Effectiveness of Performance Feedback in Stimulating User-Generated Content

Completed Research Paper

Yili Hong
Arizona State University
Tempe, Arizona, US
hong@asu.edu

Bin Gu
Arizona State University
Tempe, Arizona, US
bin.gu@asu.edu

Gord Burtch
University of Minnesota
Minneapolis, Minnesota, US
gburtch@umn.edu

Ni Huang
Temple University
Philadelphia, Pennsylvania, US
nina.huang@temple.edu

Chen Liang
Arizona State University
Tempe, Arizona, US
chen.liang.4@asu.edu

Kanliang Wang
Renmin University of China
Beijing, China
kanliang.wang@ruc.edu.cn

Dongpu Fu
Capital University of Economics and Business, Beijing, China
fudongpu@cueb.edu.cn

Bo Yang
Renmin University of China
Beijing, China
yangbo@ruc.edu.cn

Abstract

This study investigates whether and how a platform's provision of performance feedback to users about their prior content contributions can help to stimulate users' subsequent contributions. We draw on social value orientation theory to hypothesize how different framings may impact users' likelihood of producing additional content. We partnered with a major mobile crowdsourcing platform based in China to conduct a randomized field experiment involving the delivery of feedback messages with randomly determined framings, via mobile push notifications. We find that feedback framed either pro-socially or pro-self has a positive effect on content contributions, whereas feedback framed competitively has no such effect. Additionally, we observe differences across genders, such that the positive effects of pro-socially framed feedback are significantly stronger for female users. In contrast, competitively framed feedback is only effective for male users. Our findings provide implications for the design of platform-provided performance feedback to stimulate users' content contribution.

Keywords: User-generated content, randomized field experiment, mobile, length, performance feedback, social value orientation theory

Introduction

User generated content (UGC) is an important aspect of the internet (Ghose and Han 2011, Susarla et al. 2012), influencing individuals' online behavior in a variety of ways (Aral et al. 2013). UGC informs purchases (Chen et al. 2011), aids investment decisions (Park et al. 2014), provides entertainment (Leung 2009), helps firms gather customer intelligence (Lee and Bradlow 2011) and address product uncertainty (Dimoka et al. 2012; Hong et al. 2014). 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, in many cases UGC suffers from an under-provisioning problem, despite its apparent value (Burtch et al. 2015; Chen et al. 2010). UGC is scarce because it is a public good; it is typically supplied voluntarily, and its value is difficult for the producer to internalize. This fact has been widely recognized both in practice (McConnell and Huba 2006; Pew Research 2010) and academic work (Gilbert 2013; Burtch et al. 2015). It has even been reported that a mere 1% of the people who consume UGC also actively contribute it – i.e., the "1% rule" (McConnell and Huba 2006).

Stimulating users to contribute content is thus an issue of prime importance for many online platforms. Accordingly, a number of those platforms have been experimenting with different interventions. For example, LinkedIn informs its users of their popularity ranking relative to peers (Sullivan 2014), in an effort to instill competition and thereby motivate people to improve their profiles and engage more frequently with the community. DangDang provides consumers with monetary rewards in exchange for writing online product reviews. FourSquare provides users with badges recognizing their contributions and activity. Many online communities, more generally, provide users with features enabling them to craft online reputation and social image (Ma and Agarwal 2007). Notably, however, the efficacy of many of these design interventions remains unclear, primarily because of a relative dearth of rigorous academic research.

One intervention, in particular, that has yet to receive significant consideration in the literature on content production is the provision of performance feedback (Moon and Sproull 2008). Many platforms regularly provide feedback to users about the value and quality of the content they have supplied. As a few examples: Goodreads tells users how many times their posts have been read in the prior month; LinkedIn reports the number of other people looking who have recently looked at a user's profile; and Makerbot Thingiverse informs users about the number of times their 3D printing designs have been downloaded or printed by others. Here we explore the efficacy of this form of intervention, in terms of its impact on users' subsequent production of UGC. We explore the nuances of feedback provision, proposing that different message framings for performance feedback will induce different effects in the user-base, depending on users' characteristics and inherent preferences. In particular, we consider that, to maximize the likelihood that a user contributes more content in the future, if that individual is largely pro-social in nature it likely makes sense to inform him or her about performance in terms of the benefits others have derived from his or her recent contributions. Alternatively, if an individual is largely pro-self, or highly competitive, it likely makes sense to inform him or her about performance relative to other users.

In social psychology, social value orientation (SVO) refers to the relative 'weights' that an individual places on others' welfare and his or her own (Fiedler et al. 2013; Van Lange 1999). The concept refers to individuals' 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 results in a three-category typology of a) 'pro-social' (maximizing others gains), b) 'pro-self' (maximizing only the self), and c) competitive (maximizing the self, relative to others). Ample literature notes that SVO is highly correlated with gender, with males exhibiting greater pro-self tendencies and females exhibiting greater pro-social tendencies. For example, Van Lange (1999, pp. 342) observed that "women were relatively more prevalent among pro-socials and less prevalent among individualists and competitors," whereas

[1] http://www.businessinsider.com/how-much-time-people-spend-on-facebook-per-day-2015-7

[2] https://www.youtube.com/yt/press/statistics.html

Croson and Gneezy (2009) report that "women are more averse to competition than are men." Bearing the above in mind, we explore the following research questions in this work:

(1) How and to what degree does performance feedback stimulate user content contribution? How do the effects on user content contribution vary with the framing of feedback message?

(2) How does users' gender interact with the framing of feedback message (pro-social vs. pro-self vs. competitive) in stimulating user content contribution?

