an important covariate in the subsequent analysis allowing us to control for any effect of self-presentation concerns on incidence. Finding a residual effect of reviewers’ record on the valence of their next rating (evaluation) above and beyond any effect on their likelihood to evaluate an experience (incidence) would provide even stronger evidence in support of our causal reasoning.

Results

In line with our predictions, an OLS regression with average rating as the dependent variable reveals the effect of Previous Ratings in predicting the average rating given to the professor is a function of the respondent’s level of Public Self-Consciousness ($b_{PSC \times \text{PreviousRatings}} = .58, p < .01$) while controlling for the likelihood they would go online to provide a rating. See Table 1 for detailed results.

For a clearer interpretation of the results, we utilize floodlight analysis to illustrate the effect of public self-consciousness on the relationship between previous ratings and the current rating (Spiller et al. 2013). In Figure 1, we identify regions within the range of public self-consciousness in which the effect of the independent variable on the dependent variable is and is not significant (Hayes and Matthes 2009; Johnson and Neyman 1936).

Insert Table 1 about here.

Insert Figure 1 about here.
The analysis allows us to identify two distinct Johnson–Neyman points with \( p < .05 \). The first occurs at a very low level of Public Self-Consciousness (1.77) while the second occurs at a higher level of Public Self-Consciousness (3.04). In line with our prediction, the results suggest that individuals with high levels of Public Self-Consciousness (> 3.04, corresponding to 20.25% of the sample) give a significantly higher rating to Professor Smith if they are giving their first rating as opposed to their 31st rating. Unexpectedly, we observe the opposite pattern (a positive trend) for individuals with very low levels of Public Self-Consciousness (< 1.77, corresponding to 10.43% of our sample).

**Discussion**

Study 2 reveals that individuals with high levels of public self-consciousness display the negative trend in ratings that we expect; they are more positive in their first rating compared to their 31st rating. These individuals comprise about 20% of the sample. Reviewers with high levels of public self-consciousness likely account for a far greater percentage of active reviewers in online rating communities as self-presentation concerns and reputation seeking are believed to be significant drivers of participation in these communities (Lampel and Bhalla 2007).

These results are important for two reasons. First, they provide further evidence that a declining trend in online ratings is, at least in part, due to self-presentation concerns. Second, they highlight an important boundary condition such that if consumers participating in online rating communities are unconcerned about how they appear publicly, there is no reason to expect their ratings to decline as the number of reviews increases. These results offer direct support for our fourth hypothesis (H4).

Two other points are worth highlighting. First, we observe an unexpected positive trend in ratings for consumers with very low levels of public self-consciousness. One explanation, while
only speculative, is that respondents unconcerned about how they appear to others were making certain attributions to justify why they have provided such a large number of ratings (30). They may believe the reason they continued to rate their professors is that they increasingly liked their instructors and this may have had a positive effect on their 31st rating. The second point worth mentioning is that we find a positive correlation between rating valence and respondents’ likelihood to rate (see Table 1). This finding is consistent with past literature suggesting consumers are more likely to publicize their positive experiences as opposed to their negative experiences (De Angelis et al. 2012; Wojnicki and Godes 2011). Importantly, we find an effect of self-presentation concerns on rating valence that goes above and beyond that of rating incidence.

**STUDY 3: INCREASING NEGATIVITY AS RATINGS ACCUMULATE**

In Study 3, we analyze nine years (2004-2012) of real world data from Yelp, which publishes crowd-sourced ratings and reviews about a variety of local businesses, albeit principally restaurants. If the number of ratings a reviewer has already provided (rating record) influences the self-presentation tactic s/he chooses to adopt in the rating community, this should result in reviewers becoming increasingly more negative as the number of ratings they have previously provided increases. In other words, if hypothesis 3 is correct, we should observe a negative trend in individual Yelpers’ ratings as a function of the number of businesses already rated. While such a finding does not provide irrefutable evidence of the causal mechanism we propose, it would corroborate results from earlier laboratory studies, providing compelling evidence of the real world impact of self-presentation motives in rating forums such as Yelp. Our predictions are investigated in two different ways: in Study 3A, we look for a negative trend in individual reviews’ scores (ratings) while in Study 3B, we pay closer attention to the content of
the reviews investigating whether a similar trend exists with respect to the written text of the review.

