Communications of the Association for Information Systems

Volume 40 | Article 25

6-2017

Noise or Quality? Cross-nested Hierarchical Effects of Culture on Online Ratings

Thomas Mattson
University of Richmond, mattson@richmond.edu

Recommended Citation

Mattson, Thomas (2017) "Noise or Quality? Cross-nested Hierarchical Effects of Culture on Online Ratings," *Communications of the Association for Information Systems*: Vol. 40 , Article 25.
DOI: 10.17705/1CAIS.04025
Available at: http://aisel.aisnet.org/cais/vol40/iss1/25

Communications of the Association for Information Systems

Research Paper

Noise or Quality? Cross-nested Hierarchical Effects of Culture on Online Ratings

Thomas Mattson
University of Richmond
tmattson@richmond.edu

Abstract:

Previous feedback system research in a variety of contexts has focused on the impact that ratings (as proxies for quality) have on a variety of social and economic outcomes with equivocal findings. These mixed findings may be partially due to noise (factors not related to quality) embedded in aggregated or average positive and negative ratings. One significant source of ratings noise may come from culturally diverse raters' issuing ratings in virtual environments. Culture impacts how groups of individuals are socialized to behave and think, which may result in members' having different attitudes towards publicly downgrading (negative ratings) or praising (positive ratings) other members in the feedback system. In this paper, I investigate how culture influences rating practices specifically in public electronic knowledge sharing communities. Using a cross-nested hierarchical linear model, I demonstrate empirically that cultural differences at the community, occupation, and national levels interact in unique ways to increase or decrease an individual's propensity to give and receive a positive or a negative rating. My study contributes to the literature on rating systems along with having practical ramifications for the designers of feedback systems.

Keywords: Feedback Systems, Ratings, Public Knowledge Sharing Communities, Occupational Culture, National Culture, Hofstede Dimensions, Power Distance, Matthew Effect, Positive and Negative Anchors, and Culture-in-Interaction

This manuscript underwent peer review. It was received 22/06/2016 and was with the authors for 2 months for 2 revisions. John Tripp served as Associate Editor.

1 Introduction

Feedback systems[1] have been implemented in many different types of websites from electronic commerce websites such as eBay and Amazon to user-generated content websites such as Reddit and Youtube. Much of the literature on online feedback systems has focused on explaining the impact that ratings (as proxies for quality) have on a variety of social and economic outcomes (Ba & Pavlou, 2002; Chen, Xu, & Whinston, 2011; Chevalier & Mayzlin, 2006; Dellarocas, 2006; Godes & Silva, 2012; Pavlou & Dimoka, 2006; Poston & Speier, 2005; Son, Tu, & Benbasat, 2006). Unfortunately, this prior literature has reported mixed results whereby sometimes the impact of ratings is positive, negative, or not significant (Ba & Pavlou, 2002; Ghose, Smith, & Telang, 2006; Kauffman & Wood, 2005). These equivocal findings may be partially due to noise (factors not related to quality) embedded in aggregated or average positive and negative ratings (Moe & Trusov, 2011; Muchnik, Aral, & Taylor, 2013; Schlosser, 2005).

At least in theory, ratings should be a somewhat objective indication of the quality of the product, service, or user-generated content being rated, which prior literature often assumes is the case (Fichman, 2011; Sun, 2012). Prior research, however, has offered minimal empirical evidence supporting or contradicting this quality assumption. If ratings are noisy (measuring factors not related to quality), then it poses a significant problem for all types of information consumers who use these ratings to make important decisions (Aral, 2014). For instance, tutorials in online learning environments such as www.dreamincode.net, www.stackoverflow.com, or www.codeproject.com rated for personal, social, and cultural factors not related to the quality of the tutorial can cause learners to learn poor or invalid ideas related to the content of the tutorial. Therefore, we need to understand the factors (variables) in addition to or instead of quality that are involved in ratings. Researchers can use these factors to statistically tease out the portion of the ratings that are based on quality when investigating the impact of ratings, which may help explain the prior contradictory findings. Information consumers can use these factors to wade through the noise embedded in ratings in order to extract the actionable and useful portion of the ratings.

One potentially significant source of ratings noise may come from culturally diverse raters' issuing ratings in virtual environments. On eBay, for instance, consumers from all over the world provide ratings on marketplace participants, and, on www.stackexchange.com, raters from many different national and occupational cultures rate practice-related answers and industry commentary. However, with cultural diversity comes different thought patterns, values, and culturally defined norms among the different raters (Gallagher & Savage, 2013; Koh, Kim, Butler, & Bock, 2007; O'Reilly, Chatman, & Caldwell, 1991; Rui & Stefanone, 2013; Schein, 2010; Triandis, 1994). These differences may adversely influence rating practices (i.e., an individual's propensity to give positive or negative ratings) in feedback systems, because culture plays an important role in determining how groups of individuals are socialized to behave and think (O'Reilly et al., 1991; Qiu, Lin, & Leung, 2013). Certain cultures socialize their members to embrace or avoid conflict and to have different levels of respect for authority (Hofstede, 1980; Triandis, 2000), which may result in members' having different attitudes towards publicly downgrading (negative ratings) or praising (positive ratings) other members in the feedback system. Therefore, ratings may be subject to a cultural bias, which means that aggregated or average ratings may not be an objective measure of actual quality.

In this paper, I specifically address the following research question:

> **RQ**: How does culture influence individuals' propensity to give and receive positive or negative ratings in online feedback systems?

I specifically address this research question in the context of ratings in public electronic knowledge sharing communities using archival data of actual ratings in their feedback system. Public electronic knowledge sharing communities are virtual social structures focused on solving and discussing domain-specific problems concerning a skill-based craft or profession (Boland & Tenkasi, 1995; Earl, 2001; Murillo, 2011; Wasko & Faraj, 2005). Ratings in these websites are important for both the knowledge seeker and the knowledge contributor. For the knowledge seeker, noisy ratings may result in the learning of unsound practices concerning the skill-based craft due to the poor signal that the noisy ratings provide (Cheng & Vassileva, 2005). For the knowledge contributor, highly knowledgeable content contributors may stop contributing due to other members' not rating their high-quality contributions' positively or even rating them negatively (Lampe & Resnick, 2004). Therefore, in order to maximize the effectiveness of ratings for both

[1] There are many different types of feedback systems, which are also commonly referred to as rating systems, reputation systems, or moderation systems. Some systems allow rating on a continuous scale (Amazon's star rating system as of July 2016) and others enable rating in a binary thumb up (positive) and thumb down (negative) manner (YouTube and Slashdot as of July 2016).

knowledge seekers and contributors, we need to understand the behavioral antecedents (above, beyond, or instead of contribution quality) of ratings in feedback systems in order to minimize any noise associated with ratings.

In this paper, I propose that culture at the national, occupational, and community levels interact in unique ways to create a significant source of noise in positive and negative ratings. I argue that positive and negative ratings in feedback systems serve as mechanisms to reinforce national and occupational cultural similarities and differences between community members and to enforce the community's interactive culture, which distracts raters from objectively rating content based on the quality of the contribution. In addition to these main effects associated with culture, I further propose that certain cultural attributes may have a qualifying (moderating) effect on other factors such as contribution quality. In this manner, a rating is a complex social and cultural phenomenon whereby raters rate social and cultural attributes of the member contributing content in conjunction with or in some cases instead of characteristics of the contribution, which may create significant noise (non-quality indicators) in the overall ratings. I provide empirical evidence supporting these cultural effects using a series of cross-nested hierarchical linear models with a sample of the ratings from a large public electronic sharing community of software developers.

