Anonymity and Language Usage: A Natural Experiment of Social Network Integration

Research-in-Progress

Ni Huang
Temple University
1801 Liacouras Walk, Philadelphia, PA
nina.huang@temple.edu

Yili Hong
Arizona State University
300 E Lemon St, Tempe, AZ 85287
hong@asu.edu

Gordon Burtch
University of Minnesota
321 19th Ave. S., Minneapolis, MN, 55455
gburtch@umn.edu

Abstract

The Internet creates an anonymous and non-authoritarian environment where people do not have social inhibitions and can express opinions freely. However, such disinhibition at times leads to abusive use of language and uncivil behavior in the online environment. This paper leverages data from a natural experiment on an online review platform that integrated social network platform personalization features, which exposes users in an anonymous environment to a social environment. Interestingly, our preliminary findings show that after the social network platform integration, users express more emotions (specifically, they become more positive but less negative), are less likely to use inappropriate language that include sexually explicit words or words that shows rage. Further, users are less egocentric and more social in their language use. We discuss the implication of this study for creating a civil online environment.

Keywords: Anonymity, Social Network Integration, Natural Experiment, Online Reviews, Text Mining

Introduction

The Internet creates an anonymous and non-authoritarian environment where people could express opinions freely. In most online communities or websites, users engage with others using pseudonyms as opposed to their real legal names. People experience the the "online disinhibition effect" as the environment loosens or in many cases, complete abandons of social restrictions and inhibitions that would otherwise be present in normal face-to-face interaction during interactions with others on the Internet (Suler 2004). While anonymity is appealing, it does come with a cost. Perhaps one of the most serious consequences of anonymity is the advent of uncivil behavior ranging from the use of language of racism and hatred (Reader 2012; Santana 2014) to Internet Trolls (Hardaker 2010; Phillips 2011) and cyber bullies (Campbell 2005) in recent years.

Anonymity can certainly open up split personality zone and turn nice people nasty[1]. One common observation of the effect of anonymity on people's cyber behavior, in particular their language usage, is the controversial state of the comment sections on websites like *YouTube* and *Tumblr*, which are often filled with sheer non-sense, blatant racism, and inappropriate content. At the same time, another observation is that the vast majority of these reader comments are published in complete anonymity (e.g., users with pseudonyms). These anecdotal observations fit with the notions summarized by Suler (2004) as to why people sometimes act radically different on the Internet than they would do normally. For example, because when an individual realizes other people don't know him, don't see him and will never meet in real life, he is going to rant and rave anonymously.

The main reason why uncivil human behavior tends to happen under anonymity is mainly the lack of social consequences. In the offline world, if an individual constantly complains about others, he is considered a whiner and would incur such social consequences of having a bad reputation. Therefore, when anonymity is compromised, the potential social consequences will make people abstain from such uncivil behavior under complete anonymity. And lost of complete anonymity occurs with the recent development of social network platform integration in various anonymous online communities, such as online user content generation websites and online user forums.

Social network platform integration are websites' initiatives in integrating user data with social networking platforms such as *Facebook* and *Twitter*. *Facebook*'s' "Instant Personalization" is one major initiative of social network platform integration that enables its partners to use public information-name, gender, networks, and other information one has made public to serve up a personalized experience. Other practices include social login (Frutinger 2013), which is designed to simplify logins for end users as well as provide more and more reliable demographic information to web developers. Nowadays it becomes a widely observed practice that many companies have a traditional log in system also integrate social logins. Traditionally, social network platform integration is considered beneficial to the company because companies are able to obtain more reliable user profile, demographic and social network information. However, users tend to believe that their privacy is being violated.

Although social network platform integration's key objective is to personalize recommendations of products and services, in order to better serve the consumers, there is an unintended effect on user's behavior due to the salient compromise of user anonymity. For example, after social network platform integration, some of your behavior (e.g., content generation such as restaurant or hotel reviews) could be observed by your friends. When such integration happens, users might perceive that the environment has changed, and behave differently accordingly. Therefore, in this study, we seek to examine the following research question:

How does social network platform integration change users' behavior, in terms of their usage of language?