**STUDY 3A: REVIEWERS GIVE MORE NEGATIVE SCORES OVER TIME**

*Data*

The dataset used for this analysis is Yelp’s Academic Dataset, available to researchers on Yelp’s web site in 2015. The data include reviews from 30 cities across the U.S. It includes all Yelp ratings and reviews for 13,481 businesses in the U.S. between 2004, when Yelp was first introduced, and 2012. A unique feature of this dataset is the inclusion of reviewers’ first reviews ever on Yelp providing the rating records for reviewers needed for our analyses. The data consists of 330,070 reviews by 130,872 reviewers. On Yelp, reviewers rate businesses on a 5-point scale ranging from 1 (“Eek”) to 5 (“Woohoo”). Rating is the first key dependent variable in our analysis. The main independent variable is Number of Previous Ratings (Number of Previous Ratings _Rev hereafter), a count of the number of ratings posted by the reviewer before giving the focal rating. Because our interest is investigating the trend of review scores “within” reviewers, we start by focusing on a balanced subsample of the dataset comprised of the first 20 reviews of reviewers who had provided at least 20 reviews. This allows us to capture the trend within the individual reviews and leaves us with an initial dataset of 31,900 ratings provided by 1,594 reviewers. To reassure the reader, we relax these requirements later in our robustness checks.

We included a number of covariates to control for important features of the business being reviewed as well as individual reviewer characteristics likely to influence their ratings. The first set of covariates, including the average rating of the business at the time the focal rating was provided, Average of Previous Ratings for the Business (Average Ratings _Bus), and the number
of previous ratings for that business at that point, Number of Previous Ratings for the Business (Number of Prev. Ratings_Bus), help gauge a business’s quality and popularity at the time the focal rating was posted. Including these covariates reduces the sample from 31,900 to 29,651 ratings (we have missing values whenever the number of previous ratings for the business is 0). These variables also help account for social dynamics in rating behavior (i.e., the influence of others’ opinions) as highlighted in previous literature (see, for example, Godes and Silva 2012; Moe and Trusov 2011; Moon, Berge, and Iacobucci 2010; Sridhar and Srinivasan 2012). Detecting an effect above and beyond any effect of a business’s past ratings would offer especially strong evidence that ratings are influenced by the individual’s own rating record.

A second type of control variable includes the Business Type, a categorical variable indicating whether the business rated was a restaurant (80.4% of the reviews in our sample), hotel (.4%), festival (.2%), venue (.2%), or the residual category of “other businesses (18.2%),” as well as State, a categorical variable indicating the geographical location of the business. A dummy variable, Bulk, captures whether or not a rating from a specific reviewer was provided the same day as one or more other ratings from that reviewer. We control for time with the inclusion of Year dummies.

Finally, a third set of covariates in our analysis included Average Ratings of the Reviewer (Average Ratings_Rev), the running mean of the previous ratings of the reviewer before giving the focal rating, and Time Since First Review, the amount of time a reviewer has been reviewing on Yelp. We include these control variables in the robustness checks, as a way of capturing the influence of the reviewer’s previous experiences.

Method

In this context, ratings from the same reviewer cannot be treated as independent from one
another, as they are “nested.” Essentially, each rating from a specific reviewer is an occasion “i” such that occasions are clustered around the reviewer “j.” Each reviewer “j” can thus be seen as a level of a factor having a specific random effect on the specific rating score “i.” Since we are interested in inferences when both clusters (reviewers) and units (ratings) are viewed as sampled from respective populations, we employed a random-effects modeling approach (Rabe-Hesketh and Skrondal 2012). Considering the ordinal nature of our dependent variable, we use a Mixed-Effects Ordinal Logit regression model. To account for potential heteroscedasticity, we estimated the effects with robust standard errors.

