Management Science

Motivating User-Generated Content with Performance Feedback: Evidence from Randomized Field Experiments

Ni Huang, Gordon Burtch, Bin Gu, Yili Hong, Chen Liang, Kanliang Wang, Dongpu Fu, Bo Yang

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management
science, and analytics.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

informs.

Motivating User-Generated Content with Performance Feedback: Evidence from Randomized Field Experiments

Ni Huang,[a] Gordon Burtch,[b] Bin Gu,[a] Yili Hong,[a] Chen Liang,[a] Kanliang Wang,[c] Dongpu Fu,[d] Bo Yang[e]

[a] W. P. Carey School of Business, Arizona State University, Tempe, Arizona 85287; [b] Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455; [c] School of Business, Renmin University of China, Beijing 100872, China; [d] School of Information, Capital University of Economics and Business, Beijing 100070, China; [e] School of Information, Renmin University of China, Beijing 100872, China
Contact: ni.huang@asu.edu (NH); gburtch@umn.edu (GB); bin.gu@asu.edu (BG); hong@asu.edu, http://orcid.org/0000-0002-0577-7877 (YH); chen.liang.4@asu.edu (CL); klwang@ruc.edu.cn (KW); fudongpu@cueb.edu.cn (DF); yangbo@ruc.edu.cn (BY)

Abstract. We design a series of online performance feedback interventions that aim to motivate the production of user-generated content (UGC). Drawing on social value orientation (SVO) theory, we develop a novel set of alternative feedback message framings, aligned with cooperation (e.g., your content benefited others), individualism (e.g., your content was of high quality), and competition (e.g., your content was better than others). We hypothesize how gender (a proxy for SVO) moderates response to each framing, and we report on two randomized experiments, one in partnership with a mobile-app–based recipe crowdsourcing platform, and a follow-up experiment on Amazon Mechanical Turk involving an ideation task. We find evidence that cooperatively framed feedback is most effective for motivating female subjects, whereas competitively framed feedback is most effective at motivating male subjects. Our work contributes to the literatures on performance feedback and UGC production by introducing cooperative performance feedback as a theoretically motivated, novel intervention that speaks directly to users' altruistic intent in a variety of task settings. Our work also contributes to the message-framing literature in considering competition as a novel addition to the altruism–egoism dichotomy oft explored in public good settings.

1. Introduction

User-generated content (UGC) is an important aspect of the Internet, influencing individuals' online behaviors in a variety of ways. UGC informs purchases (Chen et al. 2011), aids investment decisions (Park et al. 2014), provides entertainment (Leung 2009), and helps firms gather customer intelligence (Lee and Bradlow 2011). Indeed, the demand for UGC seems to be at an all-time high—recent reports note that Facebook's 1.4+ billion users spend an average of 20 minutes per day browsing peer-generated content on the site,[1] and YouTube, which now boasts more than 1 billion users, claims the average mobile viewing session extends more than 40 minutes in duration.[2] However, despite its value, UGC often suffers from an underprovisioning problem (Burtch et al. 2017, Chen et al. 2010, Kraut et al. 2012), in large part because it is a public good (Samuelson 1954); UGC is typically supplied on a voluntary basis, and its value is difficult for the contributor to internalize. This issue is exacerbated for new online communities or platforms, where a dearth of existing content can lead a user to perceive that there is no audience to recognize or benefit from his or her contributions (Zhang and

Zhu 2011). This underprovisioning problem has been widely recognized in both practice (McConnell and Huba 2006, Pew Research Center 2010) and academic work (Gilbert 2013, Burtch et al. 2017, Goes et al. 2016). It has even been reported that a mere 1% of the people who consume UGC also actively contribute it—known as the "1% rule" (van Mierlo 2014).

Motivating users to contribute content is thus an issue of prime importance for many online platforms. Unsurprisingly, a number of platforms have been experimenting with different types of interventions— e.g., Dangdang and Rakuten pay consumers reward points to write online product reviews; Stack Overflow provides users with badges to recognize contributions and activity; and many online communities, more generally, provide users with features that enable them to craft online reputation and social image (Ma and Agarwal 2007). One intervention that has yet to receive significant consideration in the information systems (IS) literature is the provision of performance feedback (Moon and Sproull 2008, Jabr et al. 2014, Kraut et al. 2012). Many online platforms regularly provide feedback to users about the popularity of

the content they have supplied. As a few examples: LinkedIn reports the relative ranking of a user's profile in terms of viewership,[3] MakerBot Thingiverse informs users about the number of times their 3D printing designs have been downloaded by others, TripAdvisor sends users monthly updates about the total readership of users' past restaurant reviews, and scholarly websites like Academia.edu and ResearchGate inform researchers about their manuscript downloads.

In this paper, we explore the design of feedback interventions with an eye toward alternative framings that speak to a user's motivations, and we evaluate the relative impact of these alternatives on users' subsequent production of content. Specifically, we address the following research questions: *What is the relative effectiveness of alternative performance feedback message framings that speak to different motives for UGC production? How can we best target a user based on his or her characteristics, social context, and performance level?*

We consider the mix of motivations that have been noted as drivers of UGC production in the prior literature—namely, the contrast between reputation building and status seeking versus benevolence. We argue that it is prudent to supply feedback about prior contributions in a framing that aligns with the individual's self-view, in terms of his or her social value orientation (SVO). In social psychology, SVO refers to the relative "weights" that an individual may place on others' welfare versus his or her own (Fiedler et al. 2013, Van Lange 1999)—that is, the idea that individuals' preferences for combinations of outcomes, as they relate to the benefits derived by the self and others, may vary. The SVO implies a three-category typology of orientations, including (a) "cooperative" (maximizing gains of the self and others), (b) "individualist" (maximizing gains of only the self), and (c) "competitive" (maximizing gains of the self, relative to others). In our context, if an individual is largely cooperative in nature, it is more effective to supply performance feedback in terms of the benefits others have derived from his or her recent contributions. Alternatively, if an individual is self-focused, it makes sense to supply feedback specifically about his or her own performance; and if an individual is highly competitive, it is optimal to supply feedback about his or her performance relative to other users. Ample literature suggests that SVO is highly correlated with gender, with males exhibiting greater competitive tendencies (Gupta et al. 2013, Croson and Gneezy 2009) and females exhibiting greater cooperative tendencies (Stockard et al. 1988, Fabes and Eisenberg 1998). We therefore leverage gender as a reliable proxy for SVO in our analyses.

We addressed our research questions via a randomized field experiment, performed in partnership with a large mobile-app–based recipe crowdsourcing application in China. We randomly varied the framing

of performance feedback messages, which were delivered to users of the platform via mobile push notifications. Subjects were randomly assigned to one of four conditions: cooperative (e.g., "you inspired x other users"), individualist (e.g., "you are in the top x%"), competitive (e.g., "you beat $1 - x$% other users"), and control (i.e., a generic message with no performance feedback). Notifications were issued every Saturday over the course of a seven-week period. Each users' subsequent content contributions were then observed and recorded.

We find a series of interesting results. First, we observe significant heterogeneity in the effects of feedback information delivered under alternative framings; gender significantly moderates the effect of the feedback framing on the volume and quality of content production. Specifically, the cooperatively framed feedback is significantly more effective at motivating female users to produce more content, whereas competitively framed feedback is significantly more effective at motivating male users. Both genders responded positively to individualist-framed feedback, with no significant differences between the two. Second, we observe heterogeneity in the response to feedback under each type of framing over the level of performance information delivered. We find that feedback information indicating that users have low performance may be demotivating, whereas feedback information indicating that users have high or average performance is typically motivating.

This work makes a variety of important contributions to the academic literature and to practice. First, we contribute to the IS and marketing literatures on UGC production, as well as the literature on performance feedback, by introducing a novel theory-informed performance feedback design, which speaks to the altruistic motivation of workers, not only in volunteer contexts (including UGC production), but also in paid scenarios. Second, we demonstrate the potentially important role of social context as a moderator of competitively oriented feedback. In the absence of publicity/social interaction, we find no evidence that competitively framed feedback influences individual productivity. Third, we offer a first consideration of the relative effectiveness of alternative performance feedback framings, deriving from a theoretical framework pertaining to other-regarding social preferences. From a more practical perspective, we demonstrate the value of targeting users in the design and delivery of personalized communications that are aimed at facilitating UGC production, accounting for heterogeneity in user motivations, social context, and performance levels.

2. Theory and Hypotheses
In this section, we survey the literatures on user-generated content, performance feedback, and social value

orientation theory. Integrating these three literature streams, we propose a set of testable hypotheses that guide our empirical examination.

2.1. User-Generated Content

There is a long stream of literature on user-generated content (UGC) with three major themes: effects of UGC, antecedents of UGC characteristics, and UGC production. First, ample research has examined UGC's effects on, for example, consumer decision making (e.g., Chen et al. 2011), product sales (e.g., Zhu and Zhang 2010), and firm competition (e.g., Kwark et al. 2014). Second, scholars have examined the antecedents of UGC's characteristics, such as rating, content length, and linguistic features. Such work has identified multiple important factors that influence these characteristics, including the contributor's popularity (Goes et al. 2014) and cultural background (Hong et al. 2016), the activity of a contributor's social network neighbors (Zeng and Wei 2013), and the timing and sequence of UGC contributions (Godes and Silva 2012, Muchnik et al. 2013). In this paper, we seek to build on and contribute to the third major stream of literature in UGC: the antecedents of UGC production.

The literature pertaining to UGC production focuses primarily on ways that firms can foster sustained participation (Trusov et al. 2010, Zhang et al. 2013), identifying a series of motivational factors, ranging from audience size (Zhang and Zhu 2011), social capital (Kankanhalli et al. 2005, Ma and Agarwal 2007, Wasko and Faraj 2005), and network position (Zhang and Wang 2012), to civil voluntarism (Phang et al. 2015) and community commitment (Bateman et al. 2011). Other work, building on these studies, has explored the design of interventions and system features that can be used to stimulate UGC production, such as the use of financial incentives. For example, Cabral and Li (2015) report on an experiment that studies the effectiveness of this approach. Chen et al. (2010) experiment with providing users information about peers' contribution activity, and Burtch et al. (2017) compare, contrast, and combine payments and normative information to measure their relative effectiveness when it comes to the quantity and quality of elicited contributions. Other recent work has explored the effects of "calls to action" and the importance of starting small and then gradually building those calls toward more effort-intensive contributions (Oestreicher-Singer and Zalmanson 2013). Other work, employing surveys and archival data, has examined the benefits of delivering performance feedback in an online question-and-answer (Q&A) community (Moon and Sproull 2008, Jabr et al. 2014), providing evidence that such feedback can drive content contributions.