We first propose the effect of social network platform integration on different aspects of language usage, such as emotions (positivity and negativity), inappropriate language (anger, curse and sexually explicit language), and sociability (use of language related friends, social life or [lack of] egocentricity). We then draw on a data set that comprise online restaurant review data from *TripAdvisor* before and after it integrates with *Facebook* with the "Instant Personalization" feature. We found that after the social network platform integration, *TripAdvisor* users become more positive and less negative in their choice of language, controlling for their perceived quality (rating) of the restaurant. Also, users use less language that are sexually explicit and that indicates rage; instead, they show that they are more "sociable" and less "egocentric". The results provide evidence of anonymity on the use of language.

Background

Our study context is TripAdvisor. TripAdvisor has been supporting the notion of "social reviews" since June 14, 2010, when it started to allow users to log in the website via social networking platform's credentials (also known as social logins)[2]. Social logins are "opt-in" type feature, and it is up to the users whether they

---

[1] http://www.nbcnews.com/id/26837911/ns/health-behavior/t/anonymity-opens-split-personality-zone/

[2] http://www.webpronews.com/tripadvisor-goes-social-with-facebook-2010-06

would use the feature. Most notably, on December 21, 2010, TripAdvisor adopted Facebook's "Instant Personalization"[3], TripAdvisor adopted Facebook's "Instant Personalization". "Instant Personalization" is not an "opt-in" feature in that it requires users to "opt-out" if it is undesirable. Although *Facebook* has enabled privacy controls where users could choose to "opt-out" certain personalization features, it has been shown to be challenging in completely opting out "Instant Personalization"[4].

Figure 1. Screenshot of the Instant Personalization on TripAdvisor

Before the social login (June 14, 2010), there is completely no Facebook friend's activity in the TripAdvisor web pages. Between June 14, 2010 and December 21, 2014, users who opt-in to use *Facebook* login could observe limited activity about their *Facebook* friends. After the social network platform integration with "Instant Personalization" (December 21, 2010), if an individual visit TripAdvisor's website while logged into Facebook on your computer (in the past 30 days and enabled cookie), the site will act like his personalized travel planner, complete with restaurant/hotel reviews of his friends, a map showing places his friends have visited, and a list of their most popular destinations. Please see Figure 1 for a screenshot of the "Instant Personalization".

Related Literature

Social Presence Theory

Social presence is defined as individuals' awareness of interaction partners in a medium (Kehrwald 2008; Cobb 2009). According to social presence theory, the social presence that a medium deliver to its users determines the medium's social effects (Lowenthal 2010; Cui et al. 2012). Social presence has been found to be a significant predictor of user satisfaction and learning effectiveness in computer-mediated interactions (Gunawardena and Zittle 1997; Richardson and Swan 2003). In this study, we consider online review websites via the lens of social presence theory. We posit that social network integration affects users' language behavior by increasing social presence.

---

[3] http://techcrunch.com/2010/12/21/facebook-launches-instant-personalization-on-tripadvisor/
[4] https://www.eff.org/deeplinks/2010/04/how-opt-out-facebook-s-instant-personalization

Anonymity

Anonymity has its pros and cons. Earlier IS research has found that in a group setting, anonymity sparks innovative ideas, whereas has a low perceived effectiveness (Connolly et al. 1990), and anonymity increases number of comments, were more critical and probing (Jessup et al. 1990). Behavior methods researchers also found that under anonymity, users have less social desirability and higher self-esteem (Joinson 1999).

Recently, scholars have started to look at how anonymity policies (including government laws and policies and firm policies) affect user behavior. For example, Cho et al. (2012) found that Korean government's real name verification law had no effect of user participation but had a negative effect on users' uninhibited behaviors at the aggregate level. And using a case study of TechCrunch, Omernick and Sood (2013) found quality of online community comments increases after the company disallowed anonymous comments.