To answer these questions, we partnered with a large mobile crowdsourcing recipe application in China to conduct a randomized field experiment. We randomly varied the framing of performance feedback messages, which were delivered to users of the platform via mobile push notification. Subjects were randomly assigned to one of four conditions: pro-social (e.g., you helped x other users), pro-self (e.g., "you are in the top x%"), competitive (e.g., "you beat 1-x% other users"), and control (i.e., no performance feedback) (see Figure 1 in Section 4 for a visual depiction). Notifications were issued every Saturday over the course of a 7-week period. We observed each users' subsequent content contributions and found a number of interesting results. Overall, pro-socially-framed feedback drove the largest proportional increase in content generation, followed closely by pro-self-framed feedback. In contrast, in the general population, we observed no significant effects from competitively-framed feedback relative to the control condition. However, when we explored heterogeneity in the effects across genders, we observed interesting differences, consistent with our expectations. First, the proportional increase in UGC production from pro-socially framed feedback was much stronger amongst females than males. Conversely, we find that the insignificant result around competitively-framed feedback derived largely from a lack of response amongst females; we do observe a significant positive increase in UGC production from this treatment amongst males. Our findings with respect to pro-self-framed feedback are mixed, in that we observe significant positive responses amongst both genders, with a significantly stronger effect on the part of females.

This study makes a number of important contributions to both the academic literature and to practice. First, we contribute to the literature by demonstrating a clean causal effect of platform-provided performance feedback on users' subsequent UGC production. Moreover, we document heterogeneous treatment effects that depend upon a user's gender and its interaction with the framing of the feedback message, in line with the SVO theory. From a practical perspective, we demonstrate the value of accounting for user characteristics in the design and delivery of personalized communications. We demonstrate the clear benefits that platform operators from optimizing their messaging strategies when engaging with a user base.

Theory and Hypotheses

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, for example, consumer decision-making and product sales (e.g., Chevalier and Mayzlin 2006; Chen et al. 2011; Zhu and Zhang 2010), investment decisions (Park et al. 2014) and firm competition (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 a number of important factors that influence these characteristics, including the contributor's popularity (Goes et al. 2014) and cultural background (Hong et al. 2016), and the timing, sequence, and location of UGC contributions (Godes and Silva 2012; Huang et al. 2016). In this paper, we seek to build on and contribute to the third major stream of literature in UGC: the antecedents of UGC production.

Performance Feedback

The literature on online UGC production has focused primarily on fostering sustained participation. A number of studies have explored factors that impact participation in online settings (e.g., Trusov et al.

2010), identifying a series of motivational factors, such as group size and audience effects (Zhang and Zhu 2011), community commitment (Bateman et al. 2011), social networks and peer influence (Zeng and Wei 2013) and system design features, most notably with respect to the delivery of performance feedback (Moon and Sproull 2009; Jabr et al. 2015). It is the latter factor that we focus upon in this work.

Performance feedback is a commonly used approach to motivate individual performance in a variety of settings (Barankay 2011; Delfgaauw et al. 2013; Tran and Zeckhauser 2012), and numerous empirical studies speak to its efficacy (Bandiera et al. 2015; Barankay 2011; Delfgaauw et al. 2013; Tran and Zeckhauser 2012). For example, studies have shown that supplying sales people with performance feedback helps to facilitate learning and perseverance (Sujan et al. 1994), implying greater efficiency and efficacy in sales interactions, and thus sales growth (Delfgaauw et al. 2013). In our particular context, voluntary online contributions, past work tells a similar story. Moon and Sproull (2008) discussed the value of systematic performance-related feedback and found that the presence of such feedback had a distinctly positive effect on solvers (contributors) sustained participation and the quality of answers they provided over time.

At the same time, the broader literature suggests that the benefits of performance feedback are somewhat nuanced. Kluger' and DeNisi's (1996) meta-analysis of the literature observes that performance feedback does not always provide benefits; its effects are heterogeneous and highly contextual. The effects depend on how feedback is provided (Jaworski and Kohli 1991) and the personal traits of the recipient (Srivastava et al. 2008; Wozniak 2012), amongst a host of other factors. Some recent research highlights in particular that the framing of feedback information (Hossain et al. 2012) and the gender of the recipient (Barankay 2011) can play an important role in determining recipient response. Jabr et al.'s (2014) confirm these observations. These authors examined two alternative approaches to recognizing the contributions of solvers: feedback-based recognition, wherein the quality of answers was directly evaluated by questioners, and quantity-based recognition, wherein answers were all treated as equally valuable, regardless of questioner evaluations. These authors observed that the efficacy of feedback-based recognition was heterogeneous, depending heavily on solvers' situation and preferences with respect to peer recognition, social image, social comparison and social exposure.

Though there is a considerable body of research on the subject of performance feedback more generally, very little work in the field of Information Systems (IS) has explored its benefits in the context of voluntarily supplied online public goods, with two notable exceptions described above (Moon and Sproull 2008; Jabr et al. 2014). We therefore seek to build on the findings of these prior studies to understand how performance-feedback can be used to the greatest effect. In so doing, we address a number of open questions. For example, we address the possibility that the public good nature of UGC (Burtch et al. 2015) may lead particular (e.g., pro-social) message framings to be more effective. Moreover, we explore the role gender differences might play, and how gender and message framing might combine to determine individuals' subsequent UGC contributions. By addressing the latter question, we build not only on the IS literature; we also contribute back to the broader literature on performance feedback, which has yet to consider these relationships.

Social Value Orientation Theory and the Role of Gender

SVO speaks to the relative 'weights' that an individual places upon his or her own welfare, and that of others (Fiedler et al. 2013; Van Lange 1999). That is, the 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) 'pro-social' or cooperative (maximizing others' gains, in addition to one's own), b) 'pro-self' or individualistic (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 pro-social orientation, otherwise known as a cooperative orientation, refers to an individual's joint maximization of his or her own payoffs and those of others (Murphy et al. 2011), a 'pro-self' orientation refers to an individual's maximization of his own payoff, without consideration to the payoff of others (Fiedler et al. 2013; Liebrand et al. 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).