Our work aims to build on the latter work, squarely focusing on how a novel system design—the delivery of performance feedback—can affect users' content production. Moreover, we seek to understand how such a design can be optimized, by using different framings of the feedback message and further targeting performance feedback based on user characteristics, social context, and performance levels.

2.2. Performance Feedback and Message Framing

More than 100 years of research has explored the use of performance feedback interventions to motivate learning and effort in work and educational settings (e.g., Kluger and DeNisi 1996, Hattie and Timperley 2007, Jung et al. 2010). Despite the broad assumption in much of the literature that performance feedback interventions are beneficial, later reviews of the lengthy literature document that the impacts have been decidedly mixed with more than one-third of studies reporting negative outcomes (Kluger and DeNisi 1996). This observation has led to the conclusion that the impacts of performance feedback are nuanced and highly contextual. Specifically, the effects of feedback appear to depend on how feedback is provided (Jaworski and Kohli 1991), personal traits of the recipient (Srivastava and Rangarajan 2008, Wozniak 2012), and whether the feedback is presented positively versus negatively (Hossain and List 2012, McFarland and Miller 1994).

We are aware of just two studies about the impact of performance feedback mechanisms on UGC production in the IS literature (Moon and Sproull 2008, Jabr et al. 2014). Jabr et al. (2014) report on an observational study of online Q&A communities, finding that the impacts of feedback mechanisms are heterogeneous, depending heavily on the solution-provider's preferences for peer recognition and social exposure. Moon and Sproull (2008) also study online Q&A communities, and observe that the presence of a performance feedback mechanism positively relates with content production, as well as user tenure in the community.

We build on the findings of these prior studies to understand how performance feedback can be designed and delivered to the greatest effect, by exploring the potential improvements that may be achieved through novel message framing, and through the alignment of framings with individual's characteristics for participation. A few studies have explored message framing in performance feedback previously, though never in the context of UGC production, and never beyond a consideration of gain versus loss or positive versus negative framings (e.g., McFarland and Miller 1994); thus, prior work in this domain has yet to consider framings that speak directly to alternative user motivations.

However, motivation-based framings have been explored in other, related contexts. In particular, scholars have considered the distinction between altruistic and egoistic message frames in solicitations for charitable

donation. Charitable donations, like UGC production, constitute private contributions toward a public good. Scholars have observed that altruistic framings are generally more effective at eliciting donations to charities than egoistic framings (Fisher et al. 2008), particularly when donations are made publicly (White and Peloza 2009). Moreover, other work has shown that the effects of these framings vary with subject gender; females respond more to an altruistic, help-other framing, whereas males respond more to a help-self framing (Brunel and Nelson 2000).

Although charitable donations and UGC production are similar in some respects, they differ substantially in others, which leaves open the possibility that prior results may fail to generalize. Most fundamentally, a charitable donation is a request for money, where the donor is invited to provide funds with the promise of having a prospective impact by enabling the actions of the receiving organization. Performance feedback, in contrast, is retrospective in nature, communicating to a worker the results of his or her own past efforts. This distinction between prospective versus retrospective outcomes implies a significant potential for framing effects to differ across contexts.

Various studies document gender differences in relation to task performance and evaluations of said. As examples, Baer (1997) demonstrates gender heterogeneity in response to the anticipation that others will evaluate the subject's work in creativity tasks, with females responding negatively and males exhibiting no response at all, and Beyer (1990) notes that each gender exhibits a tendency to hold elevated expectations of self-performance in tasks associated with respective gender roles. Thus, we might expect, for example, that depending on a task, and its gender association, males and females would bear systematically different self-performance expectations, and that the receipt of objective performance feedback might then systematically exceed or fall short of said expectations, leading to gender heterogeneity in the impacts of feedback. These notions bear no analog in the charitable donation setting, given that it involves no notion of task performance.

Nonetheless, we consider the potential effects of similar motivation-based framing alternatives in the delivery of performance feedback about users' online content contributions. We extend on the typically considered dichotomy (altruistic versus egoistic), drawing on SVO theory to motivate a trichotomy of message framings, ranging from cooperative (altruistic), to individualistic (egoistic), to competitive. In integrating these disparate literatures, our work contributes first by considering the notion of a competitively oriented message framing in a public good setting, as well as by considering the notion of an altruistically oriented message framing in a voluntary or paid work setting.

2.3. Social Value Orientation Theory and Gender Differences

Social value orientation (SVO) speaks to the relative "weights" that an individual will place on his or her own welfare, and that of others (Fiedler et al. 2013, Van Lange 1999). SVO theory holds that individuals maintain heterogeneous preferences for combinations of outcomes as they relate to the benefits derived by the self and others. Because individuals are assumed to always be self-interested to some degree, this heterogeneity results in a three-category typology of individuals as inherently (a) "cooperative" (maximizing others' gains, in addition to one's own), (b) "individualist" (maximizing one's own gains, indifference with respect to others' gains), and (c) "competitive" (maximizing one's own gains, relative to or at the expense of others). Thus, a cooperative orientation refers to an individual's joint maximization of his or her own payoffs and those of others; an "individualist" orientation refers to an individual's maximization of his or her own payoff, without consideration to the payoff of others (Fiedler et al. 2013, Liebrand and McClintock 1988); and a competitive orientation refers to an individual's maximization of his or her own payoff relative to that of others' (Fiedler et al. 2013) Figure 1 provides a graphical depiction of the two dimensions of SVO, with self-regarding preferences on the x axis, other regarding preferences on the y axis, and the above trichotomy bolded.

The notion of SVO aligns well with our research context because it speaks, at one end of the spectrum, to individuals' motives for contributing to the public good and helping others (i.e., cooperative orientation), and at the other end of the spectrum, to individuals' desire to build image and reputation (Jabr et al. 2014, Toubia and Stephen 2013). That is, on the one hand, contributing UGC can benefit the collective by providing more content for others to consume, yet on the other hand, individuals may also obtain image-related utility if they believe they have attracted a greater share of peers' attention (Ahn et al. 2015, Iyer and Katona 2015, Toubia and Stephen 2013, Zhang and Sarvary 2014).

Figure 1. Social Value Orientations (Fiedler et al. 2013, Van Lange 1999)
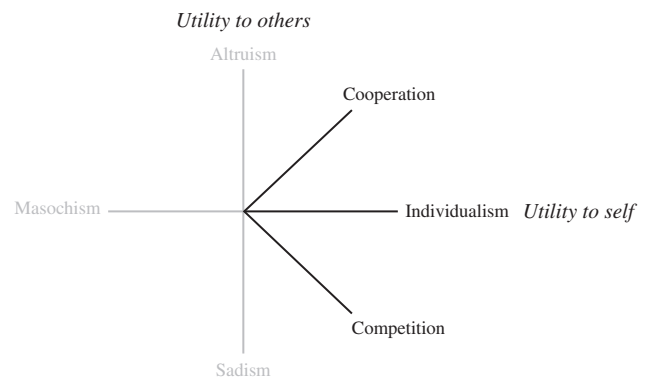
In the context of online UGC contributions, cooperative, altruistic motivations can play a particularly strong role in driving user behaviors (Zhang and Zhu 2011). A cooperative message framing should therefore provide a strong confirmation of contributors' self-view as being altruistic. This argument is consistent with the above-noted finding that altruistic message framings are generally more effective than egoistic framings when it comes to charitable contributions (Fisher et al. 2008, White and Peloza 2009). Given the similarly public good nature of UGC production, individuals who select into UGC production may be more likely to view themselves as altruists. Self-verification theory (Swann 2012, Swann and Read 1981) suggests that positive, altruistically framed feedback would strengthen contributors' motivation to take actions to sustain their self-view, which, in our context, is achieved by continuing and strengthening their UGC contribution.

At the same time, previous literature also suggests that another major motivation for UGC contribution is reputation and social recognition (Wasko and Faraj 2005, Toubia and Stephen 2013, Zhang and Zhu 2011). Such motivations are self-oriented (Roberts et al. 2006, Toubia and Stephen 2013) as they provide image-based utility to users and, as a result, individualist-framed performance feedback helps sustain contributors' self-view in this regard. As we consider that no prior literature suggests that competition is a motivating factor in UGC contribution, we have reason to believe that, broadly speaking, this framing should prove relatively weak, especially in the context of public good contribution.

We must recognize that the effect of each framing depends on the alignment between the framing and individual users' SVO—thus, evaluation of the effects of each framing would require consideration of an individual user's SVO. Although we do not directly observe SVO, the literature notes that SVOs are highly correlated with gender. First, in both economics and psychology, extensive work indicates that females are more likely to be altruistic than males. Research in child psychology notes that females are more likely to engage in prosocial behaviors than males (Fabes and Eisenberg 1998). Work in philanthropy notes that women are more likely to give money to charities and that 90% of women give more to charities than the average man (Piper and Schnepf 2008). The 2002 General Social Survey tells a similar story, indicating that women score systematically higher on altruistic values, altruistic behaviors, and empathy. The latter observation, that women tend to feel more empathy, has also been confirmed in a variety of other studies in social psychology (Baron-Cohen and Wheelwright 2004, Eisenberg and Lennon 1983).