**Results**

Model 1 (see Table 2) provides details regarding our main model, predicting Ratings through our key independent variable (Number of Previous Ratings_Rev) as well as the set of covariates related to features of the business (Average Rating_Bus and Number of Previous Ratings_Bus), as well as the dummy variable Bulk to account for bundled reviews. Results reveal that, as expected, indicators of perceived quality (Average Rating_Bus) and popularity (Number of Previous Ratings_Bus) positively affect the subsequent rating of a service. More importantly, we find evidence consistent with prior studies—after controlling for business quality and popularity, a greater number of reviews accumulated by a reviewer negatively impacts the subsequent rating of the experience (Odds Ratio= .988; p = .000). A crude synthesis of this result is that reviewers seem to become “less generous” once they accumulate more reviews. Looking at the odds ratio for the number of prior reviews, we see the addition of one additional review reduces the chance a rating falls into a higher (vs. lower) rating category by 1.2%, an effect that would likely be of interest to managers and users alike.
Table 3 shows how the likelihood of ratings falling in each of the five possible rating categories evolves as the number of previous ratings increases. What stands out is the positive relationship between the number of previous reviews accumulated and the chance of providing a negative review in particular. Consider that when a reviewer has completed two reviews, the likelihood that s/he gives a very negative rating (i.e., one or two stars) is 12.6%, while after s/he has completed 19 reviews, this likelihood increases to 14.9%. At the same time, after two reviews, the likelihood a reviewer gives a very positive score (i.e., five stars) is 17.7%, but after 19 reviews, this likelihood falls to 15.1%. These findings support the idea that reviewers become more critical and thus provide more negative ratings and less positive ratings once they have accumulated a larger number of ratings.

Robustness Checks

In order to test the robustness of our results, we ran several other models. The results confirm a consistent negative effect of prior ratings (results from Models 2 to 8 are provided in Appendix B). Model 2 adds a covariate at the reviewer level (Previous Ratings_Rev) to account for a reviewer’s past behavior and experience. Results confirm the negative effect of a larger number of prior ratings on the subsequent rating (Odds Ratio = .991; p = .000).

Model 3 replicates the random-effect analysis in an open, unbalanced sample. In this sample, we consider all ratings of each reviewer who provided at least 20 ratings. The results demonstrate a consistent effect of the number of previous ratings accumulated (Odds Ratio =
.998; \( p = .03 \)). The fact that the effect on the first 20 ratings (Model 1) is stronger than this subsequent estimation appears to suggest a diminishing effect as the number of previous reviews grows large, an issue that we explore in the next models in which we inspect the presence of a non-linear (i.e., quadratic) effect of our main independent variable.

Results from Model 4 support a positive curvilinear effect, which suggests that the negative effect of the number of prior ratings on the subsequent rating score diminishes at the increase of the number of reviews (\( \chi^2(1) = 16.01; p>\chi^2 = .000 \)). In Model 5, we remove another initial restriction (reviewers with at least 20 ratings) and re-estimate the model including all the reviewers who provided at least 5 ratings. The sample size increases to 151,996 ratings and 14,348 reviewers, but the coefficient for the number of prior ratings remains substantially in line with previous results (Odds ratio = .997; \( p = .000 \)), confirming that accumulating ratings makes reviewers less generous even for those who have provided a much smaller number of ratings.

In Model 6, we introduce a time covariate to control for differences in the time that elapsed since the first rating was provided by that specific reviewer: this helps to differentiate situations in which a reviewer accumulates many ratings in a short time from those in which a reviewer accumulates many ratings over a longer time span. Even in this case, results for our independent variable are similar to the original models (Odds Ratio = .987; \( p = .000 \)). In Model 7, we account for potential autocorrelation effects by re-estimating the analysis with the Generalized Estimated Equation approach (Liang and Zeger 1986). Replicating the analysis using a population average panel model provides consistent results (\( b = -.004; p = .000 \)).

Finally, in Model 8 we address the potential alternative explanation that reviewers start out by rating their favorite businesses when they first join Yelp. We therefore delete the first three reviews of each reviewer (their three favorite businesses based on such explanation) and
maintain a balanced sample with a 20 reviews window by looking at reviewers’ third to twenty-third reviews. Even in this case, the coefficient for our independent variable is still negative and significant (Odds Ratio = .995; $p = .030$).

Next, in Study 3B, we explore whether the declining trend in Yelpers’ individual ratings is reflected similarly in differences in the content of the written reviews that accompany these ratings. The findings from Study 3B are important because they help explain the results observed in Study 3A, further validating our proposed mechanism.