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, helping others i.e., pro-social orientation, and, at the other end of the spectrum, to individuals' desire to build image and reputation, which may derive from outperforming other users i.e., competitive orientation (Ahn et al. 2015; Iyer et al. 2015; Zhang et al. 2014). 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 also obtain "image-related" utility by attracting a greater share of peers' attention (Ahn et al. 2015; Iyer et al. 2015; Zhang et al. 2014). In the context of online UGC contribution, the literature suggests that social motivation plays a predominant role (Zhang and Zhu 2011). As such, pro-social feedback message provides a strong confirmation of contributors' self-view. Self-verification theory (Swann 2012; Swann and Read 1981) suggests that such positive feedback strengthens contributors' motivation to take actions to sustain their self-view, which in our context, is achieved by continuing and strengthening their UGC contribution. Previous literature further suggests that another major motivation for UGC contribution is reputation and social recognition (Wasko and Faraj 2005). Such motivations are self-oriented and, as a result, pro-self framed performance feedback help sustain contributors' self-view in this regard. Thus, we propose the following formal hypotheses:

H1a: pro-socially framed performance feedback has the strongest positive effect on user content contributions.

H1b: pro-self-framed performance feedback has a stronger effect on user content contributions than competitively framed performance feedback.

At the same time, a great deal of work notes that SVOs are likely to be highly correlated with individuals' gender. For example, Gupta, Poulsen, and Villeval (2013) found that men tend to be more competitive than women. This difference can be explained by a number of factors. First, research has found that men tend to exhibit lower risk aversion (Croson and Gneezy 2009) and are more likely to be overconfident (Beyer 1990; De Paola et al. 2014), because they 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 women (Gneezy et al. 2003; Morin 2015; Shurchkov 2012).

A significant body of research in both economics and social psychology also speaks to gender differences in other-regarding preferences, pro-sociality or altruism. These differences are generally explained in two ways. First, females have been found to exhibit more socially-orientated traits. It has been found that females tend to feel more empathy (Eisenberg et al. 1987) and exhibit greater sensitivity to social cues (Croson and Gneezy 2009) and others' moods and affect (Piliavin et al. 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, in turn, to respond to others' needs. Consequently, females tend to be more cooperative than males (Stockard et al. 1988) and thus more likely to contribute to the public good. Second, and conversely, males are more likely to be pro-self (Konrad et al. 2001; Van Lange et al. 1997; Weber et al. 2004). According to the gender self-schema theory (Eddleston et al. 2006; Ruble et al. 2006), males are more prone to exhibit a strong pro-self orientation and are more responsive to pro-self feedback (Gordon et al. 2000). Furthermore, studies suggest that males respond positively to competitive environment while females fail to perform or shy away from environments in which they have to compete (Gneezy et al. 2003; Niederle and Vesterlund 2007). With the above in mind, we therefore propose the following additional hypotheses:

H2a: Compared with male users, pro-socially framed performance feedback will have a stronger effect for female users.

H2b: Compared with female users, pro-self and competitively framed performance feedback will have a stronger effect for male users.

Research Methods

Research Context & Experimental Design

Our field experiment was executed in collaboration with one of the largest mobile crowdsourcing recipe applications in China (www.meishijie.net, herein referred to as our corporate partner). Our experimental treatments were designed to be 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' smart phones. Using push notifications to deliver our treatments has a number of natural advantages over other types of digital treatment delivery methods (e.g., email or SMS text). First, due to the large amount of junk and spam emails related to promotions, many users tend to ignore such emails (Burtch et al. 2015). Second, push notifications are integrated with the mobile applications, thus avoiding the concern of SMS text treatment being spam messages.

Our push notifications were designed as part of the company's weekly notification system. Within the application, push notifications (including a short message) appear on the home screen or lock screen. Once clicked or swiped, the user is taken to a landing page within the mobile application, e.g., the home page, or a specific recipe posting. The notifications we created for our experimental treatments pertain to the application's recently released "Shi Hua" (Foodie Talk) section. Foodie Talk is a functional component of the recipe application, implemented in the main mobile application interface (the second tab in Figure 1). By tapping on the top left camera icon from the main application screen, users can initiate posts related to their cooking (implementation of a recipe) or ideas for new recipes in the form of photos and text. The posts become viewable in the Foodie Talk section of the mobile application once they are submitted. Other users can "like" and "comment" on those posts. For each individual post, its current total number of comments and total number of likes are shown in the right bottom corner, near the comment icon and the heart icon.

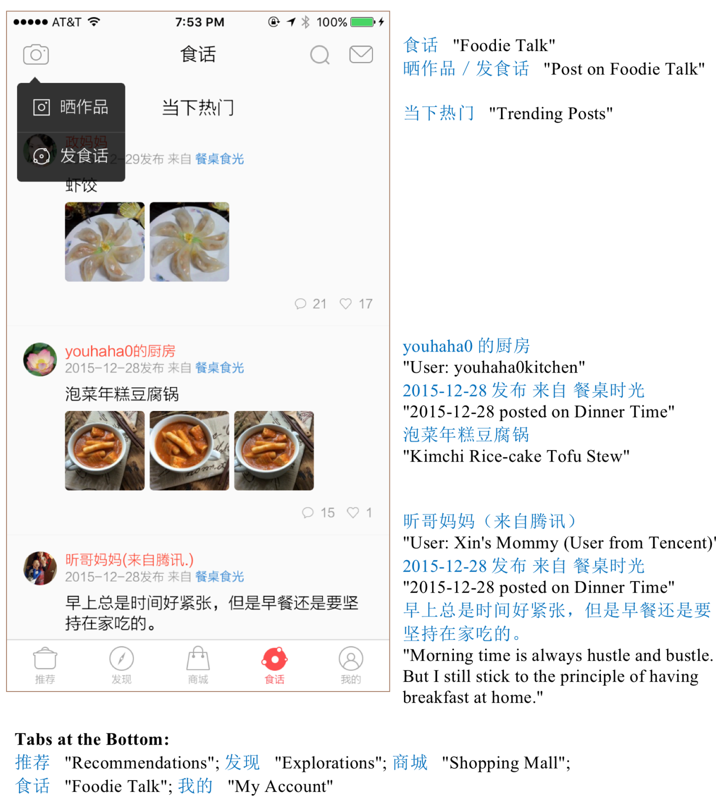
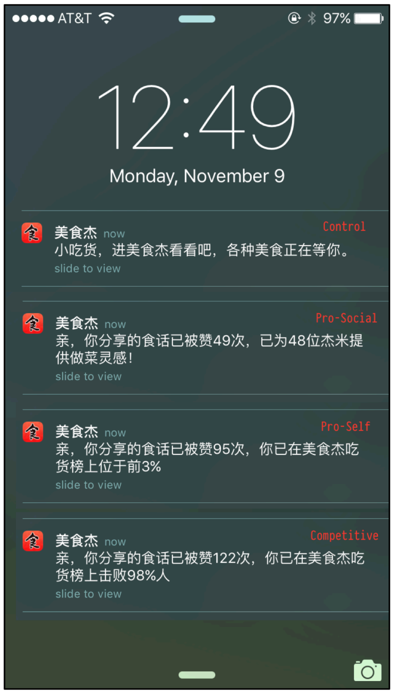