Past research has also established that females exhibit greater sensitivity to social cues and others' moods and affect (Piliavin and Charng 1990). Because females are more socially attuned, they are more likely to notice others' unfavorable circumstances (Stocks et al. 2009) and thus are more likely to respond to others' needs. Consequently, females tend to be more cooperative than males (Stockard et al. 1988). Studies in experimental economics have also found, repeatedly, that women are more inequality averse in dictator games when in the deciding role, opting for a more equal (fair) distribution of capital than men (Andreoni and Vesterlund 2001, Dickinson and Tiefenthaler 2002, Rand et al. 2016, Dufwenberg and Muren 2006). Perhaps most importantly, past studies of the effect of message framing in charitable solicitation have found evidence that females respond more strongly to altruistic framings than egoistic framings (Brunel and Nelson 2000). Considered collectively, the above leads us to the following hypothesis:

Hypothesis 1 (H1). *Cooperatively framed performance feedback will have a stronger effect on female users than on male users.*

Second, the literature suggests that males are more likely to be individualist or egoist (Van Lange et al. 1997, Weber et al. 2004). According to the gender self-schema theory (Ruble et al. 2006), males are more prone to exhibit a strong individualist orientation and are more responsive to individualist feedback (Gordon et al. 2000). More recent work has found the converse for women, such that women are more likely to exhibit an aversion to standing out from the crowd in their altruistic activities (Jones and Linardi 2014)—thus, they may be less interested in building reputation or image. Finally, if we once again consider past work on the effects of cooperative (altruistic) and individualist (egoistic) message framings in the solicitation of charitable donations, it has been found that males are more likely to respond to egoistic message framings (Brunel and Nelson 2000). This leads to our next hypothesis:

Hypothesis 2 (H2). *Individualist-framed performance feedback will have a stronger effect on male users than on female users.*

Third, and last, past work has also consistently demonstrated that males are more likely to be competitive (Gupta et al. 2013, Croson and Gneezy 2009). Studies suggest that males respond more positively to competition than females (Croson and Gneezy 2009, Gneezy et al. 2003, Niederle and Vesterlund 2007), a result that can be explained, at least in part, by the fact that men are more likely to be overconfident (Beyer 1990, De Paola et al. 2014) and to focus primarily on success and pay less attention to failure (De Paola et al. 2014). It has therefore been found that competition increases the performance of men, but not of women (Gneezy

et al. 2003, Morin 2015, Shurchkov 2012). Given prior findings that males are more likely to opt into a competitive setting, it is likely that males are more likely to view themselves as competitive or competitors. With the above in mind, we propose our final hypothesis:

Hypothesis 3 (H3). *Competitively framed performance feedback will have a stronger effect on male users than on female users.*

3. Mobile Field Experiment
3.1. Experimental Design
For our field experiment, executed in collaboration with one of the largest mobile recipe-sharing applications in China (www.Meishi.cc, hereafter referred to as our corporate partner), our experimental treatments were delivered to subjects via mobile push notifications. Push notifications are commonly used by smartphone-application operators to deliver messages to the home screen of users' smartphones. Using push notifications to deliver the treatments has multiple advantages over other types of digital treatment delivery methods (e.g., email or short messaging service (SMS) text). First, because of the large amount of junk and spam emails related to promotions, users tend to ignore such emails (Burtch et al. 2017). Second, push notifications are integrated with the mobile application, avoiding the concern that recipients may ignore SMS text messages, believing them to be spam.

Our push notifications were designed as an A/B test of the corporate partner's weekly notification system,[4] aimed at motivating users to engage with a new section of the mobile application. Our randomized experiment was the only experiment taking place during the duration of our study—i.e., the corporate partner was not running any other experiments in parallel. Users received push notifications (including the written message we designed) that would appear on the home or lock screen of their device. Once clicked or swiped, the user was taken to a landing page within the mobile application, and the same short message was shown as a pop-up once again. The notifications created for our experimental treatments pertained to the application's new (at the time) "Shi Hua" (Foodie Talk) section. Foodie Talk is a social media component of the recipe application, implemented in the main mobile application interface (the fourth tab in Figure 2).

Within the Foodie Talk section, users can initiate posts related to food, cooking, or ideas for new recipes in the form of photos and text by tapping on the top left camera icon from the main application screen. The posts would become viewable as soon they are submitted, at which point other users could "like" and "comment" on them. For each individual post, its total comment and like volume are displayed in the bottom-right corner of the post, near the comment icon and the heart icon.
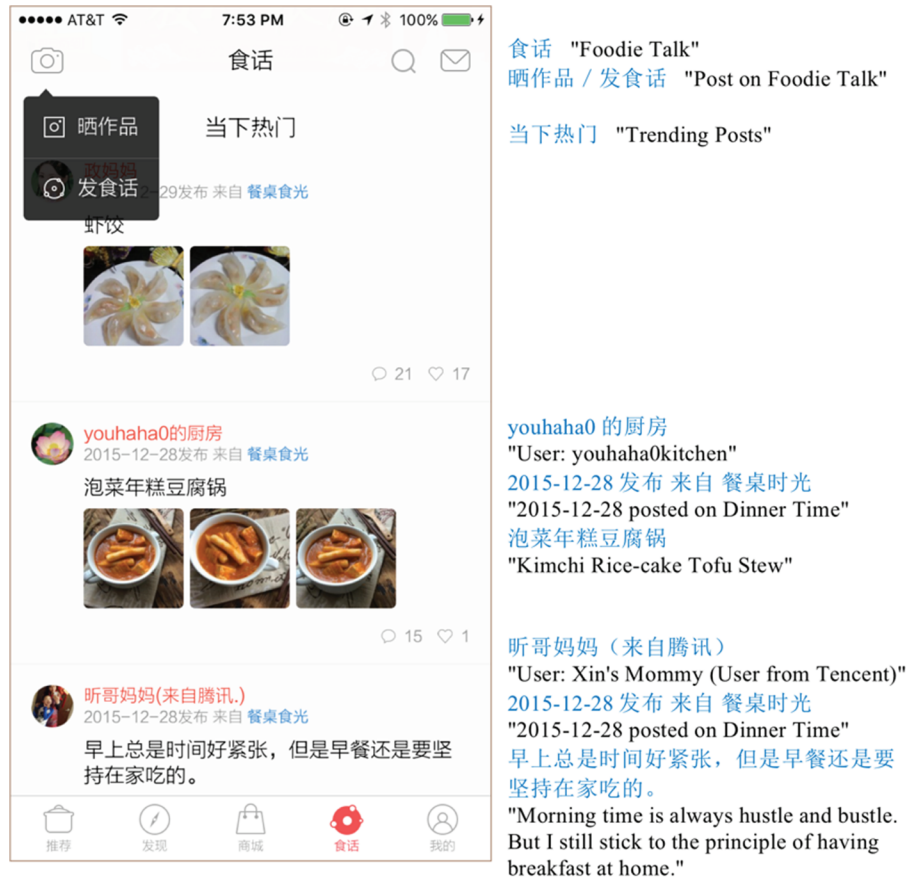
Our treatments were intended to increase users' posting volumes in the Foodie Talk section, because the corporate partner was concerned that content was initially quite scarce. The numerical values presented in the performance feedback message reflected peer engagement with a user's content, based on the total number of "likes" each user had received as of 11:59 P.M. on the day prior. We used real values when determining the content of these messages (except for users who had attracted zero likes, as detailed below), for two reasons. First, real values avoid issues related to cognitive dissonance (e.g., a very active user may be confused to find that he or she only had a very small number of likes). Second, the corporate partner was concerned about losing the users' trust, if they were to become aware that false information was being disseminated. For those users who had attracted zero likes up to the point of the first treatment, we randomly replaced the zero with an integer between 1 and 5 and updated the information on the users' profile page accordingly.

We first designed a control group notification, wherein assigned subjects were simply reminded to log in to the mobile application. We then designed three treatment group notifications, one for each SVO framing. Each user in the "cooperative" treatment group was informed about how many other users had benefited from his or her recent content postings. Each user in the "individualist" treatment group was informed about his or her percentile rank (%), in terms of the popularity of his or her recent postings (based on the aggregate count of likes across all prior posts). Finally, each user in the "competitive" treatment group was informed about the proportion of other users on the site that he or she had outperformed (% beaten), again based on the popularity of (total likes attracted by) his or her recent postings.

Figure 3 depicts a screenshot from one coauthor's mobile device, on receipt of test messages for all four treatment conditions. The text on the right provides the English translation of the message. A given user, depending on the treatment group to which they were randomly assigned at the outset of the study, would have received only one of these four message formats, each Saturday, for the duration of the study.

Assignment of subjects to treatment groups was performed one day prior to the first treatment delivery (GMT + 8 8 P.M. Beijing Time on November 7, 2015). We then observed each user weekly over the course of the seven-week study period. We worked directly with the IT and marketing departments of the corporate partner to develop a system routine for computing the number of likes and delivering the user-week–specific treatment stimuli. Current Foodie Talk users were randomly assigned using pseudo random number generators, using the approach of Deng and Graz (2002). The randomization procedure was integrated into an

Figure 2. (Color online) Foodie Talk Sample Page with Translations

食话　"Foodie Talk"
晒作品 / 发食话　"Post on Foodie Talk"

当下热门　"Trending Posts"

youhaha0 的厨房
"User: youhaha0kitchen"
2015-12-28 发布 来自 餐桌时光
"2015-12-28 posted on Dinner Time"
泡菜年糕豆腐锅
"Kimchi Rice-cake Tofu Stew"

昕哥妈妈（来自腾讯）
"User: Xin's Mommy (User from Tencent)"
2015-12-28 发布 来自 餐桌时光
"2015-12-28 posted on Dinner Time"
早上总是时间好紧张，但是早餐还是要
坚持在家吃的。
"Morning time is always hustle and bustle.
But I still stick to the principle of having
breakfast at home."

Tabs at the Bottom:
推荐　"Recommendations"; 发现　"Explorations"; 商城　"Shopping Mall";
食话　"Foodie Talk"; 我的　"My Account"

Source. Meishi application.

algorithm in the corporate partner's push notification delivery system. A total of 2,360 current users who have initiated at least one post on Foodie Talk entered the experiment. We limit our estimation sample to 1,129 users, the subset of users where we could reliably determine gender (730 users who self-reported gender information in their user profile, and an additional 399 users who did not report gender yet supplied a profile photo from which we were able to reliably determine gender using the Face++ API).[5] Randomization balance checks confirming the validity of the randomization procedure are presented in Online Appendix A. Notably, our estimation sample is roughly balanced across experiment conditions, supporting the validity of our randomization procedure—i.e., it does not appear that the probability of gender self-report or photo provision is correlated with treatment.

3.2. Interviews and Surveys to Validate the Stimuli

Prior to implementing this first experiment, we conducted interviews with five users of the application to validate the design of our treatment stimuli. Each of the interviewees was presented with all three treatment messages in sequence, and then asked a series of questions to gauge whether the messages made the them see themselves as cooperative, self-interested (i.e., personally successful), or competitive, and whether they felt each message would instigate a desire to contribute more to the platform.