Online Reviews and User Content Generation

A long stream of scholarly research exists on online reviews (WOM) across different fields, such as information systems, marketing, economics, and computer science. While earlier work on online reviews has focused on the effect of WOM on sales (Chevalier and Mayzlin 2006; Duan et al. 2008; Zhu and Zhang 2010), recent studies have been geared towards a better understanding of firm strategies (Kwark et al. 2014) and using reviews to infer product type (Hong et al. 2012) and review fraud (Luca and Zervas 2013).

However, the effect of system design on consumer review behavior remains an important yet under-studied topic. Some recent papers start to explore this topic by examining the effect of implementing a multi-dimensional rating system on reviews and consumer satisfaction (Liu et al. 2014), the effect of enabling mobile reviews on review characteristics and helpfulness (Burtch and Hong 2014; Lurie et al. 2014). More broadly, in terms of online review generation, extant studies have looked at "popularity effect" (Goes et al. 2014), "social bias" (Wang et al. 2015), "temporal and spatial distance effect" (Huang et al. 2014), among others. What is neglected is the effect of designing "social systems" on content characteristics. And we seek to address this void in the literature.

Hypotheses Development

Following social network integration, an online review platform will lose a certain degree of anonymity for some users and behavior may shift to become more socially oriented, as a result. The emotional broadcaster theory (EBT) of social sharing argues that individuals have an intrinsic drive to share experiences in a psychologically arousing manner (Harber and Cohen 2005). In a social environment, individuals are more likely to share their feelings and emotions, whether intentionally or unintentionally, a behavior known as emotional leakage (Kraut 1982). Bearing this in mind, we propose our first formal hypothesis:

Hypothesis 1: Social network platform integration will lead to an increase in emotional language.

Social network platform integration exposes users' behavior and content to their social connections. As a result, a user's behaviors become connected to his or her social identity. A social identity is a person's self-concept (Turner and Oakes 1986). Individuals derive utility from pro-social, positive behavior as a signaling mechanism, because such behavior causes an individual to be judged more positively by others and by his or herself (Gneezy et al. 2012). Social identity theory (SIT) suggests that individuals strive to achieve a positive social identity (Jackson et al. 1996; Oldmeadow and Fiske 2010). A primary avenue by which individuals signal their social identity is through the use of language (Jaspal 2009). Accordingly, following social network platform integration, we expect that users will be motivated to use more positive language, to create or maintain a positive image and social identity. Consequently, we hypothesize that:

Hypothesis 2: Social network platform integration will lead to an increase in positive language.

Social consequences are likely to increase with social network integration. Language expectancy theory (LET) proposes that individuals form expectations about what constitutes appropriate language usage, based on established social norms (Burgoon et al. 2002). Violating those norms - e.g., through the use of inappropriate language - may result in a negative reaction from one's peers (Dillard and Pfau 2002). As an anonymous interviewee reported to Kang et al. (2013): *"I posted a very bad review [of a restaurant]. And I guess I did that [anonymously]. I live in a small town, so I certainly didn't want to put my real name...."*

As this quote suggests, we would expect a decline in the use of offensive or explicit language amongst users following a reduction in their anonymity, with social network integration. We hypothesize the following:

Hypothesis 3: Social network platform integration will lead to a decline in inappropriate language.

In social settings, individuals are often driven by a desire to establish or improve their reputation and popularity (Toubia and Stephen 2013; Zhang and Zhu 2011). Scholars have applied Uses and Gratifications theory to understand the drivers of user content generation in social media. Some of the primary gratifications identified in that literature include vanity (Sepp et al. 2011) and the desire for social interaction (Shao 2009). Social network platform integration is likely to increase the exhibition of behaviors that lead to social gratifications, by facilitating a relevant, interested audience for a given user's contributions. Accordingly, we expect that individuals will produce more socially-oriented content following social networking platform integration. More formally, we propose the following hypothesis:

Hypothesis 4: Social network platform integration will lead to an increase in language that is more "social" in nature.