Figure 1. Foodie Talk Sample Page with Translations

Our treatments are intended to stimulate users' posting volumes in the Foodie Talk section of the application. We first designed a control (placebo) group, wherein assigned subjects received a notification that simply reminded them to login to the mobile application. Additionally, we designed three treatment group notifications, which are depicted below in Figure 2.



"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."

Figure 2. Treatment Messages and English Translation

Each user in the "pro-social" treatment group was informed about how many other users had benefited from his or her recent content postings. Each user in the "pro-self" treatment group was informed about their percentile rank (%) compared to other contributors, based on aggregate consumption by others' of the user's recent postings. 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 (%), again based on aggregate consumption by others of the user's recent postings.

Prior to implementing the experiment, we conducted extensive interviews with users of the site to ensure the validity of our treatment stimuli. The interviewees were asked whether the designed messages effectively primed them toward pro-sociality, pro-self or competitiveness, and whether they felt a desire to contribute more to the platform. These interviews helped to ensure that the treatment messages would be effective in stimulating additional content contributions to the application.

Assignment of subjects to treatment groups was performed one day prior to the first treatment delivery (GMT+8 8PM on Nov 7). We worked directly with the IT and marketing department of the corporate partner to develop a standard procedure for delivering our stimuli. Current Foodie Talk users were randomly assigned using pseudo random number generators, with the approach suggested by Deng and Graz (2002). The randomization procedure was integrated into an algorithm in the corporate partner's IT system. In total, 2,360[3] current users of Foodie Talk (590 per group) enter the experiment, in which 730 users provided gender information on their profiles. Randomization checks evaluating the validity of the randomization procedure are conducted to ensure proper random assignment.

Data & Empirical Specification

Besides the key dependent variable of user contribution and group indicators, we also obtain data on a number of user characteristics and behaviors, as we described in Table 1.

[3] This represent all the users who has initiated at least one post on Foodie Talk.

Descriptive statistics (Means and standard deviations) for our outcome variable and controls are presented in Table 2 for our user-level analyses. We observe users in each treatment group for a total of 7 weeks (49 days) from the initiation of the treatments. Thus, for each user, we obtain 49 observations.

Table 1. Variable Definitions

Variable	Definition
Contribution	Total number of postings per group or per user in a day.
Gender	Male=1, female=0
Age	Age of the user.
num_photo	Total numbers of photos the user has posted. We performed log transformation for this variable due to skewness.
num_recipe	Total number of recipes the user has posted. We performed log transformation for this variable due to skewness.
tenure	Number of days passed since the user's initial registration date. We performed log transformation for this variable due to skewness.
num_followers	Number of followers of the user. We performed log transformation for this variable due to skewness.
num_following	Number of users the particular user follows. We performed log transformation for this variable due to skewness.

Table 2. Descriptive Statistics (User level)

Variable	Mean	St.d.	Min	Max
Contribution	0.071	0.555	0.000	12.000
Gender	0.378	0.485	0.000	1.000
Age	25.844	6.837	15.000	60.000
Ln(num_recipe)	0.476	1.395	0.000	6.538
Ln(tenure)	4.495	1.006	1.792	7.584
Ln(num_followers)	1.056	1.379	0.000	7.367
Ln(num_following)	0.610	1.380	0.000	6.847

Our analytical approach is relatively straightforward. We begin with group-level analyses, aggregating daily average posting totals in each group. Using this data, we perform pairwise comparisons (t-tests) between groups. Subsequently, we perform a set of panel regressions, wherein we regress daily group posting volumes on a vector of group indicators. We begin with a OLS regression, and we then also incorporate a vector of day fixed effects to account for possible unobserved temporal trends. Thus, our final model specification is as per Equation 1. Here, i indexes treatment groups, and t indexes time, in days. Thus, $Treatment$ is a treatment group indicator, and Day is a vector of day indicators.

$$Contribution_{it} = Treatment_i + Day_t + \epsilon_{it} \qquad (1)$$

Following our initial group-level analyses, we break our data down to a more granular level to evaluate our hypotheses related to gender heterogeneity. We aggregate our user-level data into treatment-gender groups and again begin with a series of pairwise t-tests comparing each subgroup. When then draw on a user-day level panel to estimate a series of regressions, wherein we evaluate the interaction between each treatment indicator and a gender indicator. We also introduce a series of user-level covariates, in the interests of ensuring precision in our estimates. Our final regression specification in that set of analyses is as per Equation 2. In this case,

$$Contribution_{jt} = Treatment_j + Gender_j + Treatment_j * Gender_j + Controls_j + \epsilon_{it} \quad (2)$$

Results

We begin by evaluating hypotheses H1a and H1b. To assess these hypotheses, we consider group-level outcomes. We begin by graphically depicting daily average postings for each group in Figure 3. Here, the height of each bar indicates the daily average number of posts in each group, and the overlaid error bars reflect the standard errors of the means.