2 Literature Review

2.1 Online Ratings

Much of the literature specifically devoted to online feedback systems in a variety of virtual contexts has focused on explaining the impact that ratings have on promoting a variety of outcomes such as price premiums and trust in electronic commerce environments (Ba & Pavlou, 2002; Chevalier & Mayzlin, 2006; Clemons, Gao, & Hitt, 2006; Dellarocas, 2006; Pavlou & Dimoka, 2006; Son et al., 2006), consumer purchase errors (Godes & Silva, 2012), decision quality and time (Poston & Speier, 2005), and effort to contribute content in user-generated content websites (Chen et al., 2011). Unfortunately, this prior literature has reported mixed empirical results whereby sometimes the impact of ratings is positive, negative, or not significant (Ba & Pavlou, 2002; Ghose et al., 2006; Kauffman & Wood, 2005).

These unclear findings may be, in part, because ratings are noisy, which means that there are factors not related to quality embedded in aggregated or average positive and negative ratings (Moe & Trusov, 2011; Muchnik et al., 2013; Schlosser, 2005). Yet, prior empirical research that has focused on explaining the impact of ratings (more broadly than public electronic knowledge sharing communities) has not statistically controlled for the non-quality portion of ratings. Not controlling for these factors is problematic because the theoretical logic for predicting a positive impact of ratings often relies heavily on the assumption that ratings are highly correlated with quality (Kauffman & Wood, 2005). Prior literature, however, has not offered many theoretical explanations for what factors (variables) in addition to or instead of contribution quality are involved in individuals' giving and receiving a positive or negative rating, which makes statistically teasing out the effect of quality particularly difficult when evaluating the impact of accumulated ratings.

Of the limited research that has investigated why something (product, service, answer, response, or other form of user-generated content) receives a positive or negative rating online, the literature offers two primary explanations: 1) ratings are (more or less) an objective indicator of quality and 2) future ratings are subject to an anchoring and resulting Matthew effect. First, prior research has theorized or assumed that ratings are indicators of quality (Fichman, 2011; Sun, 2012). The underlying logic behind this conjecture is that crowds of individuals have the ability to make quality determinations or make decisions more accurately than experts even though the individuals in the crowd may use their own criteria and non-standardized processes to make ratings (Surowiecki, 2004). However, the fact that each individual in the crowd may come to a quality decision using different criteria or different processes does not mean that each individual comes inaccurate or different conclusions (Gao, Greenwood, Agarwal, & McCullough, 2015; Schultz, Mouritsen, & Gabrielsen, 2001).

Second, prior research has theorized that existing ratings may provide a positive or a negative anchor for future ratings (Li & Hitt, 2008; Moe & Trusov, 2011; Muchnik et al., 2013; Schlosser, 2005; Waguespack & Sorenson, 2011; Zhao & Zhou, 2011). These anchors create a Matthew effect whereby the rich get richer (positive anchors) and the poor get poorer (negative anchors) (Merton, 1968). In other words, a highly ranked product, service, answer, contribution, or individual will have a higher likelihood of receiving additional high ratings and vice versa for poorly ranked products, services, or individuals through a process of differential association (DiPrete & Eirich, 2006; Gould, 2002; Merton, 1968). In the context of a public electronic knowledge sharing community where members rank and compare other members based on accumulated ratings in the feedback system, anchoring can adversely impact ratings associated with a specific contribution (post), a specific contributor (individual), or both. These anchoring effects may bias future ratings in either a positive or a negative direction, which creates noise in the overall evaluation of the post—especially if the initial ratings are not true indications of quality.

Prior literature on feedback systems has not theoretically or empirically investigated other sources of noise associated with online ratings besides this theorized anchoring effect. One potentially confounding (but not yet studied) factor relates to culture. Particularly in public electronic knowledge sharing communities, culture may have a significant impact on ratings because these electronic social structures have members from all over the world with many different industry backgrounds. This cultural diversity means that individuals who interact in the virtual environment have different thought patterns, values, and culturally defined norms (Gallagher & Savage, 2013; Koh et al., 2007; O'Reilly et al., 1991; Rui & Stefanone, 2013; Schein, 2010; Triandis, 1994). For example, it may be disrespectful for individuals socialized in certain cultures to negatively rate a high-status individual even if the contribution warrants a negative rating, whereas it may be considered acceptable behavior for individuals socialized in other cultures due to cultural differences related to social inequalities. Similarly, certain cultures have a tendency to avoid conflict (Triandis, 2000), which may impact individuals' tendency to give a negative rating because a negative rating may be considered a conflict generating behavior. Therefore, ratings may be subject to a variety of cultural biases, which makes aggregated or average ratings less likely to be objective measures of quality in these culturally diverse settings.

2.2 Culture

In this paper, I broadly define culture as "the collective programming of the mind that distinguishes one group or category of people from another" (Hofstede & Bond, 1988, p. 51). Hofstede and Bond (1988) use this metaphor to refer to different groups being coded (programmed) via social, political, and economic means (modules) to process information and make sense of the world differently. Culture may be delimited at (among others) the community, occupation, and national levels. Researchers have typically defined a community's culture in terms of a community's interaction style and assumptions concerning what is acceptable and unacceptable conduct in the specific community's context (Eliasoph & Lichterman, 2003; Postmes, Spears, & Lea, 2000). Occupational culture refers to the values, ideologies, and beliefs in certain crafts or professions such as the law, medicine, aviation, and technology (Trice, 1993). National culture refers to ideological and values' differences at the nation state level. Information systems research has most commonly used Hofstede's dimensions of national culture (power distance, uncertainty avoidance, individualism-collectivism, masculinity-femininity, long-term orientation, and indulgence) to measure national culture (Kappos & Rivard, 2008; Leidner & Kayworth, 2006).[2]

Culture at each one of these levels of analysis plays an important role in determining how individuals are socialized to behave and think (Eliasoph & Lichterman, 2003; O'Reilly et al., 1991; Qiu et al., 2013; Triandis, 1972; Trice, 1993). In the online community literature, prior research has used culture (particularly at the national level) to critically examine and question the universality of individual behaviors by arguing that individuals from different cultures can be expected to exhibit different types of behaviors in electronic social structures (Gallagher & Savage, 2013). For instance, prior online community research has demonstrated significant cross-cultural differences in terms of the types of information that individuals are willing to share (Ardichvili, Maurer, Li, Wentling, & Stuedemann, 2006; Li, 2010; Siau, Erickson, & Nah, 2010), membership motivators (Madupu & Cooley, 2010; Shin, 2010), attitudes towards privacy (Marshall, Cardon, Norris, Goreva, & D'Souza, 2008; Pflug, 2011; Posey, Lowry, Roberts, & Ellis, 2010), and membership continuance behaviors (Chiou & Lee, 2008; Grace-Farfaglia, Dekkers, Sundararajan, Peters, & Park, 2006; Pfeil, Zaphiris, & Ang, 2006).

I build off these previous online cultural studies by proposing that cultural differences at the national, occupation, and community levels interact in unique ways to increase or decrease individuals' propensity to positively or negatively rate a contribution. Particularly in public electronic knowledge sharing communities, understanding what aggregated or average ratings in the feedback system mean requires understanding

[2] According to Hofstede (2001): 1) power distance refers to the extent to which a culture accepts status inequalities; 2) uncertainty avoidance refers to a culture's acceptance of ambiguous or uncertain situations; 3) individualism-collectivism is the degree of interdependence a society maintains among its members; 4) masculinity-femininity refers to a cultures competitiveness such as wanting to be the best (masculinity) or caring for others (femininity); 5) long-term orientation refers to how a culture balances its past with the challenges of the present or future; 6) indulgence refers to the extent to which a culture tries to control their impulses.

the national culture of the membership, the occupational culture (and subcultures) that the electronic social structure is situated in, and the interactive culture of the electronic social structure. Therefore, I propose that a rating is a complex social and cultural phenomenon whereby raters rate social and cultural attributes of the contributing member's culture in conjunction with or in some cases instead of the characteristics of the contribution itself because culture plays an important role in determining how individuals behave and think (O'Reilly et al., 1991; Qiu et al., 2013).