All five interviews were conducted on WeChat, a popular social networking service (SNS) application in China. On viewing the cooperative treatment message, most of the interviewees reported that the message indicated that prior contributions were valued by other users and that they would feel a sense of appreciation. The interviewees also stated that receiving the message would have motivated them to contribute more content. When the interviewees viewed the individualist treatment message, all three male respondents felt the ranking information conveyed information about a user's own past performance, and that it would likely cause them to engage in self-comparison, stimulating them to strive for better performance in the future, though one interviewee commented that

Figure 3. (Color online) Treatment Messages and English Translation

"Dear, come visit Meishijie. All kinds of yummy food recipes are waiting for you."

"Dear, your Foodie Talk posts have been 'liked' 49 times. You have provided cooking inspirations for 48 other users on Meishijie!"

"Dear, your Foodie Talk posts have been 'liked' 95 times. You are ranked top 3% among all foodies on Meishijie."

"Dear, your Foodie Talk posts have been 'liked' 122 times. You beat 98% other foodies on Meishijie."

Source. Based on test push notifications from Meishi application to an iPhone.

their response might depend on the actual level of past performance. In contrast, neither female interviewee expressed a strong interest in the individualist framing. For the competitive treatment message, one male interviewee reported that the "you have beaten" wording provided him with a sense of pride and a strong stimulus to compete with other contributors. The other two male interviewees stated explicitly that the competitiveness framing would motivate them to contribute more content. In contrast, the female interviewees reported a combination of unease and disinterest with the competitive framing. Lastly, no interviewees were impressed by or interested in the control group treatment message. Overall, the preexperiment interviews indicated that the framing of our treatment messages was well aligned with our objectives, and they provided preliminary evidence in support of our expectations; that message framing is important, as is alignment between framing and user preferences.

We also validated our treatment designs with a sample of 97 users using a survey. We began the survey with a preamble that informed the user (respondent) that the mobile app operator had calculated and provided to us the number of likes accrued by his or her content posted to the app over the prior seven days. We then asked each user to indicate his or her reactions

to the messages constructed for our four experimental conditions (i.e., the messages presented in Figure 3) on five-point Likert-type scales. The respondents were asked to indicate their agreement with the following statements, after considering each type of treatment message: (a) "I feel that my posts have helped other users," (b) "I feel that I am a real foodie," and (c) "I feel that I am a better foodie than other users." We observe significant differences in users' feelings toward the different treatment messages, such that the cooperative message received the highest average response on item (a), the individualistic message received the highest average response on item (b), and the competitive message received the highest average response on item (c). Notably, the control message received the lowest average response across on all three items. Overall, the survey results suggest that the stimuli deliver the desired manipulations. More detail on this survey is provided in Online Appendix B.

3.3. Data Description and Empirical Approach
The key measures for Experiment 1 are described in Table 1, and descriptive statistics are presented in Table 2. Consistent with our weekly treatment design,[6] we begin by estimating content volume regressions on our user-week panel, wherein we evaluate the

Table 1. Variable Definitions

Variable	Definition
PostVolume	Total number of postings by user j, in week t
Likes	Total likes of user j's posts made in week t, over the life of the posts
Comments	Total comments on user j's posts made in week t, over the life of the posts
Readership	Total unique visitors to user j's posts made in week t, over the life of the posts
Preperformance	User j's performance score as of the start of the experiment (this value determines initial treatment message information)
Gender	Female = 0; Male = 1

interaction between each treatment and a gender indicator. Our log-OLS estimation incorporates time (*Week*) fixed effects. Our initial regression specification for the user-level panel is thus as per Equation (1), where j indexes each user and t indexes weeks:

$$\mathrm{Ln}(PostVolume_{jt}) = Treat_j + Gender_j + Treat_j \times Gender_j + Week_t + \epsilon_{jt}. \quad (1)$$

We subsequently estimate a set of content quality regressions, wherein we replace our dependent variable (DV) with three quality measures, *Readership*, *Comments*, and *Likes*. Content quality has been studied in recent IS research (Qiu et al. 2016). Here, we assume that higher-quality postings would attract greater readership, more comments, and more likes from other users. We anticipate that performance feedback interventions can induce higher content quality, in addition to higher content volumes, because feedback helps users to assess progress in their pursuit of goals or objectives (Kraut et al. 2012).

As with the other variables in our main analyses reported above, we constructed these three new outcome variables at the user-week level. For each user-week observation, we identified all content contributions the user made in that period. We then calculated the total value of the measure, summing values from that day through to the end of our study period. For example, if a user-week observation had a readership value of 15 in our data, this would indicate that all postings created by that user, in that week, eventually accrued 15 unique viewers over the course of the remaining experiment period.

Table 2. Descriptive Statistics

Variable	Mean	Std. dev.	Min	Max	N
1. *PostVolume*	0.39	2.69	0.00	41.00	7,903
2. *Likes*	0.26	3.64	0.00	186.00	7,903
3. *Comments*	0.80	6.32	0.00	268.00	7,903
4. *Readership*	0.11	1.79	0.00	53.00	7,903
5. *Preperformance*	25.21	108.00	1.00	1,714.00	7,903
6. *Gender*	0.30	0.46	0.00	1.00	7,903

We estimated the regression specification reflected by Equation (2), where ρ indexes our three quality measures. Note that in addition to our other variables, we control for the number of contributions the user had made in the week, to ensure that our estimates reflect per-post effects. Note as well that content age is accounted for by the *Week* fixed effects. Once again, j indexes each user and t indexes weeks:

$$\mathrm{Ln}(PostQuality_{jt}^{\rho}) = Treat_j + Gender_j + Treat_j \times Gender_j + Controls_j + Week_t + PostVolume_{jt} + \epsilon_{jt}. \quad (2)$$

In addition to our volume and quality regressions, we also consider heterogeneity in the treatment effects over the subject performance level distribution. We construct tercile dummies reflecting whether a subject's performance level falls within the lower, middle, or top third of the performance distribution. We then interact dummies reflecting the second and third tercile with our treatment indicators to assess variation in response across the performance distribution. Moreover, we explore the nature of this heterogeneity across gender-specific subsamples.

3.4. Volume Analysis

As noted above, we observe gender for 1,129 users—thus, our estimation results pertain to those users. The results of our volume regressions are presented in Table 3. First, we observe that, on average, males contribute less content than females. Although this result may be contextual, it indicates that females, who we expect to be more cooperatively oriented, tend to be more willing to contribute UGC, and thus to the public good.

Turning to our hypotheses, we find broad support consistent with our expectations. Compared to male

Table 3. Regression Results—Gender Effects (DV = Ln(*PostVolume*))

Explanatory variable	(1)		(2)	
Cooperative	0.045***	(0.006)	0.056***	(0.005)
Individualist	0.048***	(0.007)	0.045***	(0.010)
Competitive	0.018**	(0.006)	−0.017**	(0.005)
Cooperative × Gender			−0.034**	(0.010)
Individualist × Gender			0.010	(0.014)
Competitive × Gender			0.121***	(0.011)
Gender	−0.025***	(0.006)	−0.049***	(0.004)
Constant	0.077***	(0.005)	0.084***	(0.005)
Week FE	Yes		Yes	
Observations	7,903		7,903	
F-statistic	24.10***		76.59***	
R^2	0.003		0.007	
Number of users	1,129		1,129	
Number of weeks	7		7	

Note. Cluster-robust standard errors in parentheses.
$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

users, we find that female users respond more strongly to the cooperative treatment. In contrast, compared with female users, male users respond more strongly to the competitive treatment. These two findings provide support for H1 and H3, respectively.

However, we do not find evidence in support of H2. That is, we find no statistically significant difference between males and females regarding the effect of the individualist treatment. This may reflect the fact that the individualist treatment is both "more" cooperative than the competitive treatment and "more" competitive than the cooperative treatment, in the sense that individuals who are individualist do not seek to maximize their own gains at the expense of others, nor do they seek to maximize their own gains in tandem with others. Thus, if we view the treatments, from competitive, to individualist, to cooperative, as a sliding scale of increasing (decreasing) cooperation (competition), our results can be readily rationalized.

We also explored the robustness of our estimates to alternative model specifications given the serially correlated outcomes (Bertrand et al. 2004). First, we evaluated the robustness of our estimates to a pooled estimation, collapsing our panel data to the user level. These results are reported in Table 4, where we see generally consistent results, particularly around gender heterogeneity in the cooperative and competitive treatments (column (2)). The one exception in this case is that we observe no significant effect of the individualist treatment, for either gender. Finally, we also evaluated robustness of our results to a linear OLS specification (i.e., removing the log transformations), and again observed similar results.

Our estimates are also not just statistically significant; they are economically significant. Using the weekly estimations, for the cooperative treatment, a female user produces 5.76% more posting each month than if she was to receive a generic push notification (versus 2.22% more postings each month for a male

Table 4. Regression Results—Pooled Estimation (DV = Ln(*PostVolume*))

Explanatory variable	(1)		(2)	
Cooperative	0.253**	(0.099)	0.370***	(0.127)
Individualist	0.167*	(0.098)	0.088	(0.122)
Competitive	0.147	(0.096)	0.022	(0.112)
Cooperative × Gender			−0.378**	(0.189)
Individualist × Gender			0.300	(0.203)
Competitive × Gender			0.431**	(0.212)
Gender	−0.203***	(0.076)	−0.280**	(0.126)
Constant	0.801***	(0.072)	0.825***	(0.084)
Observations	1,129		1,129	
F-statistic	3.77***		4.91***	
R^2	0.012		0.026	

Note. Robust standard errors in parentheses.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

user). In contrast, under the competitive treatment, each male user produces 10.96% more posting than if he were to receive a generic push notification (versus 1.69% fewer postings each month for a female user). This suggests very large differences in the volume of content production, if we extrapolate to the platform level.

3.5. Heterogeneous Treatment Effects

Because our subjects naturally fall at different points in the performance distribution, one question that arises is whether the effect of treatments varies across performance levels, resulting in heterogeneous responses, and whether this heterogeneity also manifests asymmetrically across the two genders. To evaluate this possibility, we took an approach akin to that of Chen et al. (2010), interacting tercile dummies capturing which baseline performance group a subject fell into at the outset of the study (based on the overall distribution of likes that users' content had as of the first day of the study) with each treatment indicator. These results are presented in Table 5, where we first report heterogeneity across the whole sample (column (1)), and then again for each gender (columns (2) and (3)).