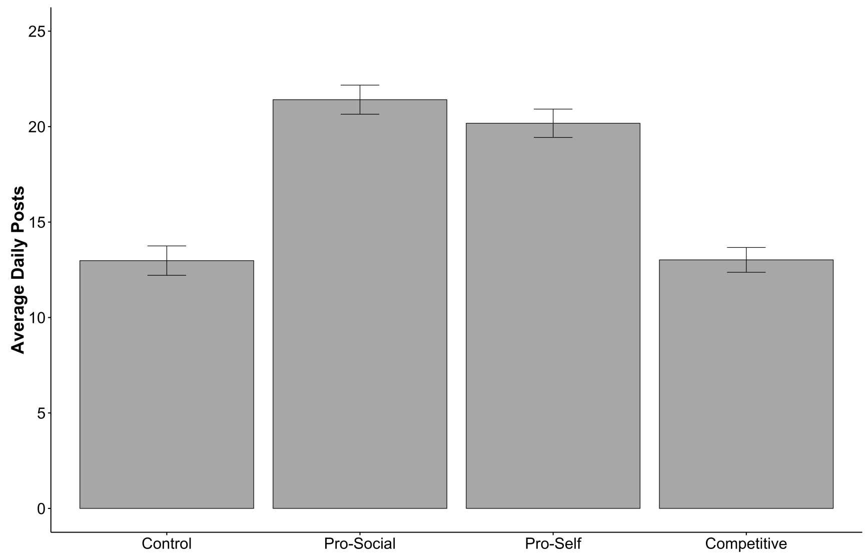


Figure 3. Group Level Treatment Effects

Table 2. Pairwise Comparison of Group Differences

	Difference	**Std. Err.**	***t***	***p* Value**		
				Unadjusted	TUKEY	Bonferroni
Pro-Social vs. Control	8.469	1.061	7.980	**0.000**	**0.000**	**0.000**
Pro-Self vs. Control	7.306	1.061	6.880	**0.000**	**0.000**	**0.000**
Competitive vs. Control	0.184	1.061	0.170	0.863	0.998	1.000
Pro-Social vs. Pro-Self	1.163	1.061	-1.100	0.274	0.692	1.000
Pro-Social vs. Competitive	8.286	1.061	-7.810	**0.000**	**0.000**	**0.000**
Pro-Self vs. Competitive	7.122	1.061	-6.710	**0.000**	**0.000**	**0.000**

Table 2 presents pairwise *t* tests between all groups. Here, we observe some support for hypothesis H1a. The Pro-Social group exhibits the highest average contribution rate, and the difference between that rate and the Competitive condition is significant ($p < 0.001$), as is the difference between the Pro-Social condition and the Control condition ($p < 0.001$). However, we do not observe statistical significant differences between the pro-social and pro-self treatments ($p = 0.274$). We also observe support for hypothesis H1b; the difference between the pro-self condition and the control condition is statistically significant ($p < 0.001$), and positive, as is the difference between the pro-self condition and the competitive condition ($p < 0.001$). Finally, we observe no statistically significant difference in contributions between the Competitive condition and the Control condition ($p = 0.863$). These results are consistent if we employ the TUKEY method or a Bonferroni correction.

Table 3 presents a regression analysis of our group-level data, employing a simple panel structure, wherein we record a single observation per day for each treatment group, reflecting total user contributions to the Foodie Talk page. The results present a fixed effect OLS regression of this dependent variable on group indicators, where the Control group is treated as the reference condition. Thus, the Constant estimate reflects the average daily contribution in the Control group). We observe a statistically significant effect in the pro-social and pro-self groups, but not the Competitive group, indicating that the former two groups are significantly different from Control ($p < 0.001$), producing ~8.5 and ~7.3 more posts per day, respectively. In contrast, the latter group, Competitive, exhibits no statistically significant difference. Once more, performing pairwise of significant differences between coefficient estimates, we observe that the effect of the Pro-Social treatment is not significantly greater than that of the pro-self treatment ($F = 1.30$, $p = 0.259$), yet both the Pro-Social ($F = 75.87$, $p < 0.001$) and pro-self ($F = 45.13$, $p < 0.001$) treatments are both significantly different from the Competitive treatment.

Table 3. Regression Results: Treatment Effects (Group Daily Average Posts)

Explanatory Variable	**FE OLS**
Pro-Social	**8.469*** (0.988)**
Pro-Self	**7.306*** (1.019)**
Competitive	0.184 (1.106)
Constant	**13.041*** (0.644)**
Observations	196
R-squared	0.452
Day Fixed Effects	Yes

Notes: Robust standard errors in parentheses, *** p<0.001.

Gender Differences

We next examine hypotheses H2a, H2b and H2c, related to the interaction between message framings (our treatments) and user gender. To examine these questions, we draw on user-level data. We once again begin by graphically presenting the individual level differences in average daily posts across treatment conditions (Figure 4). However, this time, we break the results down by gender. As noted above, we only observe gender for a subset of our users, thus our results pertain specifically to those users. Similar to the depictions in our group-level results above, the height of the bars indicates the average daily posts per user in each respective group, and the error bars indicate the standard error of the means.

We once again observe results consistent with our expectations. Most notably, we observe significant heterogeneity in male and female responses to the different treatments. Notably, both male users and female users are more active after receiving the altruism-related messages, and the effect is stronger for female users.

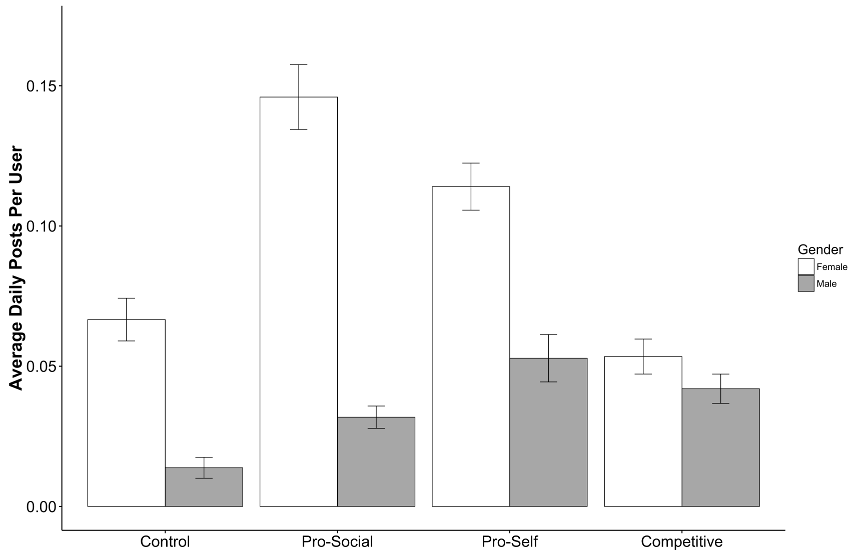


Figure 4. Individual Level Treatment Effects for Different Gender

Table 4. Regression Results: Gender Effects (DV = User Daily Posts)