3 Research Model

In this paper, I theorize about how members use feedback systems in public knowledge sharing communities (i.e., actual ratings) and not system non-usage (i.e., members' choosing not to rate a contribution). A member may choose not to rate a given contribution (i.e., system non-usage) for a similar or a different set of factors than the variables I propose in this paper. If a member chooses to rate a contribution, then the rating options will either be binary (positive or negative) or along a continuous scale. In all cases, a negative rating is the opposite of a positive rating (i.e., likelihood that a rating will be positive or negative (binary systems) or likelihood that a rating will be closer to the top (positive) or bottom (negative) end of a continuous scale). I frame all hypotheses in relation to negative ratings (see Figure 1) because prior research indicates that negative ratings have a much stronger impact on a variety of outcomes (Chevalier & Mayzlin, 2006) and that negative ratings are much less common (Chevalier & Mayzlin, 2006; Dellarocas & Wood, 2008; Hu, Pavlou, & Zhang, 2009).

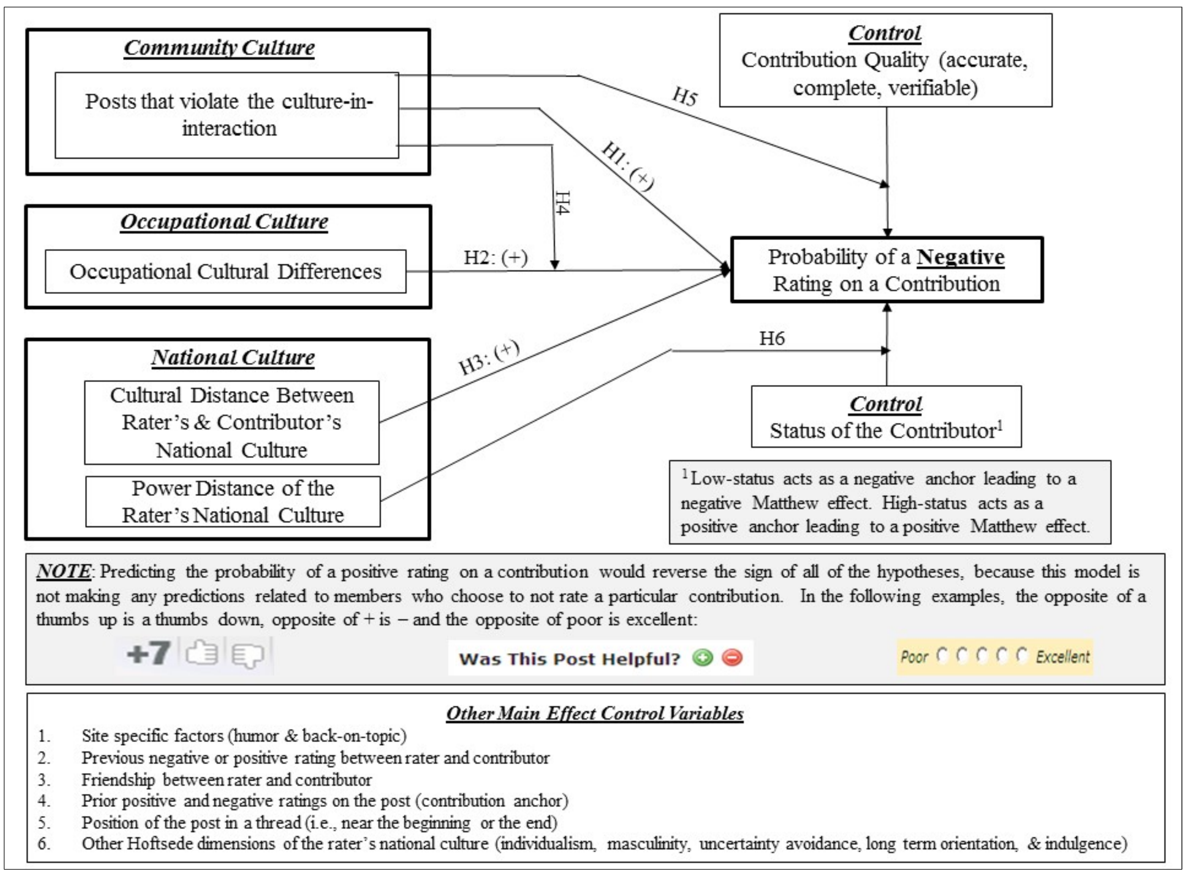

Figure 1. Research Model

3.1 Community Culture

Eliasoph and Lichterman (2003, p. 737) define a community's culture as "recurrent patterns of interaction that arise from a group's shared assumptions about what constitutes good or adequate participation in the group setting", which they refer to as culture-in-interaction. They postulate that a community's culture-in-interaction comprises three dimensions: 1) group boundaries (the group's relationship to the wider world), 2) group bonds (members' mutual responsibilities), and 3) speech (communication) norms (appropriate communication patterns). Although they define these concepts in relation to offline groups, there are roles, responsibilities, communication patterns, and normative expectations for community members whether those interactions occur online or offline (Chen et al., 2011; De Cindio, Gentile, Grew, & Redolfi, 2003; Welser, Gleave, & Fisher, 2007), which makes a group's culture-in-interaction relevant to all types of offline and online collectives. Furthermore, Eliasoph and Lichterman's (2003) model provides a granular and flexible framework to evaluate a virtual community's interactive culture. The virtual community's interactive culture (culture-in-interaction) provides the baseline set of social and cultural norms that raters may use to evaluate, contextualize, and rate online contributions.

Different virtual communities may develop a unique culture-in-interaction. For example, Reddit has a rather unique chaotic culture (Massanari, 2013), Twitter has a distinctive sharing culture surrounding the retweet (Murthy, 2013), and the now defunct Digg had a culture that was a unique blend of individualism and participation (Germonprez & Hovorka, 2013). Therefore, what is considered acceptable behavior varies from website to website because their interactive cultures differ. For example, certain deviant behaviors such as name calling or the use of crude language might be consistent with the group bonds and communication norms in one virtual community such as www.4chan.com (general discussion forum), but these types of behaviors might violate the group bonds and communication norms in another virtual community such as www.codeproject.com (public knowledge sharing community of software developers). Websites such as www.stackoverflow.com (public knowledge sharing community of software developers) have an interactive culture that strongly discourages off-topic posts (group bond), whereas similar websites such as www.dreamincode.net and www.codeproject.com have different group bonds concerning off-topic posts.

Regardless of the website's culture-in-interaction, violating any dimension of it (i.e., post written using proper English at www.4chan.com, an excessive use of emoticons in contributions at www.codeproject.com, or off-topic posts at www.stackoverflow.com) makes the contribution stand out in a negative manner, which may increase the likelihood that the post will receive a negative rating relative to a positive rating because culture-in-interaction violations negatively impact how others in the group view the violator (Eliasoph & Lichterman, 2003). Eliasoph and Lichterman (2003) demonstrate empirically that members who violate any dimension of a community's interactive culture are met with awkward silence or outright rejection from the other group members. Just as individuals perform better when they fit the organizational cultural of a company (O'Reilly et al., 1991), I propose that posts that fit or are aligned with a website's culture-in-interaction will have a reduced likelihood of receiving negative ratings (i.e., those posts will perform better).

Although Eliasoph and Lichterman (2003) made this conclusion in relation to offline groups, a wealth of existing literature supports the idea that online groups interacting in virtual environments behave similarly to offline groups (Jiang, Heng, & Choi, 2013). For instance, Yee, Bailenson, Urbanek, Chang, and Merget (2007) conclude that social interactions in online environments are governed by similar norms and interaction patterns as those in the physical world even though the mode of interaction differs. Additionally, McLaughlin and Vitak (2011) report that online social norm violations often resulted in similar patterns of social rejection as those that Eliasoph and Lichterman (2003) report, which suggests some level of generalizability to online collectives.

> **H1:** Contributions that violate a website's culture-in-interaction will have a higher probability of receiving a negative rating.