Empirical Methodology

Data

We collected data on restaurant reviews from *www.tripadvisor.com* using a customized web crawler. The reviews pertain to a random sample of restaurants located in five major cities in the United States (New York City, Los Angeles, Chicago, Philadelphia and Phoenix). The data includes review time stamps and content (ratings and text comments), in addition to reviewer profile and restaurant information. The dataset spans the period between 2005 and 2014. We excluded reviews posted in the six-month period between June 14, 2010 (the introduction of social login) and December 21, 2010 (the introduction of instant personalization). This approach provides us with a clean natural experiment, contrasting a period in which the platform was relatively non-social with a period in which the platform was highly social[5]. Moreover, because the social login feature functioned via opt-in, it suffers from clear selection effects. In contrast, the "Instant Personalization" provides an exogenous shock to all *Facebook* users, due to its opt-out design.

Variable Measures and Estimation Method

To construct our measures of linguistic content, we leveraged the latest version of Linguistic Inquiry and Word Count (LIWC), a text analysis application. We use this tool to identify sentiment, emotion, and so forth by comparing review text with a set of pre-defined keyword dictionaries (Pennebaker et al. 2001). LIWC has seen frequent use in the psychology literature, and has also recently begun to see greater use in the Information Systems and Marketing literature as well (Burtch and Hong 2014; Goes et al. 2014; Lurie et al. 2014; Yin et al. 2014). We leverage LIWC's text mining measures of emotions (emotion, positive emotion and negative emotion), inappropriate language (sexually explicit words and rage) and sociability (friends, social process[6] and first person pronouns *"i"*). The LIWC measures are normalized based on the number of words appearing in a given review. We provide some examples of the words included in each category dictionary, in Table 1. More information on the measures can be obtained from *http: //www.liwc.net/*.

Table 1. Sample Words in LIWC's Corpus		
Language Aspect	*Examples*	*Total Words*
Emotion	happy, cried, abandon	915
Positive Emotion	love, nice, sweet	406

---

[5] Please note that including the additional 6 months of data produces parameter estimates that are qualitatively similar, in terms of sign and statistical significance.

[6] Social processes include a large group of words (originally used in LIWC2001) that denote social processes, including all non-first-person-singular personal pronouns as well as verbs that suggest human interaction (talking, sharing).

Negative Emotion	hurt, ugly, nasty	499
Rage	hate, kill, annoyed	184
Sexual	horny, love, incest	96
Friends	buddy, friend, neighbor	37
Social	mate, talk, they, child	455
First person singular pronoun	I, me, mine	12

Table 1. Sample Words in LIWC's Corpus

We leverage the natural experiment at *TripAdvisor* to identify the effects of social platform integration on the linguistic characteristics of reviews. We estimate Equation (1), incorporating restaurant fixed effects (via a within transformation). In this equation, *i* indexes users, *j* indexes restaurants and *t* indexes time. We include observable user characteristics, including a user's age (5 categorical levels) and gender (0=female). Additionally, because the linguistic features of a review may be affected by the perceived quality of a restaurant, we control for both the valence of a review and the number of words in the review.

$$y_{ijt}=\beta_O+\beta_1 \cdot social\_integration_t+\beta_2 \cdot age_i+\beta_3 \cdot gender_i+\beta_4 \cdot rating_{ijt}+\beta_5 \cdot ln(words)_{ijt}+\alpha_j+\varepsilon_{ijt} \qquad (1)$$

Preliminary Findings

We provide some preliminary findings based on both ordinary least squares estimations and estimations with restaurant level fixed effects. The results generally support all of our hypotheses.

Table 2 reports our findings related to emotional language. Controlling for the review valence, number of words, observable reviewer characteristics and a restaurant fixed effect, we observe that social network platform integration leads to: (a) a greater prevalence of emotional language, (b) more positive language, and (c) less negative language.