We observe differences across each treatment framing across performance tiers, in terms of whether it has a motivating or demotivating effect on the subject, and in some cases, this also varies across genders. First, we observe relatively homogeneous results when it comes to the cooperative treatment. The treatment does not appear to demotivate users in the bottom or middle performance tier, regardless of gender, yet it also delivers no motivating benefit. The entirety of the effect from this treatment appears to derive from individuals already operating at the top of the performance distribution, with female subjects responding relatively more strongly than males.

Second, considering the individualist treatment, we see that females are demotivated when informed of a very low score, whereas males are unresponsive. However, females also shift to a positive response in the middle tier, while males again remain unresponsive. Once again, in the top performance tercile, we see a uniformly positive response across both genders. Lastly, considering the competitive treatment, we observe that females are demotivated by a low score, yet apparently are never motivated by a high score. Rather, female subjects are merely unresponsive. As such, it appears that there is no upside to using a competitive treatment for female users, under any of the conditions we examine. Conversely, for males, we observe positive response at both tails of the performance distribution, and in the upper tail, the response is particularly magnified.

Our aggregate results in column (1) of Table 5 appear to align, broadly, with expectation confirmation theory (Oliver 1980). As users likely expect themselves

Table 5. Regression Results—Heterogeneous Treatment Effects (DV = Ln(*PostVolume*))

Explanatory variable	All		Male		Female	
Cooperative (1*st Tercile*)	−0.004	(0.010)	0.010	(0.014)	−0.001	(0.015)
Cooperative × 2*nd Tercile*	−0.004	(0.007)	−0.020	(0.026)	−0.006	(0.012)
Cooperative × 3*rd Tercile*	0.110***	(0.028)	0.089**	(0.026)	0.106**	(0.036)
Individualist (1*st Tercile*)	−0.024**	(0.008)	0.022	(0.014)	−0.039***	(0.009)
Individualist × 2*nd Tercile*	0.100***	(0.013)	0.010	(0.033)	0.141***	(0.012)
Individualist × 3*rd Tercile*	0.112**	(0.033)	0.069**	(0.026)	0.129**	(0.036)
Competitive (1*st Tercile*)	−0.020	(0.013)	0.032*	(0.015)	−0.039**	(0.013)
Competitive × 2*nd Tercile*	0.009	(0.015)	−0.030	(0.017)	0.018	(0.015)
Competitive × 3*rd Tercile*	0.071**	(0.030)	0.190***	(0.035)	0.023	(0.034)
2*nd Tercile*	−0.029**	(0.008)	0.002	(0.011)	−0.035**	(0.011)
3*rd Tercile*	0.101***	(0.025)	0.111***	(0.018)	0.098**	(0.029)
Constant	0.051***	(0.008)	0.008	(0.011)	0.066***	(0.009)
Week FE	Yes		Yes		Yes	
Observations	7,903		2,387		5,516	
F-statistic	308.24***		114.47***		68.55***	
R^2	0.046		0.081		0.042	
Number of weeks	7		7		7	

Note. Cluster-robust standard errors in parentheses.
$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

to perform at the average, when performance feedback indicates that the user has fallen short of that expectation, he or she react negatively. As past performance increases, however, users' expectations are instead realized or even surpassed, and their motivation is enhanced. Therefore, the results suggest a positive feedback loop. The story becomes more nuanced, however, when we look at the gender-specific results. In the two conditions that involve some indication of social comparison (individualist and competitive), males and females appear to respond quite differently.

As noted earlier, one might be concerned that in this analysis, a user's past performance (popularity of past content) is possibly endogenous with respect to his or her treatment response. That is, it may be the case that more-active users achieve better performance, and such users may also be more likely to respond to different feedback framings. To alleviate this concern, our experimental design offers us a convenient, additional source of identification. Recall that users who received zero likes prior to the first treatment were randomly assigned a value between 1 and 5, and that such users' performance was then determined based on this assigned value. Therefore, for the subsample in question, both the message framing and the performance information were exogenously assigned.

We repeat our analyses on this subset of users, with confidence in the exogeneity of all covariates. Because we do not observe the whole distribution of performance in this sample, we replace our tercile dummies with a continuous measure of preexperiment likes, and interact it with our treatment dummies. This subsample estimation is comprised of 378 female users and 176 male users ($N = 554$). As the estimation results in Table 6 show, the coefficient estimates remain relatively

consistent. We see that low values, on average, demotivate, and that this demotivation is attenuated as the performance level increases. Of course, this sample of users is not entirely representative, given that they uniformly fall in the bottom tail of the performance distribution, but the consistency of the findings nonetheless lends credibility to our main results.

3.6. Additional Quality Analysis

We next consider whether our treatments induce a shift in the average quality of content produced by subjects (in terms of popularity). As noted by prior work, the receipt of performance feedback can have a motivating effect on users (Kraut et al. 2012), which could lead them to exert more effort when producing content, as they seek to improve on past performance. The use of

Table 6. Regression Results—Heterogeneous Treatment Effects (DV = Ln(*PostVolume*))

Explanatory variable	(1)	
Cooperative	−0.017	(0.017)
Individualist	−0.118***	(0.021)
Competitive	−0.060**	(0.023)
Cooperative × # of Likes	0.007*	(0.004)
Individualist × # of Likes	0.055***	(0.007)
Competitive × # of Likes	0.016**	0.005)
# of Likes	−0.022***	(0.004)
Constant	0.096***	(0.015)
Week FE	Yes	
Observations	3,878	
F-statistic	78.39***	
R^2	0.017	
Number of weeks	7	

Note. Cluster-robust standard errors in parentheses.
$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 7. Regression Results—Quality Effects (DV = Ln(*PostQuality*))

Explanatory variable	Ln(*Likes*)		Ln(*Comments*)		Ln(*Readership*)	
Cooperative	0.015**	(0.005)	0.024**	(0.009)	0.000	(0.007)
Individualist	0.021*	(0.009)	0.038**	(0.011)	−0.007	(0.012)
Competitive	0.002	(0.006)	0.003	(0.006)	−0.013	(0.008)
Cooperative × Gender	0.017	(0.010)	0.001	(0.012)	−0.002	(0.008)
Individualist × Gender	0.007	(0.014)	−0.002	(0.022)	0.016	(0.017)
Competitive × Gender	0.019**	(0.006)	0.076***	(0.016)	0.014	(0.012)
Gender	−0.018***	(0.004)	−0.032***	(0.006)	−0.013	(0.007)
PostVolume	0.065***	(0.005)	0.154***	(0.005)	0.026***	(0.007)
Constant	0.019***	(0.004)	0.057***	(0.005)	0.012*	(0.006)
Week FE	Yes		Yes		Yes	
Observations	7,903		7,903		7,903	
F-statistic	58.95***		2,205.52***		6.65**	
R^2	0.312		0.579		0.114	
Number of users	1,129		1,129		1,129	
Number of weeks	7		7		7	

Note. Cluster-robust standard errors in parentheses.
  $^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

number of "likes" as our performance measure could therefore motivate users to improve on both quantity and quality to achieve greater performance. Prior work points to two possible explanations for the effect of performance feedback on quality—learning and motivation. Learning seems an implausible mechanism here, because our treatment messages characterize engagement across a subject's entire corpus of prior contributions, rather than pointing to a particularly high-quality posting. As such, we attribute these effects to changes in user motivation in expending more effort. Additionally, it may be the case that learning is inherently difficult in this setting, given the taste-based nature of the content being produced (whether food tastes "good" is inherently subjective to some degree).

The results of our quality regressions are presented in Table 7. We find evidence that our treatments have positive effects on average contribution quality. Interestingly, however, the effects of the treatments are not uniform in their gender heterogeneity. Whereas only males appear to produce higher-quality content under the competitive treatment (7.90% more comments and 1.92% more likes), our results indicate that both males and females produce higher-quality content under the cooperative treatment (2.43% more comments and 1.51% more likes). This is in line with our volume results, where we observed that both males and females produce significantly higher average volumes of content (albeit the female response is much larger). Indeed, estimating our volume regression on only the male population indicates that males respond significantly, and positively, under all three treatments ($p = 0.075$ in the cooperative case, $p = 0.001$ in the individualist case, and $p = 0.002$ in the competitive case).

It is also interesting to note that we observe effects only on comments and likes, and not readership.

On reflection, this is not entirely surprising, because readership is a precondition for other users to assess quality, and to then engage with a post—i.e., like or comment on it. Thus, other users do not become aware of content quality until they have consumed it.

3.7. Follow-Up Experiment (Amazon Mechanical Turk)

The results of the main study confirmed many of our expectations. We observed the anticipated gender heterogeneity around cooperative and competitive feedback framings. However, as noted in the introduction, we are also interested in understanding the potential moderating role of social context when it comes to the observed treatment effects. Moreover, there are potential questions of generalizability when it comes to our main findings. Generalizability might be a concern if the observed altruistic behavior is dependent on gender roles. A long-held stereotype is that cooking is a female-oriented task (Blair and Lichter 1991). It may be the case that females are more likely to engage in altruistic behavior when that behavior aligns with culturally assumed feminine gender roles. As such, we might be concerned that the greater positive response by females to the cooperation condition would depend on this feature of the study context. The Chinese context of the main field experiment may also pose a generalizability concern, as numerous prior studies speak to cultural differences in social behavior (e.g., Triandis 1989, 1994).

3.7.1. Experimental Design. We undertook a follow-up randomized experiment on Amazon Mechanical Turk (MTurk), to explore generalizability. Notably, we observe a predominantly male, North American subject pool (we report descriptive statistics below). Our experiment took the form of a crowdsourced ideation

task, wherein we initially solicited "one piece of advice for new users of Amazon Mechanical Turk" from 999 distinct members of the existing MTurk population. We compensated workers in this first stage at a rate of $0.40 USD per submission.

We then took the 999 pieces of advice from the first stage and invited an additional set of Turkers (namely, excluding individuals who had participated in the first stage) to evaluate each first-stage submission. Specifically, we asked five separate workers to indicate their perception of the usefulness of each piece of advice on a scale from 1 to 5, with 5 being most useful. The summed usefulness scores for each piece of submitted advice thus ranged from 5 to 25.