Explanatory Variable	**(1)**	**(2)**	**(3)**
Pro-Social	**0.058*****(0.005)	**0.080*****(0.008)	**0.050*****(0.007)
Pro-Self	**0.047***** (0.005)	**0.047*****(0.007)	**0.036*****(0.007)
Competitive	0.006 (0.005)	**-0.013***(0.007)	-0.003(0.006)
Gender		**-0.054*****(0.007)	**-0.013***(0.007)
Pro-Social X Gender		**-0.062*****(0.009)	**-0.041*****(0.009)
Pro-Self X Gender		-0.008(0.009)	-0.012(0.009)
Competitive X Gender		**0.043*****(0.008)	**0.036*****(0.007)
Age			**0.003*****(0.001)
Ln(num_recipe)			**0.059*****(0.004)
Ln(tenure)			0.001(0.003)
Ln(num_followers)			**0.066*****(0.005)
Ln(num_following)			**0.007****(0.003)
Constant	**0.044*** (0.003)**	**0.067*****(0.005)	**-0.136*****(0.025)
Day FE	Yes	Yes	Yes
Observations	35,770	35,770	35,770
F Statistic	89.18***	100.83***	161.97***
R-squared	0.002	0.010	0.094
Number of Days	49	49	49

Notes: Cluster-robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$

We next analyze the data at the user level. First, the pairwise comparisons show evidence consistent with that reported in our regressions below. We construct a user-day panel and estimate the main and interaction effects between treatment and gender on daily posting volumes (Table 4). The results of these analyses are consistent with our group-level results, providing similar support for hypotheses H1a and H1b. Next, considering column 2, we observe that, on average, males contribute less content than females. This is perhaps unsurprising, as it indicates that females, who we expect to be more pro-socially oriented, tend to be more willing to contribute UGC, and thus to the public good.

Perhaps the most interesting results are the estimates we observe for our interaction terms in Column 3. These results indicate that, compared with female users, male users respond more strongly to the "Competitive" treatment, yet they respond less strongly to the "Pro-Social" treatment. These two findings provide support for hypotheses H2a and H2c. Interestingly, however, we find evidence contrary to our expectations for H2b. That is, we find that the effect of the pro-self treatment is actually stronger for females than for males. This may simply be a reflection of the fact that the pro-self treatment is merely 'more' pro-social than the competitive treatment, in the sense that individuals who are pro-self do not seek to maximize their own gains at the expense of others. Thus, if we view the treatments, from competitive, to pro-self to pro-social as a sliding scale of increasing pro-sociality, our results are readily rationalized.

Discussion

We have drawn upon SVO theory to hypothesize how different performance feedback messages may interact with recipient gender to produce different effects on individuals' production of UGC. Conducting a randomized mobile field experiment in partnership with a large mobile online crowdsourcing application based in China, we examined the causal effects of different performance feedback messages delivered via mobile push notification. Moreover, we have explored heterogeneity in these treatment effects across user genders. We demonstrate that pro-socially framed performance feedback messages are particularly effective at stimulating user content contributions in this context. However, we also show that these effects vary significantly across genders. We have found that female users are more responsive to pro-socially framed performance feedback, whereas male users are more responsive to pro-self and competitively-framed performance feedback.

Our research builds upon past work dealing with the effects of performance feedback on individual engagement by exploring the importance of feedback message framing, recipient gender, and the interaction between the two. We demonstrate the importance of aligning message framing with a user's characteristics and preferences. Moreover, our work builds on past research in IS on the design and implementation of performance feedback mechanisms in online contexts (Moon and Sproull 2008; Jabr et al. 2014), with an eye toward stimulating UGC production. We identify important factors that platforms should consider in the implementation of these mechanisms, in order to optimize user response. In this same vein, we also contribute to recent work in IS that has examined interventions that businesses might employ to stimulate greater production of UGC (e.g., Chen et al. 2010; Burtch et al. 2015), to resolve the under-provisioning problem. Additionally, whereas prior research has primarily focused on the average effects of various monetary (Wang et al. 2013) or social interventions (Toubia et al. 2013), here we have uncovered important heterogeneous treatment effects over user characteristics. Our findings demonstrate the value of considering the nuances of treatment effects to deliver personalized user interactions. More generally, our study provides some additional empirical evidence confirming the oft reported observation that female and male users have different social value orientations. In this sense, our work contributes to past work dealing with SVO (Fiedler et al. 2013; Murphy et al. 2011).

Our work is of course subject to a number of limitations. First, our study is conducted in the context of crowdsourcing recipe application. While such a context bears resemblance to many other UGC sites, subtle contextual differences may limit generalizability of the findings. Thus future research could explore the effectiveness of performance feedback in other UGC contexts. Second, we limited our considerations of performance feedback to ones based on the SVO theory. However, there are other forms of performance feedback that future research could explore. Third, although pro-social and pro-self performance feedback messages were shown to stimulate users' content contribution in seven weeks, the long run effects of performance feedback were not clear. It is likely that such effects may decay over time. Therefore, future research could examine the dynamic effects of different types of performance feedback.