Table 2: Regression Results for Emotional Expression						
Model:	(1) OLS	(2) FE	(3) OLS	(4) FE	(5) OLS	(6) FE
DV:	*emotion*	*emotion*	*positive*	*positive*	*negative*	*negative*
social integration	0.215***	0.201***	0.235***	0.212***	-0.019***	-0.015***
	(0.030)	(0.033)	(0.030)	(0.033)	(0.004)	(0.004)
age	-0.028**	-0.056***	-0.009	-0.040***	-0.008***	-0.006***
	(0.011)	(0.011)	(0.010)	(0.011)	(0.001)	(0.001)
gender	-0.298***	-0.273***	-0.341***	-0.310***	0.013***	0.012***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.002)	(0.003)
rating	0.573***	0.554***	0.900***	0.888***	-0.093***	-0.097***
	(0.009)	(0.011)	(0.009)	(0.011)	(0.001)	(0.001)
ln(words)	-1.660***	-1.711***	-1.761***	-1.802***	0.260***	0.257***
	(0.016)	(0.019)	(0.016)	(0.019)	(0.001)	(0.001)
Constant	12.381***	12.775***	10.694***	11.025***	-0.287***	-0.266***
	(0.102)	(0.111)	(0.100)	(0.110)	(0.010)	(0.012)
Observations	124,168	124,168	124,523	124,523	127,641	127,641
R²	0.152	0.156	0.203	0.205	0.225	0.219
Restaurant FE	No	Yes	No	Yes	No	Yes
F statistic	3856	2265	6012	3257	11664	9875
Number of restaurants	–	3,688	–	3,688	–	3,691

Notes: 1. Robust standard errors in parentheses; 2. *** p<0.001, ** p<0.01, * p<0.05, +p<0.1

Table 2: Regression Results for Emotional Expression

Table 3 reports our findings relating to the use of inappropriate or explicit language. Controlling for the review valence, number of words, observable reviewer characteristics and a restaurant fixed effect, we observe that social network platform integration leads to: (a) a lower proportion of sexually explicit words, and (b) a lower proportion of words reflecting anger or hatred (i.e., rants and raves).

Table 3: Regression Results for Inappropriate Language				
Model:	(1) OLS	(2) FE	(3) OLS	(4) FE
DV:	*rage*	*rage*	*sexual*	*sexual*
social integration	-0.017***	-0.017***	-0.009**	-0.006[+]
	(0.003)	(0.003)	(0.003)	(0.003)
age	-0.011***	-0.009***	-0.015***	-0.014***
	(0.001)	(0.001)	(0.001)	(0.001)
gender	0.006***	0.005**	-0.064***	-0.063***
	(0.002)	(0.002)	(0.002)	(0.002)
rating	-0.049***	-0.049***	0.038***	0.037***
	(0.001)	(0.001)	(0.001)	(0.001)
ln(words)	0.106***	0.105***	0.069***	0.068***
	(0.001)	(0.002)	(0.001)	(0.001)
Constant	-0.092***	-0.090***	-0.229***	-0.229***
	(0.008)	(0.009)	(0.008)	(0.009)
Observations	127,655	127,655	127,637	127,637
R^2	0.104	0.101	0.044	0.041
Restaurant FE	No	Yes	No	Yes
F statistic	2098	1406	1130	777.3
Number of restaurants	–	3,691	–	3,691

Notes: 1. Robust standard errors in parentheses;
2. *** $p<0.001$, ** $p<0.01$, * $p<0.05$, [+] $p<0.1$

Table 3: Regression Results for Inappropriate Language

Table 4 reports our findings relating to language reflecting sociability. Controlling for the review valence, number of words, observable reviewer characteristics and a restaurant fixed effect, we find that social network platform integration leads to: (a) an increase in references to friends and other "social" words, and (b) a decline in the use of first-person pronouns, which indicates collectivism and a decline in egocentrism.