Based on these evaluation scores, in the third stage of the experiment, we constructed our performance measures for use in treatment messages. Given our relatively smaller sample (compared to the main field experiment), we employed just three experimental conditions: control, cooperative feedback, and competitive feedback. We randomized participants who completed the first-stage submission into one of these three groups ($N = 333$ in each group), and we then evaluated intergroup balance on self-reported gender and geographic location to ensure the validity of our randomization procedure (see Online Appendix D).

After a 72-hour delay, we contacted each subject via bulk email using the MTurk API. In the email communication, we reminded each subject of his or her earlier submission, quoting the advice supplied by the subject in the first stage. If the subject was assigned to the cooperative treatment, we also informed him or her that, of five other Turkers we showed the advice to, X found it useful, where X was determined by the number of evaluators who assigned the advice a 4 or 5 out

of 5 on the usefulness scale. Once again, we reported the true performance values; these values were not manipulated in any way. If the subject was assigned to the competitive treatment, we informed him or her that the advice was rated as being more useful than Y% of all the advice collected. Again, Y was the true value observed in the data; it was not manipulated. Importantly, the underlying distribution incorporated numerous "ties." Whenever this happened, we randomized the percentage values within the ties. Thus, if 10 pieces of advice tied for first place, we would have randomly assigned percentage values of 99.0%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, and 99.9% across the 10 subjects' treatment messages.

Subjects in all three conditions were then invited to complete a follow-up task on MTurk. In the follow-up task, the subject was paid to submit at least one piece of new advice (again at a rate of $0.40 USD per user), but was also encouraged to provide up to four additional pieces of advice without compensation. Sample email messages issued via the MTurk API are provided below for the cooperative condition (Figure 4) and the competitive condition (Figure 5). The email transmitted in our control condition was formatted similarly but made no mention of performance information. This basic design is inspired by a curiosity study performed by Law et al. (2016), which enables us to assess whether our treatments are effective at inducing subjects to submit uncompensated (freely provided) content, for the public good.

Finally, we again recruited other workers of MTurk to collectively evaluate these final submissions. Each piece of advice was evaluated by a group of five other Turkers, to assess usefulness. If a subject submitted less than four pieces of free advice, evaluation scores

Figure 4. MTurk Study: Cooperative Treatment Message

Hello,

On Mon Mar 20 17:47:49 PDT 2017 you completed a HIT for me where you provided the following advice for new Turkers about how they can work most efficiently / effectively:

"Start doing your work with simple surveys and simple transcriptions (audio & text) work."

Of the 5 other Turkers I asked to evaluate your advice, 1 said that it was helpful! Because the advice we collected earlier was so helpful, we are now hoping to collect some more advice from you. Please navigate to the URL below to access our follow-up HIT on Mechanical Turk:

https://www.mturk.com/mturk/searchbar?selectedSearchType=hitgroups&searchWords=Provide+more+advice+for+new+Turker

Thank you!

Greetings from Amazon Mechanical Turk,

The message above was sent by an Amazon Mechanical Turk user.
Please review the message and respond to it as you see fit.

Sincerely,
Amazon Mechanical Turk
https://www.mturk.com

Figure 5. MTurk Study: Competitive Treatment Message

Hello,

On Mon Mar 20 18:01:03 PDT 2017 you completed a HIT for me where you provided the following advice for new Turkers about how they can work most efficiently / effectively:

"Read the instructions clearly and do the Task"

Based on an evaluation by other Turkers, you advice was better than 9.4% of all the advice we collected! Because the advice we collected earlier was so excellent, we are now hoping to collect some more advice from you. Please navigate to the URL below to access our follow-up HIT on Mechanical Turk:

https://www.mturk.com/mturk/searchbar?selectedSearchType=hitgroups&searchWords=Provide+more+advice+for+new+Turker

Thank you!

Greetings from Amazon Mechanical Turk,

The message above was sent by an Amazon Mechanical Turk user.
Please review the message and respond to it as you see fit.

Sincerely,
Amazon Mechanical Turk
https://www.mturk.com

of 0 were assigned for each omitted submission. Thus, a subject who submitted three pieces of free advice would receive 15 evaluation scores (five per submission), and a value of 0 would be assigned as a placeholder to the fourth.[7]

3.7.2. Data Description and Empirical Approach. Our primary outcome measures include a binary indicator of whether a subject provided any uncompensated advice, as well as an analogous count measure, reflecting the number of freely provided pieces of advice. Our independent variables include our vector of treatment group indicators, as well as a geography fixed effect. These measures are described in Table 8, and

Table 8. MTurk Study: Variable Definitions

Variable	Definition
FreeAdvice	A binary indicator of whether subject *i* supplied any uncompensated advice
CountFreeAdvice	The count of uncompensated advice provided (0 to 4) by subject *i*
FreeAdviceQuality	The summation of crowdsourced usefulness evaluations for all uncompensated advice provided by subject *i* (evaluation scores are set to 0 if no free advice was provided)
Gender	Male = 1; Female = 0
NorthAmerica	A binary indicator of whether subject *i* reported residing in North America
SouthAmerica	A binary indicator of whether subject *i* reported residing in South America
Asia	A binary indicator of whether subject *i* reported residing in Asia
Europe	A binary indicator of whether subject *i* reported residing in Europe

Table 9. MTurk Study: Descriptive Statistics

Variable	Mean	Std. dev.	Min	Max	N
FreeAdvice	0.21	0.41	0.00	1.00	999
CountFreeAdvice	0.48	1.09	0.00	4.00	999
FreeAdviceQuality	7.33	17.03	0.00	80.00	999
Gender	0.62	0.49	0.00	1.00	999
NorthAmerica	0.59	0.49	0.00	1.00	999
SouthAmerica	0.03	0.18	0.00	1.00	999
Asia	0.32	0.47	0.00	1.00	999
Europe	0.04	0.19	0.00	1.00	999

descriptive statistics (means and standard deviations) are presented in Table 9.

We first estimate a linear probability model (LPM) based on our binary measure of freely provided advice, and we evaluate robustness of the result to logistic regression. We then repeat our analyses using the count outcome, employing ordinal logistic regression. Equation (3) reflects our linear model specification. Finally, we again evaluate the quality of the resulting "free" content, based on an OLS, as per the specification in Equation (4). Here, the outcome variable, *FreeAdviceQuality*, reflects the summation of crowdsourced usefulness evaluation scores of subjects' freely provided advice, from the final phase of the experiment. We regressed this value on the same set of covariates, in addition to controlling for *CountFreeAdvice*, to ensure that any observed treatment effects would be interpretable on a per-post basis.

$$FreeAdvice_j = Treat_j + Gender_j + Treat_j \times Gender_j + Geography_j + \epsilon_j, \tag{3}$$

Table 10. MTurk Study: Regression Results—Binary Gender Effects (LPM; DV = *FreeAdvice*)

Explanatory variable	(1)		(2)	
Cooperative	0.053*	(0.032)	0.120**	(0.052)
Competitive	−0.031	(0.030)	−0.034	(0.046)
Cooperative × Gender			−0.109*	(0.067)
Competitive × Gender			0.007	(0.061)
Gender	0.033	(0.027)	0.067	(0.044)
Constant	0.078	(0.113)	0.825***	(0.084)
Geography FE	Yes		Yes	
Observations	999		999	
F-statistic	29.73***		24.41***	
R^2	0.018		0.022	

Note. Robust standard errors in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

$$FreeAdviceQuality_j = Treat_j + Gender_j + Treat_j \times Gender_j$$
$$+ Geography_j + CountFreeAdvice_j$$
$$+ \epsilon_j. \qquad (4)$$

3.7.3. Volume Analysis. Our initial estimation, an LPM using the binary outcome of free advice provision, is presented in Table 10 (note that these results were also robust to the use of a logit estimator). Additionally, Table 11 presents our results using ordinal logistic regression, for the count outcome. For each treatment, we observe that our estimates bear the same sign as was observed in the main field experiment. However, we find that only our estimates related to the cooperative treatment are statistically significant, with females' response being significantly stronger than males'. Thus, our findings provide a partial replication of our main results, and speak to the generalizability of those findings. The fact that the competitive treatment has no significant effect in this context could plausibly be attributed to the relative lack of social interaction amongst subjects, and thus the lack

Table 11. MTurk Study: Regression Results—Count Gender Effects (OLOGIT; DV = *CountFreeAdvice*)

Explanatory variable	(1)		(2)	
Cooperative	0.305*	(0.182)	0.706**	(0.300)
Competitive	−0.168	(0.201)	−0.229	(0.354)
Cooperative × Gender			−0.639*	(0.378)
Competitive × Gender			0.097	(0.432)
Gender	0.226	(0.163)	0.461	(0.289)
/*cut*1	2.124	(1.175)	2.248	(1.198)
/*cut*2	2.739	(1.179)	2.865	(1.203)
/*cut*3	3.275	(1.187)	3.402	(1.209)
/*cut*4	3.522	(1.191)	3.650	(1.214)
Geography FE	Yes		Yes	
Observations	999		999	
F-statistic	1,708.92***		1,855.11***	
R^2	0.012		0.014	

Note. Robust standard errors in parentheses.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

of opportunity to establish status and reputation in this context. Of course, it may also be the case that our null result derives from a lack of study power.[8]

3.7.4. Additional Quality Analysis. We again explored the quality effects of our treatments. We crowdsourced evaluations of the resulting "freely provided" advice from the final stage of our experiment, showing each submission to five other Turkers and obtaining evaluations of usefulness on a scale from 1 to 5, with 5 being most useful. We then sum those evaluation values for each of the pieces of free advice to arrive at a total evaluation score for each (again, between 5 and 25). Finally, we sum across the total evaluation scores for a given subject, substituting a 0 value when a piece of advice was not submitted (e.g., if a subject submitted two pieces of free advice, this implies that two text fields of the total were left blank by the subject; accordingly, we would assign a 0 score as a placeholder for each of the two missing submissions).[9] Finally, we regress the summed total evaluation scores for each user on our treatment indicators, a gender indicator, their interaction, a geography fixed effect, and a control for the count of freely submitted pieces of advice (to ensure that our estimates of the treatment effects are on a per-post basis). The results of our regression are presented in Table 12.

Whereas we observed evidence of a positive volume effect from the cooperative treatment, particularly for females, here we observe evidence of a negative quality effect from the competitive treatment, again particularly for females. Although the interaction with gender is not statistically significant, suggesting that males do not respond differently, the coefficient on the interaction term is relatively large and positive. Thus, our results in this case suggest that the use of a "correct" framing can significantly improve the volume of content produced, while the use of an "incorrect" framing can significantly reduce the quality of the content produced.