Conclusion

With a mobile randomized field experiment, this study empirically examines the effectiveness of using different framings of performance feedback to stimulate users' content contribution in the context of a crowdsourcing mobile application. When framed "pro-socially" or "pro-self", performance feedback could effectively drive users towards the desired behaviors. Based on the heterogeneous treatment effects, our study also highlights the importance of gender differences in designing such performance feedback. As some conclusions were drawn, many open questions remain. It is our hope that future research could build on this work to further explore effective methods to incentivize or stimulate UGC.

References

Ahn, D.-Y., Duan, J. A., and Mela, C. F. 2015. "Managing User-Generated Content: A Dynamic Rational Expectations Equilibrium Approach," *Marketing Science*, (35:2),pp 284-303.

Aral, S., Dellarocas, C., and Godes, D. 2013. "Introduction to the special issue-social media and business transformation: A framework for research," *Information Systems Research* 24(1), pp 3-13.

Bandiera, O., Larcinese, V., and Rasul, I. 2015. "Blissful ignorance? A natural experiment on the effect of feedback on students' performance," *Labour Economics* (34), pp 13-25.

Barankay, I. (2011). "Gender differences in productivity responses to performance rankings: Evidence from a randomized workplace experiment," *Unpublished working paper, The Wharton School.*

Beyer, S. 1990. "Gender differences in the accuracy of self-evaluations of performance," *Journal of Personality and Social Psychology* (59:5), p 960.

Burtch, G., Hong, Y., Bapna, R., and Griskevicius, V. 2015. "What Are Social Incentives Worth? A Randomized Field Experiment in User Content Generation," *Proceedings of the International Conference on Information Systems*, Dallas, TX.

Chen, Y., Harper, F. M., Konstan, J., and Xin Li, S. 2010. "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens," *The American Economic Review* (100:4), pp 1358-1398.

Chen, Y., Wang, Q., and Xie, J. 2011. "Online social interactions: A natural experiment on word of mouth versus observational learning," *Journal of Marketing Research,* 48(2), pp 238-254.

Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," Journal of Marketing Research (43:3), pp. 345-354.

Croson, R., and Gneezy, U. 2009. Gender Differences in Preferences. *Journal of Economic Literature*, *47(*2), 448-474.

Delfgaauw, J., Dur, R., Sol, J., and Verbeke, W. (2013). "Tournament incentives in the field: Gender differences in the workplace," *Journal of Labor Economics* 31(2), pp 305-326.

Deng C, Graz J (2002) Generating randomization schedules using SAS programming. *Proceedings of the 27th Annual SAS Users Group International Conference.* Cary, NC: SAS Institute, 2002: 267-27.

De Paola, M., Gioia, F., and Scoppa, V. 2014. "Overconfidence, omens and gender heterogeneity: Results from a field experiment," *Journal of Economic Psychology* (45), pp 237-252.

Dimoka, A., Hong, Y. and Pavlou, P.A., 2012. "On product uncertainty in online markets: Theory and evidence," *MIS Quarterly* (36:2), pp. 395–426.

Eddleston, K. A., Veiga, J. F., and Powell, G. N. 2006. "Explaining sex differences in managerial career satisfier preferences: the role of gender self-schema," *Journal of Applied Psychology* (91:2), p 437.

Eisenberg, N., and Miller, P. A. 1987. "The relation of empathy to prosocial and related behaviors," *Psychological Bulletin* (101:1), p 91.

Fiedler, S., Glöckner, A., Nicklisch, A., and Dickert, S. 2013. "Social Value Orientation and information search in social dilemmas: An eye-tracking analysis," *Organizational Behavior and Human Decision Processes* 120(2), pp 272-284.

Ghose, A. and Han, S.P., 2011. An empirical analysis of user content generation and usage behavior on the mobile Internet. *Management Science*, *57* (9), pp.1671-1691.

Gilbert, E. (2013, February). Widespread Underprovision on Reddit. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW)*, pp. 803-808. ACM.

Gneezy, U., Niederle, M., and Rustichini, A. 2003. "Performance in Competitive Environments: Gender Differences," *Quarterly Journal of Economics* 118(3), pp 1049-1074.

Goes, P. B., Guo, C., and Lin, M. 2016. "Do Incentive Hierarchies Induce User Effort? Evidence from an Online Knowledge Exchange," Information Systems Research, Article in Advance.

Gordon, F. M., Welch, K. R., Offringa, G., and Katz, N. 2000. "The complexity of social outcomes from cooperative, competitive, and individualistic reward systems," *Social Justice Research* (13:3), pp 237-269.

Gupta, Datta N., Poulsen, A., & Villeval, M. C. (2013). "Gender matching and competitiveness: Experimental evidence." *Economic Inquiry*, *51*(1), 816-835.

Hong, Y., Huang, N., Burtch, G., and Li, C., 2016. "Culture, Conformity and Emotional Suppression in Online Reviews," Journal of the Association for Information Systems, forthcoming.

Hong, Y. and Pavlou, P.A., 2014. "Product fit uncertainty in online markets: nature, effects, and antecedents," *Information Systems Research*, 25(2), pp 328-344.

Hossain, T., and List, J. A. 2012. "The behavioralist visits the factory: Increasing productivity using simple framing manipulations," *Management Science* (58:12), pp 2151-2167.

Huang, N., Burtch, G., Hong, Y., and Polman, E. 2016. "Effects of Multiple Psychological Distances on Construal and Consumer Evaluation: A Field Study of Online Reviews," *Journal of Consumer Psychology*. Forthcoming

Iyer, G., and Katona, Z. (2015). Competing for Attention in Social Communication Markets. *Management Science, pp1-17.*

Jabr, W., Mookerjee, R., Tan, Y., & Mookerjee, V. (2014). "Leveraging Philanthropic Behavior for Customer Support: The Case of User Support Forums." *MIS Quarterly* 38(1), 187–208.

Jaworski, B. J., and Kohli, A. K. 1991. "Supervisory feedback: Alternative types and their impact on salespeople's performance and satisfaction," *Journal of Marketing Research*, pp 190-201.

Kluger, A. N., and DeNisi, A. (1996). "The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory," *Psychological bulletin* 119(2), p 254.