Table 4. Regression Results for Sociality Language						
Model:	(1) OLS	(2) FE	(3) OLS	(4) FE	(5) OLS	(6) FE
DV:	*friend*	*friend*	*social*	*social*	*i*	*i*
social integration	0.028***	0.029***	0.028***	0.029***	-0.021***	-0.013***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)
age	0.007***	0.007***	0.007***	0.007***	-0.053***	-0.055***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
gender	-0.024***	-0.021***	-0.024***	-0.021***	-0.019***	-0.021***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
rating	0.015***	0.016***	0.015***	0.016***	-0.002	-0.002[+]
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln(words)	0.160***	0.161***	0.160***	0.161***	0.297***	0.297***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Constant	0.119***	0.112***	0.119***	0.112***	-0.435***	-0.429***
	(0.008)	(0.010)	(0.008)	(0.010)	(0.010)	(0.010)
Observations	127,656	127,656	127,656	127,656	127,638	127,638
R^2	0.186	0.181	0.186	0.181	0.243	0.238
Restaurant FE	No	Yes	No	Yes	No	Yes
F statistic	3213	1726	3213	1726	12238	8245
Number of restaurants	–	3,691	–	3,691	–	3,691

Notes: 1. Robust standard errors in parentheses; 2. *** $p<0.001$, ** $p<0.01$, * $p<0.05$, [+] $p<0.1$

Table 4. Regression Results for Sociality Language

Discussion

This paper provides an initial effort to understand the effects of social network platform integration on user-generated content production. We have found that TripAdvisor's attempt to provide personalized services to its users by integrating with Facebook significantly impacted the nature of online reviews. After the adoption of *Facebook* "instant personalization," *TripAdvisor*'s users became more positive (less negative) less sexually explicit, less abrasive (angry), more "sociable" and less "egocentric."

We attribute our findings to an increase in user inhibitions following a relative decline in their perceived anonymity. That is, once behaviors and content became exposed to a user's social network *Facebook*, they began to abstain from behavior that had the potential to be perceived as socially undesirable. The results suggest that reducing anonymity has a strong impact on language usage and thus the content of online reviews. Our initial assessment suggests that this effect is beneficial to both the platform and its users. However, a more rigorous examination can provide greater clarity on the net benefits and costs.

An alternative factor that helps explaining our findings is platform sociability. Platform sociability is the ability of a platform to support social connections and interactions (Goel et al. 2011). To improve sociability, many online platforms incorporate user profiles, behavioral information, and the ability to form social connections, such as social login or instant personalization. The ability of a platform to support social connections and interactions heightens perceived social presence; users' awareness of other users – e.g., an audience for their reviews. We believe that social presence, in turn, can affect how users author restaurant reviews, in terms of their reliance on affective (emotional) processes. Because social reviews that involves friends' opinions may reduce consumers' uncertainty about product quality (Dimoka et al. 2012) and uncertainty about their preferences with the product (Hong et al. 2014, Liu et al. 2014), it is important to understand how reduction in anonymity and increase in sociability will affect online reviews.

Going forward, we have several plans for extension. First, our current measures of linguistic content derive from packaged text mining software (LIWC). Being commercial software, LIWC ensures that we have employed a standardized approach to measurement. However, it would be useful to establish the external validity of such measures via cross-validation, comparing our metrics with actual users' perceptions of reviews' textual characteristics. It may also be fruitful to conduct data mining, to build our own classification models and to undertake a more nuanced attempt at text mining. Second, although we have leveraged a natural experiment to identify the effects of social network integration on user-generated content, due to the observational nature of our data, these analyses may nonetheless suffer from issues of endogeneity (e.g., an unobserved correlated shock). Accordingly, going forward, we plan to supplement this analysis with an experimental manipulation of reviewer anonymity - e.g., priming subjects with privacy concerns (Burtch et al., 2015), to better establish a causal relationship. Third, due to the preliminary nature of this study, we have only examined three salient aspects of review content. We therefore plan to further this study by looking at other aspects of user behavior.

References

Andreoni, J., and Bernheim, B. D. 2009. "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects," *Econometrica* (77:5), pp. 1607–1636.