**STUDY 3B: REVIEWERS (ALSO) WRITE MORE NEGATIVE REVIEWS OVER TIME**

For this study, we exploit the fact that our original Yelp dataset also includes the text of the reviews associated with each rating and enough text to analyze in a rigorous fashion. The goal is to test whether, once a reviewer accumulates a greater number of ratings and reviews, the content of what s/he writes to accompany a rating also changes. This is important since previous literature has highlighted that synthetic ratings can often be inconsistent with the written reviews that accompany them, which can negatively influence the credibility of a reviewer (Schlosser 2011). Specifically, we set out to test whether the emotional valence of a reviewer’s early reviews is more positive than the valence of later reviews. If this is indeed the case, it would provide additional evidence that reviewers become more critical as they produce a greater number of reviews. In Study 3B, we analyzed the content of individual Yelpers’ written reviews with the aid of LIWC (Linguistic Inquiry and Word Count) computerized text analysis software.

*Data*

In order to obtain a comparable and manageable set of written reviews to analyze, we
extracted all reviewers who provided at least 20 reviews in our dataset. This is the same sample size in terms of reviewers as in study 3A from 1,594 reviewers, however we have 29,648 reviews (recall we had 29,651 ratings) due to the fact that three ratings were provided without text. The written content of these 29,648 reviews was analyzed using the standard dictionary of the LIWC software (Pennebaker, Francis, and Booth 2001).

The focal independent variable is the number of the review in the reviewer’s review record (Number of Previous Reviews), which ranges from 2 to 20 (as in study 3a). Our dependent variable is Net positivity, captures the overall positive emotional valence of each review. For each review, the LIWC software provided a positive emotion and negative emotion score, two indicators of the emotionally charged words embedded in the language of the review (Pennebaker, Francis, and Booth 2001). We computed the difference between these scores in order to obtain a single measure of the overall emotional valence of the review (more positive or more negative).

Our analysis includes a number of covariates also used in Study 3A, intended to capture the perceived quality and the popularity of the business being evaluated. This includes the reviewer’s rating of the focal restaurant (Stars) to control for the overall assessment of the experience. Also included are Number of Prev. Ratings_Bus (i.e., the popularity of the specific restaurant), and Average Ratings_Bus (i.e., the average quality of the restaurant up to the point of the focal review). Our model also includes controls for Business Type and Year (covariates used in Study 3A as well).

**Method and Results**

Similar to our analysis of ratings, we estimated a mixed effects linear regression with robust standard errors and a random effect at the reviewer level to capture the nested nature of
our data (complete results in Appendix C). Results reveal that later reviews have a smaller positive valence (are more negative) than earlier reviews ($b = -.021; p = .000$): our Net Positivity index reaches a value of 4.42 for the second review and moves down to 4.03 for the 20th review, a reduction of about 9.3% of the standard deviation of the index. These results suggest that reviewers tend to become more negative in the tone of their written reviews as the number of past reviews increases and provides further support for our conceptualization.

As a related finding, we also note that early reviews are, on average, shorter ($M_{2nd} = 129$ vs. $M_{20th} = 136$) by about 5.3% (7 of 129 words, on average) compared to late reviews ($p = .000$). Adding one review in the review record, on the average, increases the word count of reviews by .41 words ($b = .414; p = .000$). This implies that, later on, reviewers are more comprehensive in their assessment and use more words in the attempt to be discriminating.

Discussion of Study 3A and 3B

Together, the results of Study 3A and Study 3B provide significant real-world evidence consistent with earlier laboratory findings. Reviewers on Yelp become increasingly negative as the number of ratings and reviews they have provided increases. This negative trend is reflected by both lower synthetic ratings and more negative content accompanying those ratings. While the data from Study 3 cannot speak to the causal mechanism directly, the well-controlled laboratory studies reported earlier suggest this effect is driven by self-presentation motives—in this case, reviewers’ desire to be perceived as knowledgeable by fellow Yelpers. While new to the community and wanting to appear knowledgeable, reviewers are more concerned with displaying their ability to make good choices and provide comparatively more positive ratings and reviews. After accumulating a sizable number of ratings and reviews in their profile, reviewers become increasingly concerned with displaying critical skills and an ability to discriminate, resulting in
more negative ratings and less positive content in the written reviews they post.