3.2 National and Occupational Cultural Differences

Instead of having competing institutions that govern the ratings process as is typically the case with offline ratings (Espeland & Sauder, 2007; Fleischer, 2009; Hsu, Roberts, & Swaminathan, 2012; Waguespack & Sorenson, 2011), rating contributions in a public electronic knowledge sharing community may involve competing occupational and/or national cultures (Blanchard, 2008; Dholakia, Bagozzi, & Pearo, 2004; Stewart & Gosain, 2006). Generally, individuals are ethnocentric by nature, which means that they tend to think of their own culture (occupation or national) as the gold standard and other cultures as "good" only to the extent that they are similar to their own (Triandis, 1994, 2000). As such, individuals tend to have a sense of pride about their own culture, help members of their own cultures, and reject members from different cultures (Matsumoto & Hwang, 2011; Mork, Aanestad, Hanseth, & Grisot, 2008; Triandis, 1994). In this manner, dissimilarity increases negativity towards others and favoritism towards culturally similar others (O'Reilly et al., 1991; Tajfel & Turner, 1986; Triandis, 2000). Therefore, it would be reasonable to expect raters to have a propensity to positively rate similar others and to negatively rate dissimilar others because a means of showing this type of favoritism is through ratings in a feedback system (Stewart, 2005).

Cultural similarities and differences along national and occupational boundaries form a basis for in-group and out-group comparisons (Rao & Ramachandran, 2011; Trice & Beyer, 1993; Yuki, 2003). Prior research in a variety of disciplines has demonstrated empirically that individuals tend to favor other members in their in-group while discriminating against members in their out-group (Brewer, 1979; Diehl, 1990; Tajfel, 1970; Turner, 1975). In this context, the source of in-group favoritism among social groups formed along national or occupational cultural boundaries may be the result of a desire to maintain reciprocal relationships with other in-group members and/or a desire to benefit the in-group as a whole (Brewer & Yuki, 2007; Buchan, Croson, & Dawes, 2002). For example, Barnett, Goodman, and Stewart (2012) argue in the context of open source software communities that individuals' positively rating other individuals with a competing ideology is tantamount to negatively rating their own ideology. This explanation is logical because positively rating an individual with a competing ideology or competing occupational culture increases the social standing of not only the individual being rated but also the social group to which the ideology or occupational culture belongs (Barnett et al., 2012; Stewart & Gosain, 2006).

In public electronic knowledge sharing communities such as www.dreamincode.net and www.codeproject.com, a visible flag icon associated with an individual's national culture typically identifies an individual's national culture and other visible icons such as a Microsoft, a Python, or a Linux logo identify an individual's occupational (sub) cultures. These icons are typically prominently visible next to all contributions so raters can see these cultural attributes when making rating decisions. However, these group distinctions in online environments are more depersonalized than those in typical offline in-groups of close friends and out-groups of strangers, which makes them conceptually similar to minimal groups. A minimal group refers to a set of depersonalized social categories based on seemingly arbitrary category distinctions between in-groups and out-groups (Brewer, 1979; Brewer & Yuki, 2007) such as Java developers or .NET developers in a public electronic knowledge sharing community. Despite the apparent arbitrary nature of minimal groups, however, research has consistently shown the idea of in-group favoritism and out-group derogation to be evident across multiple cultures even without a strong personal interdependence among in-group members and even when the individuals in the minimal in-group and out-group do not directly interact or compete with each other (Brewer & Yuki, 2007; Wetherell, 1982; Yamagishi, Jin, & Kiyonari, 1999).

One theoretical explanation for this mere categorization effect in minimal groups is that there is a social norm to align with one's in-group classification regardless of the strength of the in-group classification (Hertel & Kerr, 2001; Wilder, 1986). The social norm is for members to show a distinct favoritism towards in-group members even when the in-group is based on depersonalized categories (Mullen, Brown, & Smith, 1992; Tajfel, Billig, Bundy, & Flament, 1971). For instance, a public electronic knowledge sharing community of software developers may have a social norm for a minimal group of .NET developers to favor other .NET developers with positive ratings and disfavor out-group Java developers with negative ratings even though the developers may not know each other. The software development industry has a social norm for Microsoft and Oracle to exhibit some level of animosity towards each other, which creates a minimal group categorization of .NET (Microsoft) and Java (Oracle) developers. Although Raymond (1999) does not use the minimal group terminology, his research implies that commercial software developers have an unwritten norm to favor other commercial software developers and that a similar norm exists for open source software developers. This effect in minimal groups may not be as sharp as with more interdependent groups, but ample reported empirical evidence demonstrates that shared in-group membership in minimal groups provides a basis for individuals' inferring relational connections even to relatively unknown others (Brewer & Yuki, 2007; Tajfel et al., 1971).

H2: The greater the occupational cultural differences between the rater and the author of the contribution that is being rated, the higher the probability of a negative rating.

H3: The greater the national cultural differences between the rater and the author of the contribution that is being rated, the higher the probability of a negative rating.

3.3 Interaction of Culture-in-Interaction and Occupational Culture

Given the minimal group nature of online categorizations in public electronic knowledge sharing communities, it may take some form of counter-normative behavior to intensify or activate the categorization effect (Hertel & Kerr, 2001). For instance, posting an answer to a discussion topic in "textspeak" (i.e., language that adopts a lot of abbreviations and slang) instead of in "proper" English when the culture-in-interaction for the website is to use proper English might make the minimal group categorization more pronounced. The counter-normative posting style (culture-in-interaction violation) may further activate the normative expectation for individual group members to show loyalty to one's minimal group (Hertel & Kerr, 2001; Wilder, 1986). Without the culture-in-interaction violation, the minimal group categorization effect may be less evident because the individual has less of a reason to seek out minimal similarities or differences between the members (Mullen et al., 1992; Tajfel et al., 1971).

Furthermore, Triandis (1994) argues that behavioral predictability leads to familiarity and trust, while behavioral unpredictability often leads to group conflicts and tensions. Triandis (1994) made this observation in relation to offline cultures, but the social dynamics in online collectives and groups have become increasingly similar to offline interactions (Jiang et al., 2013; Yee et al., 2007). Prior virtual community and computer-mediated communication research has demonstrated that online discussion forums have social norms that are expected to be followed and not following them may result in flaming and stronger in-group identification (Kayany, 1998; Kim, 2011). Following the social norms and adhering to the forum's culture-in-interaction is a form of behavioral predictability (Lea, O'Shea, Fung, & Spears, 1992). However, once the social norms have been violated and the behaviors of the violators become more unpredictable, then the minimal in-group and out-group categorization becomes sharper (Kayany, 1998; Kim, 2011; Lea et al., 1992).

Therefore, small occupational cultural differences may be amplified (be more pronounced) in the presence of a culture-in-interaction violation because the culture-in-interaction violation may make small occupational cultural differences salient when they may otherwise have been unnoticeable due to behavioral predictability. For instance, if a post violates the community's culture-in-interaction at www.codeproject.com, then programming ideology or paradigm differences between the rater and the contributor may become evident in light of the community culture violation. Without the community culture violation, these differences (minimal group categorizations) may have been otherwise unnoticeable.

> **H4:** A contribution that violates a website's culture-in-interaction will amplify the impact that small occupational cultural differences have on the probability that the contribution will receive a negative rating.

3.4 Interaction of Culture-In-Interaction Violations and Contribution Quality

An objective rating in a feedback system is one that rates the contribution on its own merits while not being distracted or biased by any personal, social, or cultural factors. When raters rate contributions based primarily on quality, feedback systems serve as important mechanisms to filter out or separate the high-quality knowledge contributors from the low-quality knowledge contributors (Lampe & Resnick, 2004; Preece & Shneiderman, 2009; Resnick, Kuwabara, Zeckhauser, & Friedman, 2000). The entire community (both knowledge seekers and knowledge contributors) benefits when high-quality contributions are objectively rewarded with positive ratings and low-quality contributions are punished with negative ratings (Cheng & Vassileva, 2005; Lampe & Resnick, 2004; Poston & Speier, 2005). However, this type of objective rating activity can be difficult to achieve due to individuals' susceptibility to social forces and assorted cognitive biases that may be involved with rating contributions (Cialdini & Goldstein, 2004; Thaler, 2015).