Burgoon, M., Denning, V. P., and Roberts, L. 2002. "Language expectancy theory," *The persuasion handbook: Developments in theory and practice* pp. 117–136.

Burtch, G., Ghose, A., and Wattal, S. 2015. "The Hidden Costs of Accommodating Crowdfunder Privacy Preferences: A Randomized Field Experiment," *Management Science* .

Burtch, G., and Hong, Y. 2014. "What Happens When Word of Mouth Goes Mobile?" *in Proceedings of International Conference on Information Systems, Auckland, New Zealand* .

Campbell, M. A. 2005. "Cyber Bullying: An Old Problem in a New Guise?" *Australian Journal of Guidance and Counselling* (15:01), pp. 68–76.

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.

Cho, D., Kim, S., and Acquisti, A. 2012. "Empirical analysis of online anonymity and user behaviors: the impact of real name policy," in *System Science (HICSS), 2012 45th Hawaii International Conference on*, IEEE.

Connolly, T., Jessup, L. M., and Valacich, J. S. 1990. "Effects of anonymity and evaluative tone on idea generation in computer-mediated groups," *Management science* (36:6), pp. 689–703.

Das, S., and Kramer, A. 2013. "Self-Censorship on Facebook," in *Association for the Advancement of Artificial Intelligence (AAAI)*, .

Dillard, J. P., and Pfau, M. W. 2002. *The persuasion handbook: Developments in theory and practice*, Sage Publications.

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.

Duan, W., Gu, B., and Whinston, A. B. 2008. "Do online reviews matter?—An empirical investigation of panel data," *Decision support systems* (45:4), pp. 1007–1016.

Frutiger, M., Overby, E., and Wu, D. 2014. "Is Social Network Platform Integration Valuable for an Online Service? A Randomized Field Experiment and Archival Data Analysis," *International Conference on Information Systems*.

Gneezy, A., Gneezy, U., Riener, G., and Nelson, L. D. 2012. "Pay-What-You-Want, Identity and Self-Signaling in Markets," *Proceedings of the National Academy of Sciences (PNAS)* (109:19), pp. 7236–7240.

Goes, P. B., Lin, M., and Au Yeung, C.-m. 2014. ""Popularity Effect" in User-Generated Content: Evidence from Online Product Reviews," *Information Systems Research* (25:2), pp. 222–238.

Harber, K. D., and Cohen, D. J. 2005. "The emotional broadcaster theory of social sharing," *Journal of Language and Social Psychology* (24:4), pp. 382–400.

Hardaker, C. 2010. "Trolling in asynchronous computer-mediated communication: from user discussions to theoretical concepts," *Journal of Politeness Research* (6:2), pp. 215–242.

Hong, Y., Chen, P.-Y., and Hitt, L. M. 2012. "Measuring product type with dynamics of online product review variance," in *Proceedings of the 33th International Conference on Information Systems (ICIS), Orlando, Florida*.

Hong, Y., Pavlou, P. A. 2014. Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents. *Information Systems Research* (25:2), pp. 328-344.

Huang, N., Hong, Y., and Burtch, G. 2014. "The Nature and Effects of Distance Bias in Consumer Evaluation," *Available at SSRN 2514962*.

Jackson, L. A., Sullivan, L. A., Harnish, R., and Hodge, C. N. 1996. "Achieving positive social identity: Social mobility, social creativity, and permeability of group boundaries." *Journal of Personality and Social Psychology* (70:2), p. 241.

Jaspal, R. 2009. "Language and social identity: A psychosocial approach," *Psych-Talk* (64), pp. 17–20.

Jessup, L. M., Connolly, T., and Galegher, J. 1990. "The effects of anonymity on GDSS group process with an idea-generating task," *MIS Quarterly* pp. 313–321.

Joinson, A. 1999. "Social desirability, anonymity, and Internet-based questionnaires," *Behavior Research Methods, Instruments, & Computers* (31:3), pp. 433–438.