One alternative explanation we address is the possibility that Yelpers start out by evaluating only their positive experiences. This was dealt with in Study 3A by omitting the first three reviews of each reviewer, and we still observe the effect. Also recall in Study 2 we found a significant effect of the number of previous ratings on rating valence controlling for incidence. Consequently, we believe self-presentation’s direct influence on rating valence to be a significant, albeit most likely a contributory explanation for the observed downward trend.

GENERAL DISCUSSION

An enormous number of opinions posted online are provided in the context of rating communities such as Yelp and TripAdvisor. Is it realistic to assume that consumer ratings provided in such contexts are unaffected by the reviewers’ concerns about what other community members will think of them? Across four studies, combining laboratory studies with the analysis of real-world data, our research shows that such an assumption is unwarranted. Our results reveal how members of online rating communities are concerned about appearing knowledgeable to other community members. This is likely exacerbated by the introduction of reputation systems (Dellarocas 2003) in online rating sites. While reputations systems have undoubtedly encouraged consumer participation in online forums, and arguably contributed to keeping these platforms flourishing, our work highlights an unexplored and unexpected repercussion—anonymous reviewers manifesting self-presentation concerns. The concern with appearing knowledgeable results in self-presentation tactics that evolve as reviewers accumulate an expanding portfolio of reviews. The end result is increasingly negative evaluations.

Reputation systems are intended to incentivize community members to build an online
persona and to care about what others think about their (quasi anonymous) persona. The desire for this persona to be perceived as knowledgeable and gain a reputation as such in the community (thereby increasing their number of followers, procuring “Elite” status, etc.) leads consumers to alter their ratings and reviews in a systematic and dynamically evolving fashion. In particular, we show online reviewers who are new to a community are particularly concerned with enhancing their reputation by displaying their ability to make good choices. Such a self-presentation tactic inevitably leads them to become more generous early on in their ratings and reviews. Later, this self-presentation concern is replaced by a desire to display critical skills and an ability to discriminate, which in turn leads reviewers to become more negative in their ratings and reviews. The end result is an overall negative trend in individual ratings, an effect we document across a wide variety of product and service categories in our studies and that is also reflected in the written reviews that accompany such ratings.

These findings contribute to several related literatures. First, we provide a meaningful contribution to the literature on impression management and self-presentation by identifying specific self-presentation tactics that individuals adopt in order to appear knowledgeable and, more importantly, showing that an individual’s use of different self-presentation tactics evolves over time in a predictable fashion. Second, these findings make an important contribution to the literature on WOM by showing that an individual’s online ratings are not just influenced by the rating behavior of other community members, as previously shown in the literature (Godes and Silva 2012; Moe and Trusov 2011; Moon, Berge, and Iacobucci 2010; Schlosser 2005; Sridhar and Srinivasan 2012) but are also influenced by their own past ratings. Third, our work also has interesting implications for the literature on sequential evaluations. Historically, that literature has focused on the importance of assimilation and/or contrast to prior experiences within the sequence and the resultant effect on evaluations (see Ghoshal et al. 2014). What we show is much
different; an individual’s self-presentation concerns can also have an impact on how public evaluations of a sequence of experiences unfold.

Our work raises important questions for future research. First, future research might attempt to integrate the impact of one’s own past behavior with the impact of others’ rating behavior. For example, Schlosser (2005) shows people become more negative following exposure to someone else’s negative review, but would this effect hold when someone is new to a community and trying to prove that they make good choices? When and how what I have done matters more than what others have done when deliberating impression management strategies appears to be a fruitful avenue for future research.

Second, we have argued that reviewers alter their evaluations in order to gain a reputation in a given online rating community. This reputation can come in several different ways, such as earning a “badge,” increasing one’s number of followers, or simply obtaining positive feedback on a review from other users. An interesting question that remains is what happens after reviewers obtain one or more of these validations: does it affect the choice of self-presentation tactic? Do reviewers become less concerned about the impression they make on others? These are important questions that future research could address.

A third set of questions that this work raises is to what extent consumers are aware of the fact they may be altering their ratings in order to enhance their reputation within the online community. We know from previous research that self-presentation behaviors can be habitual and automatic (Brown 2007). Knowing whether reviewers in online communities are acting in an automatic, unconscious fashion while providing ratings due to implicit self-presentation goals is essential when deliberating whether and how platforms such as Yelp and TripAdvisor might reduce or eliminate their influence.