Table 12. MTurk Study: Regression Results—Quality Effects (OLS; DV = *FreeAdviceQuality*)

Explanatory variable	(1)		(2)	
Cooperative	0.002	(0.217)	−0.237	(0.313)
Competitive	−0.257	(0.205)	−0.559*	(0.287)
Cooperative × Gender			0.379	(0.421)
Competitive × Gender			0.482	(0.402)
Gender	−0.023	(0.181)	−0.312	(0.260)
CountFreeAdvice	15.47***	(0.200)	15.419***	(0.197)
Constant	−0.445	(0.448)	−0.288	(0.462)
Geography FE	Yes		Yes	
Observations	999		999	
F-statistic	769.45***		684.30***	
R^2	0.975		0.975	

Note. Robust standard errors in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

4. General Discussion

We have drawn on SVO theory to hypothesize how different framings of performance feedback messages (cooperative, individualist, or competitive) may interact with recipient gender to produce different effects on individuals' production of UGC. By conducting two randomized experiments, one in partnership with a mobile-app–based recipe crowdsourcing platform, and a follow-up experiment on Amazon Mechanical Turk involving an ideation task, we examined the causal effects of different performance feedback message framings. We demonstrate that cooperatively framed performance feedback messages are particularly effective at motivating user content contributions, and that females are consistently more responsive to these messages than males. In contrast, we have demonstrated that competitively framed feedback is more effective at motivating male users than female users, though apparently only when there is opportunity for status and reputation building.

Our work makes several important contributions to the literatures on UGC and performance feedback. First, our research builds on past work dealing with the effects of performance feedback on individual engagement by exploring the importance of feedback message framing (e.g., Fisher et al. 2008, White and Peloza 2009) and recipient gender (Brunel and Nelson 2000). We consider the application of these alternative message framings in a novel context involving the production of a public good, UGC. Moreover, we contribute to the literature on framing effects via a systematic consideration of SVO theory, based on which we develop a trichotomous typology of framings that goes beyond the altruism–egoism distinction considered in past work, leading us to consider a third—namely, the "competitive" message framing, which we show to be effective in some situations.

Second, our work contributes to the performance feedback literature by providing a first consideration of a performance feedback type that speaks to cooperative, altruistic motives for user participation. Importantly, we demonstrate that this novel form of performance feedback can be a very effective motivator in both volunteer (Meishijie) and paid (MTurk) work settings. This pair of results attests to the value of tapping into the altruistic, benevolent motives of individuals in many different organizational contexts. Intuitively, if an organization can find a way to communicate performance feedback to an employee in a way that connects their effort and work product to benefits derived by others (e.g., the broader social good, other employees), it may be possible to improve their effort and performance a great deal.

Third, and last, our work builds on past research in IS on the influence of performance feedback mechanisms in online UGC contexts (Moon and Sproull 2008,

Jabr et al. 2014), with an eye toward designing such systems, to improve both the volume and quality of UGC production. Moreover, whereas past work in the IS literature has primarily relied on surveys or archival data, we use a combination of novel field experimental designs. More generally, we also contribute broadly to recent work in IS that has examined interventions that businesses might employ to motivate greater production of UGC (e.g., Chen et al. 2010, Burtch et al. 2017, Goes et al. 2016), to resolve the underprovisioning problem.

Our work also has significant managerial implications. The value proposition of many online platforms depends critically on UGC. For such platforms, motivating contributions early on can be of critical importance, for a platform to reach critical mass. Sustained user contributions are of paramount importance, else underprovisioning may eventually lead to the demise of the platforms. Our study shows that one effective way to work toward this is via the delivery of personalized performance feedback about the popularity of contributors' existing content. Platform managers can utilize our approach, in tandem with others, to design interventions that optimally engage users, to motivate sustained content contributions, by the delivery of personalized performance feedback to account for users' characteristics, social contexts, and performance levels.

Our work is subject to several limitations. First, our treatment messages admittedly reflect but one reasonable set of phrasings intended to capture the desired message framings. It is possible that the use of alternative phrasings or numerical representations would drive differences in subjects' response. Accordingly, future work might seek to evaluate subjects' mental processing in response to alternative phrasings reflecting our intended frames, employing psychometric measures, in a controlled laboratory setting, to arrive at optimal wordings. Second, based on the results of the two experiments, the quality effects appear across contexts. The results suggest interesting opportunities around the context-dependent effect of supplying differentially framed performance feedback on enhancing content quality, by enabling learning or inducing motivation.

5. Concluding Remark

In this work, we report on two theory-informed randomized field experiments that demonstrated a causal effect of performance feedback on users' content contributions in the context of a mobile recipe-sharing application and also on Amazon Mechanical Turk. Our study highlights the importance of gender differences, and shows that effectiveness of this intervention can be significantly enhanced by tailoring message framings based on user gender. However, many open questions remain in this space, and there is ample opportunity

to build on this line of research. It is our hope that future research can utilize our experimental procedure to reproduce and contrast our findings in other contexts, or build on our study, to further explore effective digital interventions to motivate or incentivize UGC production, user engagement, and, more broadly, desirable individual behaviors.

Acknowledgments

Endnotes

[1] http://www.businessinsider.com/how-much-time-people-spend-on-facebook-per-day-2015-7 (*Business Insider*, July 8, 2015).

[2] https://www.youtube.com/yt/press/statistics.html (accessed May 30, 2017).

[3] http://www.fastcompany.com/3030941/most-innovative-companies/linkedin-will-now-rank-your-profile-based-on-popularity (*Fast Company*, May 21, 2014).

[4] Subjects had the ability to disable push notifications in the app. However, randomization should guarantee that the probability of users' having disabled app notifications would not vary systematically across treatment groups at the experiment outset. Some treatments may have induced higher rates of notification disabling once treatments were delivered; however, this would only prevent us from identifying statistically significant effects, given that these users would cease to receive the treatment.

[5] Details of our gender-prediction process are provided in Online Appendix C. Note that our main analyses are robust to the exclusion of subjects where gender was predicted from profile image. Moreover, the key characteristics of the users in our estimation sample exhibit no systematic differences from the rest of the Foodie Talk population.

[6] The number of likes for the same users can vary in different weeks' push notifications. This is particularly relevant for the heterogeneous treatment effects models. For consistency, the main analyses utilize a user-week panel. Nonetheless, to demonstrate robustness to our choice of data structure, we provide the results of a pooled, user-level regression in Table 4.

[7] We ultimately repeated our quality analyses using a different outcome measure—namely, the average crowdsourced evaluation score based only on submitted advice (i.e., not populating zeroes). Thus, these later analyses excluded subjects who did not make any free advice submissions. Our results using the alternative measure were qualitatively similar.

[8] We also considered heterogeneity in the treatment effects once again, with respect to the subjects' performance level. However, in this case, we observed no statistically significant moderating effects. Again, it may be the case that the null result derives from a lack of study power.

[9] Our approach is similar to that taken by Burtch et al. (2017) in their study of online review production; those authors assign a 0 value for quality measures in cases where no content is provided. We would note that, repeating our analysis replacing the dependent variable with the average crowdsourced quality evaluation for any free-advice submitted (i.e., excluding nonsubmitters), we arrive at qualitatively similar results.

References

Ahn D-Y, Duan JA, Mela CF (2015) Managing user-generated content: A dynamic rational expectations equilibrium approach. *Marketing Sci.* 35(2):284–303.

Andreoni J, Vesterlund L (2001) Which is the fair sex? Gender differences in altruism. *Quart. J. Econom.* 116(1):293–312.

Baer J (1997) Gender differences in the effects of anticipated evaluation on creativity. *Creativity Res. J.* 10(1):25–31.

Baron-Cohen S, Wheelwright S (2004) The empathy quotient: An investigation of adults with Aspergers syndrome or high functioning autism, and normal sex differences. *J. Autism Developmental Disorders* 34(2):163–175.

Bateman PJ, Gray PH, Butler BS (2011) The impact of community commitment on participation in online communities. *Inform. Systems Res.* 22(4):841–854.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *Quart. J. Econom.* 119(1):249–275.

Beyer S (1990) Gender differences in the accuracy of self-evaluations of performance. *J. Personality Soc. Psych.* 59(5):960–970.

Blair SL, Lichter DT (1991) Measuring the division of household labor gender segregation of housework among American couples. *J. Family Issues* 12(1):91–113.

Brunel F, Nelson M (2000) Explaining gendered responses to "help-self" and "help-others" charity ad appeals: The mediating role of world-views. *J. Advertising* 29(3):15–28.

Burtch G, Hong Y, Bapna R, Griskevicius V (2017) Stimulating online reviews by combining financial incentives and social norms. *Management Sci.* Article in Advance.

Cabral L, Li L (2015) A dollar for your thoughts: Feedback-conditional rebates on eBay. *Management Sci.* 61(9):2052–2063.

Chen Y, Wang Q, Xie J (2011) Online social interactions: A natural experiment on word of mouth versus observational learning. *J. Marketing Res.* 48(2):238–254.

Chen Y, Harper FM, Konstan J, Li XS (2010) Social comparisons and contributions to online communities: A field experiment on MovieLens. *Amer. Econom. Rev.* 100(4):1358–1398.

Croson R, Gneezy U (2009) Gender differences in preferences. *J. Econom. Literature* 47(2):448–474.

De Paola M, Gioia F, Scoppa V (2014) Overconfidence, omens and gender heterogeneity: Results from a field experiment. *J. Econom. Psych.* 45(December):237–252.

Deng C, Graz J (2002) Generating randomization schedules using SAS programming. Paper 267-27, *Proc. 27th Annual SAS Users Group Internat. Conf.* (SAS Institute, Cary, NC).

Dickinson DL, Tiefenthaler J (2002) What is fair? Experimental evidence. *Southern Econom. J.* 69(2):414–428.

Dufwenberg M, Muren A (2006) Gender composition in teams. *J. Econom. Behav. Organ.* 61(1):50–54.

Eisenberg N, Lennon R (1983) Sex differences in empathy and related capacities. *Psych. Bull.* 94(1):100–131.