Konrad, A. M., and Hartmann, L. 2001. "Gender differences in attitudes toward affirmative action programs in Australia: Effects of beliefs, interests, and attitudes toward women," *Sex Roles* (45:5-6), pp 415-432.

Kwark, Y., Chen, J., and Raghunathan, S. 2014. "Online Product Reviews: Implications for Retailers and Competing Manufacturers," Information Systems Research (25:1), pp. 93-110.

Lee, T. Y., and Bradlow, E. T. (2011). Automated marketing research using online customer reviews. *Journal of Marketing Research*, *48*(5), 881-894.

Leung, L. (2009). User-generated content on the internet: an examination of gratifications, civic engagement and psychological empowerment. *New Media & Society*, *11*(8), 1327-1347.

Liebrand, W. B., and McClintock, C. G. 1988. "The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation," *European journal of personality* 2(3), pp 217-230.

Ma, M., and Agarwal, R. 2007. "Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities," *Information Systems Research* 18(1), pp. 42-67.

McConnell, Ben; Huba, Jackie (May 3, 2006). "The 1% Rule: Charting citizen participation". Church of the Customer Blog. Archived from the original on 11 May 2010. Retrieved 2010-07-10.

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

Morin, L.-P. 2015. "Do men and women respond differently to competition? Evidence from a major education reform," *Journal of Labor Economics* (33:2), pp 443-491.

Murphy, R. O., Ackermann, K. A., and Handgraaf, M. 2011. "Measuring social value orientation," *Judgment and Decision Making* 6(8), pp 771-781.

Niederle, M., and Vesterlund, L. 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?," *Quarterly Journal of Economics* (122:3), pp. 1067-1101.

Park, J. H., Gu, B., Leung, A. C. M., and Konana, P. 2014. "An investigation of information sharing and seeking behaviors in online investment communities," *Computers in Human Behavior* (31), pp 1-12.

Pew Internet Research, Online Product Research. Retrieved from: http://www.pewinternet.org/2010/09/29/online-product-research-2/

Piliavin, J. A., and Charng, H.-W. 1990. "Altruism: A review of recent theory and research," *Annual review of sociology*, pp 27-65.

Ruble, D. N., Martin, C. L., and Berenbaum, S. A. 2006. "Gender development." *Handbook of child psychology* (1998).

Shurchkov, O. 2012. "Under pressure: gender differences in output quality and quantity under competition and time constraints," *Journal of the European Economic Association* 10(5), pp 1189-1213.

Srivastava, R., and Rangarajan, D. 2008. "Understanding the salespeople's "feedback-satisfaction" linkage: what role does job perceptions play?," *Journal of Business & Industrial Marketing* (23:3), pp 151-160.

Stockard, J., Van De Kragt, A. J., and Dodge, P. J. 1988. "Gender roles and behavior in social dilemmas: Are there sex differences in cooperation and in its justification?," *Social Psychology Quarterly*, pp 154-163.

Stocks, E. L., Lishner, D. A., and Decker, S. K. 2009. "Altruism or psychological escape: Why does empathy promote prosocial behavior?," *European Journal of Social Psychology* (39:5), pp 649-665.

Sujan, H., Weitz, B. A., and Kumar, N. 1994. "Learning orientation, working smart, and effective selling," *Journal of Marketing*, 58(3), pp 39-52.

Sullivan, M. 2014. "New LinkedIn feature ranks your popularity — just what we all need."

Susarla, A., Oh, J.-H., and Tan, Y. 2012. "Social networks and the diffusion of user-generated content: Evidence from YouTube," *Information Systems Research* 23(1), pp 23-41.

Swann, W. B., Jr. 2012. "Self-Verification Theory". In P. Van Lange, A. Kruglanski, & E.T. Higgins (Eds.) pp 23-42. *Handbook of Theories of Social Psychology*, Sage: London.

Swann, W. B., Jr., & Read, S. J. 1981. Self-Verification Processes: How We Sustain Our Self-Conceptions. *Journal of Experimental Social Psychology* (17), 351-372

Toubia, O., and Stephen, A. T. 2013. "Intrinsic vs. image-related utility in social media: Why do people contribute content to twitter?" *Marketing Science* 32(3), pp 368-392.

Tran, A., and Zeckhauser, R. 2012. "Rank as an inherent incentive: Evidence from a field experiment," *Journal of Public Economics* 96(9), pp 645-650.

Trusov, M., Bodapati, A. V., & Bucklin, R. E. (2010). "Determining influential users in internet social networks." *Journal of Marketing Research*, *47*(4), 643-658.

Van Lange, P. A. 1999. "The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation," *Journal of Personality and Social Psychology*, 77(2), p 337.

Van Lange, P. A., De Bruin, E., Otten, W., and Joireman, J. A. 1997. "Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence," *Journal of Personality and Social Psychology* 73(4), p 733.

Wang, J., Ghose, A., and Ipeirotis, P. 2012. "Bonus, Disclosure, and Choice: What Motivates the Creation of High-Quality Paid Reviews?" International Conference on Information Systems (ICIS), Orlando, FL: AIS.

Wasko, M. and S. Faraj. 2005. "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly* (29:1) pp. 35-57

Weber, J. M., Kopelman, S., and Messick, D. M. 2004. "A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness," *Personality and Social Psychology Review* (8:3), pp 281-307.

Wozniak, D. 2012. "Gender differences in a market with relative performance feedback: Professional tennis players," *Journal of Economic Behavior & Organization* (83:1), pp 158-171.

Zeng, X., and Wei, L. (2013). Social Ties and User Content Generation: Evidence from Flickr. *Information Systems Research*, 24(1), pp 71–87.

Zhang, K., and Sarvary, M. 2014. "Differentiation with user-generated content," *Management Science* (61:4), pp 898-914.

Zhang, X., and Zhu, F. (2011). "Group size and incentives to contribute: A natural experiment at Chinese Wikipedia." *American Economic Review*, 07-22.

Zhu, F., and Zhang, X., 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," Journal of Marketing (74:2), pp. 133-148.