Violating a website's social and/or cultural norms serves as a distraction that can bias objective or rational decision making (Thaler, 2015). This idea is very noticeable in the review process for academic research, particularly in cross-disciplinary work. If, for instance, an academic scholar is reviewing or evaluating a research paper organized in a counter-normative manner (i.e., methodology first, then results, then study motivation, and then theory development), then it becomes more difficult to evaluate the quality of the science and the voracity of the paper's findings. When the authors violate the presentation norm (paper organization) for that particular academic audience, then the evaluators will be much more likely to rate the authors' paper negatively even if the science is of supreme quality. In other words, the rater, reviewer, or evaluator becomes biased or distracted by the counter-normative mode of presentation (Cialdini, 2003; Thaler, 2015).

I propose that the same type of bias may happen online when raters evaluate contributions. For instance, if a member of a public electronic knowledge sharing community posts a three-paragraph answer that accurately answers a knowledge seeker's question but the communication norm for this particular virtual community is to answer questions with only one or two sentences, then the cultural norm violation may distract the rater from objectively rating the contribution based on the quality of the response. In this case, the culture-in-interaction violation may bias the rater into not seeing the high-quality response, which may mitigate the impact of contribution quality on this particular rating.

H5: A contribution that violates a website's culture-in-interaction will mitigate the impact of contribution quality on the probability that the contribution will receive a negative rating.

3.5 Interaction of Power Distance and Contributor Matthew Effect

Hofstede's (1980) power distance dimension of national culture captures the extent to which a culture accepts status inequalities or, said differently, how much respect a culture has for power and authority. High power distance cultures such as Russia, India, and China have a high degree of respect for authority and accept status differences as a cultural norm, whereas low power distance cultures such as Australia, Israel, and Canada have a lower degree of respect for authority and do not accept status differences as a cultural norm (Hofstede, 1980; Triandis, 2000). Previous literature has primarily investigated the concept of power distance in offline environments where there is a face-to-face interaction in order to demonstrate status inequalities and the associated respect (or lack thereof) for authority among the individuals involved in the interaction (Kappos & Rivard, 2008; Leidner & Kayworth, 2006). However, researchers have also demonstrated that the idea of power distance is an important factor (albeit less pronounced) in determining online patterns of social interactions where individuals virtually interact in manners consistent with their power distant cultures (Ardichvili et al., 2006; Jackson & Wang, 2013; Zhang, 2013). For instance, in online learning environments, which are conceptually similar to public electronic knowledge sharing environments, Zhang (2013) qualitatively concluded that power distance had a distinct influence on how individuals interacted with others (peers and instructors) in the online learning setting.

Both the online and the offline ratings' literature have consistently found that existing ratings provide a positive or a negative anchor for future ratings, which leads to a Matthew effect for content contributors (Li & Hitt, 2008; Moe & Trusov, 2011; Muchnik et al., 2013; Schlosser, 2005; Waguespack & Sorenson, 2011; Zhao & Zhou, 2011). However, I propose that the power distance of the rater's national culture will moderate this effect because individuals socialized in high power distance cultures have a greater tendency to act deferentially to those in high-status positions, whereas individuals socialized in low power distance cultures generally have less respect for those in high-status positions (Erez & Gati, 2004; Hofstede & Bond, 1988; Triandis, 2000). In a public electronic knowledge sharing community, high-status individuals are those members who have accumulated high scores (sum of positive ratings minus sum of negative ratings) in the feedback system (Lampe & Resnick, 2004; Stewart, 2005). In the feedback system, positively rating a high-status individual is an act of deference, whereas negatively rating a high-status individual is an act of disrespect (Barnett et al., 2012; Stewart, 2005). Given that individuals from high power distance cultures have more respect for authority, then individuals socialized in these cultures should logically reward posts authored by high-status members of the community with more positive ratings relative to individuals socialized from low power distance cultures. Therefore, I propose that raters from high power distance cultures will have a greater likelihood of perpetuating the Matthew effect for contributions made by high-status members of the community.

To exemplify this proposed moderating effect, consider an example of a public electronic knowledge sharing community of academic researchers who range from assistant professors (low-status) to full professors (high-status) from a diverse set of cultural backgrounds. Further assume that a full professor posts an obviously wrong (low-quality) response that suggests that a researcher use an OLS regression to analyze non-negative count data. The power distance moderating hypothesis that I propose in this paper predicts that a rater from a low power distance culture such as Australia or Israel would have less difficulty negatively rating the full professor's incorrect response relative to a rater from a high power distance culture because low power distance cultures view everyone as more or less equals, which would reduce the impact of the positive anchor that comes from the full professor's high-status. However, a rater who comes from a high power distance culture such as Korea or Malaysia would have a greater likelihood of holding the high-status full professor in high esteem even in connection with the obviously wrong response. Therefore, I suggest that this rater would have a higher likelihood of positively rating the high-status full professor relative to a rater from a low power distance culture, which would amplify the positive anchor that stems from the contributor's high-status. If a low-status assistant professor provided the above OLS response instead of the high-status full professor, then this moderating hypothesis would predict that the rater from the high power distance culture would have a greater likelihood of negatively rating the contribution relative to positively rating it because the negative rating would reaffirm the contributor's low-status in the existing social structure.

H6: The Matthew effect of the content contributor will be moderated by the power distance of the country that the rater identifies with.

4 Research Design and Methods

To test this research model, I downloaded and analyzed a sample of positive and negative ratings in the feedback system from posts in the technology industry discussion forum at TPC (pseudonym) (a large public electronic knowledge sharing community of software developers). The technology industry discussion forum (referred to as the TPC Cubicle in the remainder of the paper) is a forum for members to discuss non-programming specific topics such as enterprise applications, new product announcements, technology consulting, and other technology topics. I selected this discussion forum to maximize the potential generalizability of the findings to related electronic social structures, which may not be designed around question-and-answer forums but where diverse members may still share a cultural affinity (Arrigara & Levina, 2008).

At the time of the study, the feedback system at TPC allowed any registered member to rate any post either up or down in a binary manner, but it did not offer the ability to leave comments along with the rating. TPC did not allow anonymous posting or anonymous ratings so anonymity was not a confounding factor in my empirical analysis. The feedback system at TPC was fully transparent whereby anybody could view the date, the post, and the members involved in the rating interaction. At TPC, each member’s score in the feedback system was calculated as the sum of the total positive ratings minus the sum of the total negative ratings and was visibly displayed next to all posting activity. This simple subtraction of negative ratings from positive ratings is the most common way public electronic knowledge sharing communities determine community status and reputation (Barnett et al., 2012; Stewart, 2005).

To obtain the data, I used a hybrid manual and automated digital collection process whereby I respected and followed all of the website’s policies and procedures (Allen, Burk, & Davis, 2006). I downloaded and analyzed all rating activity in the feedback system in the TPC Cubicle between March 1 and May 31, 2011. This time period represents a typical three-month period in terms of the number of posts that take place in the TPC Cubicle[3], and this three-month period provided a sufficient sample of positive and negative ratings to statistically test my research model. Several ratings contained either a rater or a post author with a blank or deleted profile or profiles with unidentified national cultures, so I excluded those interactions from the analysis. The final data set included 1,017 ratings (132 negative ratings and 885 positive ratings) across 520 posts between 786 unique combinations of members with 136 unique post authors’ receiving ratings and 153 unique raters, which resulted in 207 unique members in my final sample. The low percentage of negative ratings relative to positive ratings is consistent with previous empirical research (Chevalier & Mayzlin, 2006; Dellarocas & Wood, 2008; Hu et al., 2009), and the regression models that I used to analyze these data do not require the baseline probabilities to be equal (Snijders & Bosker, 2012).