Marketers, business owners, and online rating platforms should be aware of the potential
downsides of the introduction of reputation systems in online rating forums identified in this work (i.e., distorted ratings, even if modestly so). The more reviewers care about making a good impression on others, the more likely they are to alter their ratings and reviews in the hope of strategically improving their reputation within the community. Interestingly, many online platforms are trying to incentivize their users to connect with their offline social networks (e.g., Yelp’s Find Your Friends function). Managers should be aware that this may have important repercussions on the way in which individuals provide ratings, as this is likely to have an impact on their self-presentation concerns and the tactics they adopt. Recall, for instance, that individuals are typically more boastful with friends and more modest with strangers (Tice et al. 1995).

Interestingly, if an online community such as Yelp were to witness its growth rate slow and its reviewer base plateau, this should be reflected in a consistently more negative set of reviews (i.e., few if any new reviewers who are more positive). Such a phenomenon would be consistent with this research and literature documenting an overall declining macro-trend in book ratings over time (see Godes and Silva 2012). Marketers should consider drivers of trends over time whenever relying on online ratings to assess the success of their initiatives. While their ratings may degrade over time, their performance may not be deteriorating; the decline may be an artifact of impression management strategies among reviewers.
REFERENCES


### TABLE 1: REGRESSION COEFFICIENTS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>S.E.</th>
<th>T</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous Ratings (1=None, 0=30)</td>
<td>-1.46**</td>
<td>(.55)</td>
<td>-2.65</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>PCS</td>
<td>-.38**</td>
<td>(.15)</td>
<td>-2.62</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Previous Ratings X PSC</td>
<td>58**</td>
<td>(.20)</td>
<td>2.84</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Likelihood to Rate</td>
<td>.07*</td>
<td>(.03)</td>
<td>2.38</td>
<td>.02</td>
</tr>
<tr>
<td>Constant</td>
<td>6.58**</td>
<td>(.41)</td>
<td>15.96</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

* p<.05; **p<.01

### TABLE 2: MIXED EFFECTS ORDINAL LOGIT REGRESSION

<table>
<thead>
<tr>
<th>Model 1</th>
<th>OR</th>
<th>p</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Prev. Ratings_Rev</td>
<td>.988*</td>
<td>.000</td>
<td>.984 .992</td>
</tr>
<tr>
<td>Bulk</td>
<td>1.020</td>
<td>.419</td>
<td>.972 1.070</td>
</tr>
<tr>
<td>Average Ratings_Bus</td>
<td>3.698*</td>
<td>.000</td>
<td>3.523 3.874</td>
</tr>
<tr>
<td>Number of Prev. Ratings_Bus</td>
<td>1.000*</td>
<td>.000</td>
<td>1.000 1.001</td>
</tr>
<tr>
<td>Reviewer Var. (_cons)</td>
<td>.348</td>
<td></td>
<td>0.307 0.394</td>
</tr>
</tbody>
</table>

Observations 29,651
Groups 1,594
Wald $\chi^2$ 4,274.82

* p<.01. Business Type, State, and Year dummies included
TABLE 3: PREDICTED PROBABILITIES OF THE DIFFERENT RATING VALUES IN THE CASE OF 3 (SMALL) VS. 20 (BIG) PRIOR REVIEW RECORD*

<table>
<thead>
<tr>
<th>Stars</th>
<th>2 Prior Reviews</th>
<th>19 Prior Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>=1</td>
<td>3.3%</td>
<td>4.0%</td>
</tr>
<tr>
<td>=2</td>
<td>9.3%</td>
<td>10.9%</td>
</tr>
<tr>
<td>=3</td>
<td>26.9%</td>
<td>29.2%</td>
</tr>
<tr>
<td>=4</td>
<td>42.8%</td>
<td>40.8%</td>
</tr>
<tr>
<td>=5</td>
<td>17.7%</td>
<td>15.1%</td>
</tr>
</tbody>
</table>

* All other variables kept at their mean level

FIGURE 1: THE MODERATING ROLE OF PUBLIC SELF-CONSCIOUSNESS
APPENDIX A: Scenario Study 2

This semester you have decided to take Professor Smith's Consumer Psychology class. This was an elective class that you were really keen on taking.

Here is what you think of the class.