Fabes R, Eisenberg N (1998) Meta-analyses of age and sex differences in children's and adolescents' prosocial behavior. Damon W, ed. *Handbook of Child Psychology*, 5th ed., Vol. 3, Social, Emotional, and Personality Development, Eisenberg N, ed. (John Wiley & Sons, Chichester, UK).

Fiedler S, Glöckner A, Nicklisch A, Dickert S (2013) Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organ. Behav. Human Decision Processes* 120(2): 272–284.

Fisher R, Vandenbosch M, Antia K (2008) An empathy-helping perspective on consumers' responses to fund-raising appeals. *J. Consumer Res.* 35(1):519–531.

Gilbert E (2013) Widespread underprovision on Reddit. *Proc.* 2013 *Conf. Comput. Supported Cooperative Work* (*CSCW*), 803–808.

Gneezy U, Niederle M, Rustichini A (2003) Performance in competitive environments: Gender differences. *Quart. J. Econom.* 118(3):1049–1074.

Godes D, Silva JC (2012) Sequential and temporal dynamics of online opinion. *Marketing Sci.* 31(3):448–473.

Goes PB, Guo C, Lin M (2016) Do incentive hierarchies induce user effort? Evidence from an online knowledge exchange. *Inform. Systems Res.* 27(3):497–516.

Goes PB, Lin M, Au Yeung C-m (2014) "Popularity effect" in user-generated content: Evidence from online product reviews. *Inform. Systems Res.* 25(2):222–238.

Gordon FM, Welch KR, Offringa G, Katz N (2000) The complexity of social outcomes from cooperative, competitive, and individualistic reward systems. *Soc. Justice Res.* 13(3):237–269.

Gupta DN, Poulsen A, Villeval MC (2013) Gender matching and competitiveness: Experimental evidence. *Econom. Inquiry* 51(1): 816–835.

Hattie J, Timperley H (2007) The power of feedback. *Rev. Educational Res.* 77(1):81–112.

Hong Y, Huang N, Burtch G, Li C (2016) Culture, conformity and emotional suppression in online reviews. *J. Assoc. Inform. Systems* 17(11):308–329.

Hossain T, List JA (2012) The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Sci.* 58(12):2151–2167.

Iyer G, Katona Z (2015) Competing for attention in social communication markets. *Management Sci.* 62(8):2304–2320.

Jabr W, Mookerjee R, Tan Y, Mookerjee V (2014) Leveraging philanthropic behavior for customer support: The case of user support forums. *MIS Quart.* 38(1):187–208.

Jaworski BJ, Kohli AK (1991) Supervisory feedback: Alternative types and their impact on salespeople's performance and satisfaction. *J. Marketing Res.* 28(2):190–201.

Jones D, Linardi S (2014) Wallflowers: Experimental evidence of an aversion to standing out. *Management Sci.* 60(7):1757–1771.

Jung JH, Schneider C, Valacich JS (2010) Enhancing the motivational affordance of group collaboration environments: The effects of real-time performance feedback and goal setting. *Management Sci.* 56(4):724–742.

Kankanhalli A, Tan BC, Wei KK (2005) Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS Quart.* 29(1):113–143.

Kluger AN, DeNisi A (1996) The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psych. Bull.* 119(2):254–284.

Kraut RE, Resnick P, Kiesler S, Burke M, Chen Y, Kittur N, Konstan J, Ren Y, Riedl J (2012) *Building Successful Online Communities: Evidence-Based Social Design* (MIT Press, Cambridge, MA).

Kwark Y, Chen J, Raghunathan S (2014) Online product reviews: Implications for retailers and competing manufacturers. *Inform. Systems Res.* 25(1):93–110.

Law E, Yin M, Goh J, Chen K, Terry MA, Gajos KZ (2016) Curiosity killed the cat, but makes crowdwork better. *Proc. 2016 CHI Conf. Human Factors in Comput. Systems* (ACM, New York), 4098–4110.

Lee TY, Bradlow ET (2011) Automated marketing research using online customer reviews. *J. Marketing Res.* 48(5):881–894.

Leung L (2009) User-generated content on the Internet: An examination of gratifications, civic engagement and psychological empowerment. *New Media Soc.* 11(8):1327–1347.

Liebrand WB, McClintock CG (1988) The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *Eur. J. Personality* 2(3):217–230.

Ma M, Agarwal R (2007) Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Inform. Systems Res.* 18(1):42–67.

McConnell B, Huba J (2006) The 1% rule: Charting citizen participation. *Church of the Customer Blog* (May 3), https://web.archive.org/web/20100511081141/http://www.churchofthecustomer.com/blog/2006/05/charting_wiki_p.html.

McFarland C, Miller D (1994) The framing of relative performance feedback: Seeing the glass as half empty or half full. *J. Personality Soc. Psych.* 66(6):1061–1073.

Moon J, Sproull L (2008) The role of feedback in managing the Internet-based volunteer work force. *Inform. Systems Res.* 19(4):494–515.

Morin L-P (2015) Do men and women respond differently to competition? Evidence from a major education reform. *J. Labor Econom.* 33(2):443–491.

Muchnik L, Aral S, Taylor SJ (2013) Social influence bias: A randomized experiment. *Science* 341(6146):647–651.

Niederle M, Vesterlund L (2007) Do women shy away from competition? Do men compete too much? *Quart. J. Econom.* 122(3): 1067–1101.

Oestreicher-Singer G, Zalmanson L (2013) Content or community? A digital business strategy for content providers in the social age. *MIS Quart.* 37(2):591–616.

Oliver R (1980) A cognitive model of the antecedents and consequences of satisfaction decisions. *J. Marketing Res.* 17(4):460–469.

Park JH, Gu B, Leung ACM, Konana P (2014) An investigation of information sharing and seeking behaviors in online investment communities. *Comput. Human Behav.* 31(February):1–12.

Pew Research Center (2010) Online product research. (September 29), http://www.pewinternet.org/2010/09/29/online-product-research-2.

Phang CW, Kankanhalli A, Tan BC (2015) What motivates contributors vs. lurkers? An investigation of online feedback forums. *Inform. Systems Res.* 26(4):773–792.

Piliavin JA, Charng H-W (1990) Altruism: A review of recent theory and research. *Annual Rev. Sociol.* 16:27–65.

Piper G, Schnepf S (2008) Gender differences in charitable giving in Great Britain. *Voluntas Internat. J. Voluntary Nonprofit Organ.* 19(2):103–124.

Qiu L, Tang Q, Whinston A (2016) Two formulas for success in social media: Social learning and network effects. *J. Management Inform. Systems* 32(4):78–108.

Rand DG, Brescoll VL, Everett JA, Capraro V, Barcelo H (2016) Social heuristics and social roles: Intuition favors altruism for women but not for men. *J. Experiment. Psych. General* 145(4):389–396.

Roberts JA, Hann IH, Slaughter SA (2006) Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the Apache projects. *Management Sci.* 52(7):984–999.

Ruble DN, Martin CL, Berenbaum SA (2006) Gender development. *Handbook of Child Psychology* (John Wiley & Sons, New York).

Samuelson PA (1954) The pure theory of public expenditure. *Rev. Econom. Statist.* 36(4):387–389.

Shurchkov O (2012) Under pressure: Gender differences in output quality and quantity under competition and time constraints. *J. Eur. Econom. Assoc.* 10(5):1189–1213.

Srivastava R, Rangarajan D (2008) Understanding the salespeople's "feedback-satisfaction" linkage: What role does job perceptions play? *J. Bus. Indust. Marketing* 23(3):151–160.

Stockard J, van de Kragt AJ, Dodge PJ (1988) Gender roles and behavior in social dilemmas: Are there sex differences in cooperation and in its justification? *Soc. Psych. Quart.* 51(2):154–163.

Stocks EL, Lishner DA, Decker SK (2009) Altruism or psychological escape: Why does empathy promote prosocial behavior? *Eur. J. Soc. Psych.* 39(5):649–665.

Swann WB Jr (2012) Self-verification theory. Van Lange P, Kruglanski A, Higgins ET, eds. *Handbook of Theories of Social Psychology* (Sage, London), 23–42.

Swann WB Jr, Read SJ (1981) Self-verification processes: How we sustain our self-conceptions. *J. Experiment. Soc. Psych.* 17(4): 351–372.

Toubia O, Stephen A (2013) Intrinsic vs. image-related utility in social media: Why do people contribute content to Twitter? *Marketing Sci.* 32(3):368–392.

Triandis HC (1989) The self and social behavior in differing cultural contexts. *Psych. Rev.* 96(3):506–520.

Triandis HC (1994) *Culture and Social Behavior.* McGraw-Hill Series in Social Psychology (Trafalgar House Publishing, New York).

Trusov M, Bodapati AV, Bucklin RE (2010) Determining influential users in Internet social network. *J. Marketing Res.* 47(4):643–658.

Van Lange PA (1999) The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *J. Personality Soc. Psych.* 77(2):337–349.

Van Lange PA, De Bruin E, Otten W, Joireman JA (1997) Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *J. Personality Soc. Psych.* 73(4): 733–746.

van Mierlo T (2014) The 1% rule in four digital health social networks: An observational study. *J. Medical Internet Res.* 16(2):e33.

Wasko M, Faraj S (2005) Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quart.* 29(1):35–57.

Weber JM, Kopelman S, Messick DM (2004) A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality Soc. Psych. Rev.* 8(3):281–307.

White K, Peloza J (2009) Self-benefit versus other-benefit marketing appeals: Their effectiveness in generating charitable support. *J. Marketing* 73(4):109–124.

Wozniak D (2012) Gender differences in a market with relative performance feedback: Professional tennis players. *J. Econom. Behav. Organ.* 83(1):158–171.

Zeng X, Wei L (2013) Social ties and user content generation: Evidence from Flickr. *Inform. Systems Res.* 24(1):71–87.

Zhang C, Hahn J, De P (2013) Continued participation in online innovation communities: Does community response matter equally for everyone? *Inform. Systems Res.* 24(4):1112–1130.

Zhang K, Sarvary M (2014) Differentiation with user-generated content. *Management Sci.* 61(4):898–914.

Zhang X, Wang C (2012) Network positions and contributions to online public goods: The case of Chinese Wikipedia. *J. Management Inform. Systems* 29(2):11–40.

Zhang XM, Zhu F (2011) Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *Amer. Econom. Rev.* 101(4):1601–1615.

Zhu F, Zhang X (2010) Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *J. Marketing* 74(2):133–148.