Table 1. ANOVAs Comparing Missing Data

		Ratings	**Occupational cultural differences**	**National cultural distance**	**Culture-in-interaction violations**	**Power distance of the rater’s national culture**
Between group ANOVAs	Mean square	0.34	0.14	2.25	0.16	145.90
	F	2.96	0.27	1.38	0.66	2.65
	Sig.	0.09	0.60	0.24	0.42	0.10
Descriptive statistics						
Final sample of ratings	N	1017	1017	1017	520	1017
	Non-centered mean	# Positives: 885 # Negatives: 132	1.3	0.9	0.2	39.6
	Std. deviation	Not applicable	0.7	1.3	0.5	6.9
Removed from sample ratings	N[1]	87	29	55	51	73
	Mean	# Positives: 70 # Negatives: 17	1.4	1.1	0.2	41.1
	Std. deviation	Not applicable	0.7	1.5	0.5	12.7

[1] There were a total of 87 ratings where I could not calculate at least one of the variables, which resulted in my removing those data points from the final sample. In 29 out of the 87 bad data points, I could calculate the occupational cultural difference variable, in 55 out of the 87 bad data points I could calculate the cultural distance variable, and in 73 out of the 87 bad data points I could determine the power distance of the rater’s national culture variable. The 87 ratings with missing data were made across 51 unique posts.

[3] For the seven quarters starting on 1 January, 2010, there were an average of 108 threads and 1,272 posts made by 214 members in the TPC Cubicle per quarter. From 1 March, 2011, to 31 May, 2011 (my study’s sampling period), there were 115 threads and 1,460 posts made by 224 unique members.

Given that I had to remove certain data points due to an inability to calculate or code certain variables, I tested for a potential selection bias in the final sample using a series of ANOVAs (i.e., testing whether excluded rating interactions are statistically different from the included rating interactions). None of the ANOVAs revealed significant differences at the 0.05 level (see Table 1 above). Therefore, at least in these data during this time period, members who fully filled out their online profiles did not exhibit statistically significant different rating practices from those members who did not fully fill out their online profiles.

Table 2 summarizes the operational definitions of each variable, which I explain in the subsequent sections.

Table 2. High-level Summary of the Operationalization of Each Hypothesized Variable

	Operationalization	Citation(s) (if applicable)
Ratings	Binary positive (0) or negative (1) rating	The feedback system at TPC only allows for non-anonymous binary ratings.
Culture-in-Interaction violations	A sum of the binary evaluation of each culture-in-interaction dimension (group bonds, communication norms, and group boundaries). The values range from 0 (no violations) and 3 (violations along all three culture-in-interaction dimensions)	I used Eliasoph and Lichterman's (2003) conceptual definition but the specific instantiation is unique to my study because no other paper has used their framework in a virtual community. Users of online discussion forums and theoretical academics further informed this operational definition.
Occupational cultural differences	A sum of a binary comparison of the rater and receiver of the rating along three dimensions: programming paradigm, ideology, and primary programming language. The values for this variable range from 0 (no differences) to 3 (differences along all three dimensions).	This operational definition is unique to my paper. However, certain aspects were informed by Barnett et al. (2012), Raymond (1999), Spinellis, Drossopoulou, and Eisenback (1994), Stewart and Gosain (2006), and Trice and Beyer (1993). Programmers and technical academics provided additional guidance for this operational definition.
National cultural distance	$CD_{ij} = \sqrt{\sum_{k=1}^{K} \{(I_{kj} - I_{ki})^2 / V_k\}}$, where I_{kj} is the rater's country score on the kth cultural dimension, I_{ki} is the contributor's country score on the kth cultural dimension, and V_k is the variance of the kth cultural dimension. I measured each of Hofstede's dimensions of culture at the national level using Hofstede's scores based on the country of origin field in each member's profile.	Geletkanycz (1997), Mitchell, Smith, Seawright, and Morse (2000), and Kogut and Singh (1988)
Power distance of the rater's national culture	I measured power distance at the national level using Hofstede's scores based on the country of origin field in each rater's profile.	Geletkanycz (1997) and Mitchell et al. (2000)
Main effect control variables but each are involved in a hypothesized moderating effect		
Contribution quality	Using a three-point scale (-1 for low quality, 0 for neither high nor low quality (or not applicable), and 1 for high quality), three PhD researchers coded each post along each of Fichman's (2011) dimensions of quality (accuracy, completeness, and verifiability). After coding, I summed the values across all three dimensions, which resulted in a scale from -3 (low quality on all three dimensions) to +3 (high quality on all three dimensions).	Fichman (2011) but the three-point scale is unique to my paper.
Status of the contributor	Each member's accumulated score in the feedback system (sum of positive ratings minus the sum of negative ratings), which I then mapped to TPC's custom clustering of membership.	This is the status/ranking process of members at TPC. Other research by Barnett et al. (2012) and Stewart (2005) used a similar method.

4.1 Culture-in-Interaction

Determining and coding each post that received a rating against the culture-in-interaction for this forum involved two steps: 1) determining each culture-in-interaction dimension of the TPC Cubicle and 2) coding each post that received a rating against the culture-in-interaction dimensions. Consistent with the coding processes that Strauss and Corbin (1990) outline, two PhD researchers iteratively analyzed all (non-archived) discussion threads between 1 January and 28 February, 2011, and between 1 June and 31 August, 2011, in the TPC Cubicle in order to determine each culture-in-interaction dimension. These two time periods included 188 discussion threads and 1,805 posts in those discussion threads (61 threads and 607 posts in the first time period and 127 threads and 1,198 posts in the second time period).

Table 3. Culture-in-Interaction of the TPC Cubicle

	Communication norms	Group bonds	Group boundaries
Eliasoph and Lichterman's (2003) conceptual definition	A group's assumptions about what appropriate communication is in the group context	A group's assumptions about what members' mutual responsibilities should be while in the group context	A group's assumptions about what the group's relationship to the wider world should be while in the group context
Primary elements of each dimension	1) Posts that are neither excessively long nor short 2) Posts that use minimal chains of quotations 3) Proper punctuation and capitalization (within reason) 4) Cautious use of emoticons 5) Occasional use of crude/foul language 6) Use of spoiler tags 7) Matching the title of the thread with the purpose	1) Make insightful comments pertaining to the topic of the thread 2) Keep the mood fairly light but on point 3) Not lobby for ratings in the feedback system (4) Start threads which may generate significant discussion 5) Have a degree of respect for thread participants	1) Posting topics in the correct forum in TPC 2) Posting links only internal to TPC (i.e., no self-referral links or links to competing electronic knowledge sharing communities) 3) Discussing non-technical disciplines as inferior to programming 4) Discussing only ethical issues related to the IT discipline (Taboo topics include the creation of malware or viruses) (5) Correcting code only when asked
Sample of evidence	**Members directly responding to posts that violated the communication norms or the group bonds** 1) "We don't do that around these parts" or "we use spoiler tags in here". 2) "What's up with all of the emoticons?" or "what's up with the embedded quotations?". 3) Some of the responses used crude/foul language to identify that a previous post was out of the norm. This type of foul language was used most frequently with posts concerning the feedback system and why a particular post received negative ratings. 4) "I stopped reading after the first sentence in that wall of text". **Observations** 1) Posts using foul language in order to make a point related to the discussion topic were not met with any complaints from other members and were common. 2) Moderators would frequently remark that they changed the title of the thread to match the actual topic of the thread. 3) Members who started discussion threads with minimal discussion value were personally attacked ("hey <<expletive>>" or "why was this thread started?").		**Members directly responding to posts related to group boundaries** 1) Other threads involved members directly commenting about the hubris and egos in the programming profession. For example, "I don't want to go into the programming forums because I am scared someone will think I am dumb". 2) Comments such as "did you try posting that in one of the java forums where it belongs" or "take this thread to the [off-topic forum]". 3) Statements such as "If this were to be posted in the HTML/CSS section then that action would be warranted if explicitly requested". **Observations** 1) Links to outside sources were deleted and other members would complain about such posting activity. 2) Threads concerning malware or viruses revealed that these topics were off limits in the TPC Cubicle.