- The content is relevant and you find it inherently interesting;
- Professor Smith is personable but a little disorganized. Sometimes he goes off on a tangent during class and you find it hard to follow his explanations;
- The workload is reasonable. He gives weekly assignments that are never too difficult or time-consuming;
- You like the guest speakers he has invited to class. Their lectures were both current and insightful;
- Lecture slides are pretty useful, but the textbook adopted for the class is really boring;
- Professor Smith has a tendency to cold call people in class, which you are not a big fan of;
- Professor Smith is a tough but fair grader.
### APPENDIX B: ROBUSTNESS CHECKS, MIXED EFFECTS ORDINAL LOGIT REGRESSION

<table>
<thead>
<tr>
<th></th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Model 7^</th>
<th>Model 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
<td>b</td>
<td>OR</td>
</tr>
<tr>
<td>Number of Prev. Ratings_Rev</td>
<td>.991**</td>
<td>.998*</td>
<td>.961**</td>
<td>.997**</td>
<td>.987**</td>
<td>-.004**</td>
<td>.995*</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.030</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.030</td>
</tr>
<tr>
<td>Number of Prev. Ratings_Rev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sq</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulk</td>
<td>1.039</td>
<td>.993</td>
<td>1.026</td>
<td>1.019</td>
<td>1.022</td>
<td>.016</td>
<td>.998</td>
</tr>
<tr>
<td></td>
<td>.123</td>
<td>.732</td>
<td>.300</td>
<td>.079</td>
<td>.372</td>
<td>.198</td>
<td>.686</td>
</tr>
<tr>
<td>Average Ratings_Bus</td>
<td>3.684**</td>
<td>3.674**</td>
<td>3.694**</td>
<td>3.792**</td>
<td>3.696**</td>
<td>.631**</td>
<td>.374**</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Number of Prev. Ratings_Bus</td>
<td>1.000**</td>
<td>1.001**</td>
<td>1.000**</td>
<td>1.000**</td>
<td>1.000**</td>
<td>.000**</td>
<td>1.001**</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Average Ratings_Rev</td>
<td>1.109*</td>
<td>1.001**</td>
<td>1.000**</td>
<td>1.000**</td>
<td>1.000**</td>
<td>.000**</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>.010</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Time Since First Review</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.407</td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.065</td>
</tr>
<tr>
<td>Reviewer Var. (_cons)</td>
<td>.300</td>
<td>.336</td>
<td>.348</td>
<td>.404</td>
<td>.348</td>
<td>.363</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>28,200</td>
<td>50,254</td>
<td>29,651</td>
<td>151,996</td>
<td>29,651</td>
<td>29,651</td>
<td>22,020</td>
</tr>
<tr>
<td>Groups</td>
<td>1,594</td>
<td>1,595</td>
<td>1,594</td>
<td>1,594</td>
<td>1,594</td>
<td>1,594</td>
<td>1,191</td>
</tr>
<tr>
<td>Wald χ²</td>
<td>4,209.13</td>
<td>6,996.59</td>
<td>4,266.61</td>
<td>25,610.52</td>
<td>4,279.48</td>
<td>3,748.22</td>
<td>2,835.71</td>
</tr>
</tbody>
</table>

**p<.01; *p<.05. Business Type, State, and Year dummies included
^ GEE Population Average
### APPENDIX C: MIXED EFFECTS REGRESSION, REVIEWS TEXTS

**DV= Net Positivity Emotional Valence**

<table>
<thead>
<tr>
<th></th>
<th>$b$</th>
<th>$p$</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Previous Reviews</td>
<td>-.022*</td>
<td>.000</td>
<td>-.031 -.012</td>
</tr>
<tr>
<td>Stars</td>
<td>1.166*</td>
<td>.000</td>
<td>1.097 1.234</td>
</tr>
<tr>
<td>Average Ratings_Bus</td>
<td>.055</td>
<td>.187</td>
<td>-.027 .136</td>
</tr>
<tr>
<td>Number of Prev. Ratings_Bus</td>
<td>.000</td>
<td>.979</td>
<td>-.001 .001</td>
</tr>
<tr>
<td>Business Type included</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>4.192</td>
<td>.000</td>
<td>2.651 5.733</td>
</tr>
</tbody>
</table>

| Observations                         | 29,548  |
| Groups                                | 1,594   |
| Wald $\chi^2$                         | 21,054.60 |

* $p<.01$