Kang, R., Brown, S., and Kiesler, S. 2013. "Why do people seek anonymity on the internet? Informing policy and design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, , ACM.

Kontaxis, G., Polychronakis, M., and Markatos, E. P. 2012. "Minimizing information disclosure to third parties in social login platforms," *International Journal of Information Security* (11:5), pp. 321–332.

Kraut, R. E. 1982. "Social presence, facial feedback, and emotion." *Journal of Personality and Social Psychology* (42:5), p. 853.

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.

Leshed, G. 2008. *Silencing the Clatter: Removing Anonymity from a Corporate Online Community*, Online Deliberation: Design, Research, and Practice., CLSI Publications.

Liu, Y., Chen, P.-Y., and Hong, Y. 2014. "Value of Multi-Dimensional Rating Systems: An Information Transfer View," in *Proceedings of the 35th International Conference on Information Systems (ICIS), Auckland, New Zealand.*

Luca, M., and Zervas, G. 2013. "Fake it till you make it: Reputation, competition, and Yelp review fraud," *Harvard Business School NOM Unit Working Paper* (:14-006).

Lurie, N., Ransbotham, S., and Liu, H. 2014. "The Content and Impact of Mobile vs. Desktop Reviews," in *Association for Consumer Research, Chicago, IL*, .

Oldmeadow, J. A., and Fiske, S. T. 2010. "Social status and the pursuit of positive social identity: Systematic domains of intergroup differentiation and discrimination for high-and low-status groups," *Group Processes & Intergroup Relations* (13:4), pp. 425–444.

Omernick, E., and Sood, S. O. 2013. "The impact of anonymity in online communities," in *Social Computing (SocialCom), 2013 International Conference on*, IEEE.

Pennebaker, J. W., Francis, M. E., and Booth, R. J. 2001. "Linguistic inquiry and word count: LIWC 2001," *Mahway: Lawrence Erlbaum Associates* (71), p. 2001.

Phillips, W. 2011. "LOLing at tragedy: Facebook trolls, memorial pages and resistance to grief online," *First Monday* (16:12).

Reader, B. 2012. "Free press vs. free speech? The rhetoric of "civility" in regard to anonymous online comments," *Journalism & Mass Communication Quarterly* (89:3), pp. 495–513.

Santana, A. D. 2014. "Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards," *Journalism Practice* (8:1), pp. 18–33.

Sepp, M., Liljander, V., and Gummerus, J. 2011. "Private Bloggers' Motivations to Produce Content: A Gratification Theory Perspective," *Journal of Marketing Management* (27:13-14), pp. 1479–1503.

Shao, G. 2009. "Understanding the Appeal of User-Generated Media: A Uses and Gratification Perspective," *Internet Research* (19:1), pp. 7–25.

Sleeper, M., Balebako, R., Das, S., McConahy, A. L., Wiese, J., and Cranor, L. F. 2013. "The Post That Wasn't: Exploring Self-Censorship on Facebook," in *Computer Supported Cooperative Work (CSCW)*, , ACM.

Suler, J. 2004. "The online disinhibition effect," *Cyberpsychology & behavior* (7:3), pp. 321–326.

Toubia, O., and Stephen, A. 2013. "Intrinsic vs. Image-Related Utility in Social Media: Why Do People Contribute Content to Twitter?" *Marketing Science* (32:3), pp. 368–392.

Turner, J. C., and Oakes, P. J. 1986. "The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence," *British Journal of Social Psychology* (25:3), pp. 237–252.

Wang, A., Zhang, M., and Hann, I.-H. 2015. "Socially Nudged: A Quasi-Experimental Study of Friends' Social Influence in Online Product Ratings," *forthcoming, Information Systems Research* .

Yin, D., Bond, S., and Zhang, H. 2014. "Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews," *Mis Quarterly* (38:2), pp. 539–560.

Zhang, M., and Zhu, F. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia," *The American Economic Review* (101:4), pp. 1601–1615.

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.