The two PhD researchers who determined the culture-in-interaction for the forum did so by observing what people did in terms of their posts (i.e., use of language, types of topics, and types of responses) and by observing what other members said concerning specific posts and their posting styles. For example, members would often respond to posts written in "textspeak" format or posts with excessive emoticons with comments such as "we don't do that around these parts" or "what's up with all of the emoticons?". Members would often explicitly "call out" other members when they did something that was out of the norm. Some of the conclusions, however, came from observations and not direct commentary about a behavior. For example, the two PhD researchers who determined the culture-in-interaction for the forum observed and made conclusions based on member commentary concerning threads written about malware or about writing viruses. After the two PhD researchers made all of the conclusions concerning each dimension of culture-in-interaction, I discussed the findings with individuals who regularly participate in online discussion forums and with theoretical academics in order to challenge and critically evaluate each culture-in-interaction dimension. This exercise resulted in another iteration of analysis in order to refine the elements of each culture-in-interaction dimension. Table 3 displays the elements of each culture-in-interaction dimension. Of the 1,805 posts that I used to determine each culture-in-interaction dimension, 144 violated a group bond, 163 violated a communication norm, 112 violated a group boundary, and 1,691 did not violate any culture-in-interaction dimension.

After determining each culture-in-interaction dimension, three PhD researchers coded each post that received a rating against each culture-in-interaction dimension in a binary manner. Each dimension was independent and a post could violate none or any combination of the three dimensions. The three coders coded roughly 10 percent of the posts that received a rating together in order to define and refine the systematic coding process, but they found that they needed to further refine the process so they coded an additional 20 posts together. This process resulted in 73 total posts coded together. I then created a random sample of 100 posts that received ratings that all three coders analyzed and coded independently in order to statistically evaluate the reliability of the process (see Table 4). During this step, the three coders discussed and resolved any coding discrepancies collectively. If, after discussion, all three coders did not come to an agreement, then they used a simple majority to reconcile the differences. I then randomly assigned the remaining posts that received ratings to one of the three coders. This coding process resulted in 33 posts coded with a group violation, 54 posts coded with a communication norm violation, and 35 posts coded with a group boundary violation. Finally, I summed up the binary values to determine the overall culture-in-interaction for the particular post. The final variable had a scale of 0 (no violations) to 3 (violations of each culture-in-interaction dimension).

Table 4. Reliability Statistics for Coded Variables

	Sample size	Cohen's kappa (Simple)			Krippendorff's alpha
		Coders 1 & 2	Coders 1 & 3	Coders 2 & 3	
Culture-in-interaction[1]					
Group bond violations	100	0.78	0.84	0.94	0.85
Communication norm violations	100	0.90	0.91	0.81	0.87
Group boundary violations	100	0.94	0.81	0.86	0.87
Contribution quality[2]					
Accuracy	100	0.90	0.87	0.77	0.84
Completeness	100	0.89	0.87	0.86	0.88
Verifiability	100	0.85	0.82	0.92	0.86
Website-specific factors[3]					
Humor	100	0.91	0.92	0.83	0.89
Back on topic	100	0.90	0.82	0.90	0.88
Programming paradigm	50	0.87	0.91	0.79	0.86

[1] The final culture-in-interaction variable was a sum of the binary values of each of the coded dimensions. Therefore, I calculated the reliability statistics for each coded dimension and not for the aggregated variable.
[2] These variables had three possible values (-1, 0, and 1), so these Cohen's Kappa values were weighted (not simple). The final quality variable was the sum of the three coded dimensions, so I calculated the reliability statistics for each coded dimension.
[3] The final website-specific variable was the sum of the two coded dimensions, so I calculated the reliability statistics for each coded dimension.

4.2 Occupational Culture and Degree of Difference

Through consultation with programmers and technical academics, I used three dimensions (see Table 5) to determine each member's occupational culture: 1) the programming paradigm a particular programmer subscribes to, 2) the programmer's software ideology, and 3) the programmer's primary programming language. Programmers are often protective of their programming paradigm, ideological preferences, and dominant programming language (Barnett et al., 2012; Raymond, 1999; Stewart & Gosain, 2006), which make them effective dimensions for in-group and out-group comparisons between developers made in the context of positive and negative ratings.

A programming paradigm offers "linguistic abstractions and proof theories for expressing program implementations" (Spinellis et al., 1994, p. 191). Each programming paradigm differs in its underlying assumptions, the types of programmatic solutions available, and is a defining professional characteristic of a given developer (Spinellis et al., 1994). Three PhD researchers coded the personal profiles of each member to determine each member's programming paradigm using the same process that was used to determine the culture-in-interaction violations. They first coded 25 member profiles together to define and refine the process. Then, I created a random sample of 50 profiles that each coder independently coded in order to statistically evaluate the process, which revealed a highly reliable coding process (see Table 4). During this step, any coding discrepancies were discussed and resolved collectively. If, after discussion, all three coders did not come to an agreement, then they used a simple majority to reconcile the differences. Finally, I randomly assigned the remaining profiles to one of the three coders.

Trice and Beyer (1993, p. 33) define ideology as "shared, relatively coherently interrelated sets of emotionally charged beliefs, values, and norms that bind some people together and help them make sense of their worlds." Examples of two programming ideologies are the commercial approach typically linked with the Microsoft community and the open source approach typically linked with the Linux community (Raymond, 1999). In my study, I determined a programmer's dominant ideology using the operating system preference field in each member's personal profile. I chose this field as the ideology proxy for two primary reasons. First, programmers who take the time to list an operating system preference subscribe, on at least some level, to the set of principles governed by the company that developed the operating system (even if they do so begrudgingly). I confirmed this idea through consultation with programmers. Second, the operating system preference field is visible when logged-in members read and rate posts.

Table 5. Sample of Values for Each Dimension of Programming Occupational Culture

Dimension	Sample values	Coding details	Descriptive statistics
Programming paradigm: captures differences in conceptualizing problems and general problem classification (Spinellis et al., 1994).	Imperative, object oriented, functional, multi-paradigm	A combination of the "about me" field in the member's profile, the programming language field in the member's profile, and posting activity at the website (tutorials, code snippets, and programming forum activity).	There were 131 developers coded as objected oriented, 37 as imperative, and 39 as multi-paradigm in the final sample.
Ideology: captures differences in taken-for-granted assumptions associated with different programming ideologies (Raymond, 1999; Stewart & Gosain, 2006)	Microsoft, Linux, Mac, ideology neutral	Operating system preference field in the member's personal profile, which is visible when logged-in members view contributions.	In the final sample, 114 developers identified with the Microsoft ideology, 49 with the Linux ideology, 16 with the Macintosh ideology, and 28 had no ideology preference.
Programming languages: captures idiosyncratic differences associated with each programming language (Van Roy & Haridi, 2004)	Java, C#, C++, C, VB .NET, VB 6.0, PHP, and so on	Which forum the member spent most of their time based on posts displayed in the "view posts" link off of each member's personal profile page.	In the final sample, there were nine different programming languages: C (19), C++ (30), C# (50), Clojure (1), Java (61), PHP (24), Python (2), VB .NET (15), and VB 6.0 (5).

Any combination of the three dimensions is a member's occupational culture. For example:

1. An *imperative* developer subscribing to the *Microsoft* ideology whose primary language is *VB .NET.*
2. A *multi-paradigm* developer subscribing to the *Linux* ideology whose primary language is *Java.*
3. An *object-oriented* developer subscribing to the *MAC* ideology whose primary language is *C++.*

There were 50 unique combinations of the three dimensions among the 207 members in my final sample.

Each programming language has its own nuances and implementations of certain paradigmatic features, which makes creating a direct link between languages and paradigms problematic (Van Roy & Haridi, 2004). For instance, constructors (subroutines called when objects get created) written in Java are similar to but not identical to constructors written in, say, C++ or C# even though they are all object-oriented programming languages. Programmers are also highly protective of their dominant programming language, which makes them important elements for minimal in-group and out-group comparisons (Van Roy & Haridi, 2004). I confirmed this dominant programming language dynamic by consulting with software developers. In this paper, I determined each member's primary programming language by the programming forum where the member had the majority of their posts.

After determining each member's occupational culture along these three dimensions, I compared the rater and the author of the post receiving the rating on each dimension (see Equation 1).

$$\sum_{k=1}^{3} if((Value\ for\ Receiver\ of\ the\ Rating) = (Value\ of\ the\ Rater), 1,0) \quad (1)$$

A difference score of 3 means that the two programmers had differences on all three dimensions and a difference score of 0 means that the two programmers were the same on all three dimensions.

4.3 National Culture Differences and Hofstede's Dimensions of Culture

There is considerable debate in the literature about measuring national culture (Kirkman, Lowe, & Gibson, 2006; McCoy, Galletta, & King, 2005; Sivakumar & Nakata, 2001). Some scholars argue that culture, particularly Hofstede's dimensions, should be measured at the individual level of analysis (Brockner, 2005; Srite & Karahanna, 2006), whereas other scholars including Hofstede are adamantly opposed to measuring culture at the individual level (Bochner & Hesketh, 1994; Hofstede, 2001; Palich, Horn, & Griffeth, 1995). Part of the contention rests on the definition of Hofstede's dimensions. For instance, if power distance is defined as a property of the culture (i.e., Korea is a high power distance culture), then measuring power distance at the individual level may be misleading. However, if power distance is defined as an individual's perception of status inequalities, then measuring power distance at the individual level is justified. Hofstede (1980) is clear that the methodological and theoretical underpinnings of his cultural dimensions mean that researchers should use his dimensions at the country level and not at the individual level. At the national level, Hofstede's study has been extensively replicated in many contexts using different samples at different points in time, which has led multiple researchers to conclude that Hofstede's reported country scores meet rigorous standards of reliability and validity (Geletkanycz, 1997; Kogut & Singh, 1988; Shackleton & Ali, 1990). Therefore, I use Hofstede's reported values for each cultural dimension for each member's identified country similar to Geletkanycz (1997), Mitchell et al. (2000), and many others.

I used the country of origin field in each member's profile as the national culture proxy, which I then mapped to Hofstede's values for each of his six cultural dimensions. This field, just like the operating system preference field, was visible when logged-in members read and rate posts, so raters could visibly see the content contributor's country of origin when making rating decisions. The final sample contained 29 different countries, but most of the developers were from the US (125) and Great Britain (27). I mean-centered Hofstede's dimensions in all models in order to reduce the skewness of the values. To determine the degree of national cultural differences between the receiver of the rating and the rater's national culture, I used Kogut and Singh's (1988) measure of cultural distance (see Equation 2).

$$CD_{ij} = \sqrt{\sum_{k=1}^{K} \{(I_{kj} - I_{ki})^2 / V_k\}}, \quad (2)$$

where I_{kj} is the rater's country score on the kth cultural dimension (power distance, uncertainty avoidance, individualism-collectivism, masculinity-femininity, long-term orientation, and indulgence), I_{ki} is the contributor's country score on the kth cultural dimension, and V_k is the variance of the kth cultural dimension. I calculated CD_{ij} for each rating interaction in the final sample.

4.4 Contribution Quality and Status of the Contributor

Prior literature has established contribution quality and the status of the contributor (Matthew effect) as the two primary factors that impact ratings in online feedback systems. Therefore, I control for the main effects of these two variables in all of my analysis. Both are also included in a hypothesized interaction effect with a cultural variable.

I determined contribution quality using the dimensions of quality defined by Fichman (2011). Fichman (2011) examined contribution quality based on the accuracy (correctness of a statement), completeness (thoroughness of the statement), and verifiability (proving a reference to a source) of each answer in her study. All three dimensions, however, were not applicable to each post in my sample. For instance, a post about Microsoft releasing a product that it did not release could be coded as being factually inaccurate and as being verifiable (i.e., this claim can be verified via a quick internet search). However, the completeness of this post is not relevant. In other instances, a member may post a well-articulated post that details the member's opinion about the relevance of certifications for programming careers or a detailed description of a personal experience. The accuracy and verifiable nature of the member's opinion concerning certifications and the member's personal story are not relevant, but the completeness of the response is relevant. Therefore, I coded each post that received a rating along each of the three dimensions of quality using a three-point scale (-1 for low quality on the dimension, 0 for neither high nor low quality (or not applicable), and 1 for high quality on the dimension). I followed the same coding process that I used to code the culture-in-interaction variables with these quality variables (see Table 4 for reliability statistics). After the three PhD researchers coded each post that received a rating along the three dimensions of quality, I summed the values across all three dimensions, which resulted in a scale from -3 (low quality on all three dimensions) to +3 (high quality on all three dimensions).

I determined the status of the member whose contribution was rated in two steps. First, I calculated each member's score (sum of total positive ratings minus sum of total negative ratings) in the feedback system as of 15 April, 2011, which was the midpoint of the rating observation period. Second, I determined the status grouping associated with each score based on TPC's clustering of membership. During my study's observation period, TPC clustered members with similar scores together (see Table 6). I used the numerical value instead of the absolute score in the data analysis due to excessive skewness of the raw scores.

Table 6. Status of Content Contributor Based on Scores in the Feedback System

Numerical value	Name	Score greater than or equal to	Score less than	Number of members in the sample
1	Exiled	Not bounded	-20	0
2	Disgraced	-20	-10	2
3	Shunned	-10	-5	2
4	Dishonored	-5	0	2
5	Apprentice	0	5	62
6	Worker	5	10	16
7	Tradesman	10	25	20
8	Craftsman	25	50	17
9	Whiz	50	100	21
10	Stalwart	100	250	24
11	Architect	250	500	17
12	Enlightened	500	750	8
13	Master	750	1000	3
14	Grandmaster	1000	2500	10
15	Guru	2500	5000	3

4.5 Control Variables

Many additional non-quality factors may logically result in either a positive or a negative rating (even though they have not been explicitly investigated in the prior literature), so I controlled for the following factors in my study: 1) website-specific factors (humor and back on-topic posts), 2) prior positive and negative ratings between the rater and the member whose contribution was being rated, 3) friendship between the rater and content contributor, 4) the number of prior positive and negative ratings on the post (contribution anchor), 5) position of the post in the thread (i.e., post near the beginning, middle, or end), and 6) the other Hofftstede (2001) dimensions of the rater (discussed previously).

The TPC Cubicle had website-specific factors that do not fall into any of Fichman's (2011) dimensions of quality but nevertheless impacted the propensity that a contribution would receive a positive or a negative rating. These factors included humor and back-on-topic posts. For example, certain threads may have begun discussing one topic and ended up discussing a different topic, and a member may have written a respectful post that attempts to bring the discussion thread back to its original purpose (back-on-topic posts). In the TPC Cubicle, members frequently commented about their appreciation for those types of posts and often rewarded those posts with positive ratings. I followed the same coding process that I used to code the culture-in-interaction variables with these two website-specific variables. These two variables were highly correlated (i.e., humorous posts were often posts that attempted to bring a thread back on topic in my sample of posts), so I aggregated these variables into a single website-specific variable that had a minimum value of 0 (neither humorous nor back-on-topic posts) and a maximum value of 2 (both humorous and back-on-topic posts).

I used two binary variables to account for the past rating history between two members: 1) previous negative rating and 2) previous positive rating. I determined both by examining all previous ratings (not time bound) for both the rater and the author whose contribution was being rated. If the rater was previously negatively (positively) rated by the current member who was being rated at any point previously, then I set the negative (positive) rating binary variable to 1 (otherwise it was set to 0). TPC also had a section where members can "friend" each other similar to Facebook and other social networking websites, so there was a distinct possibility that friends had a lower propensity to negatively rate each other. For each rating interaction, I checked the friend's section of the website to determine if the two members were friends. I set this binary variable to 1 if the two members were friends and to 0 if the two members were not friends.

In addition to anchoring on the high or low status of the post author, raters could anchor on previous ratings on a given contribution (post). To control for this alternative explanation, I calculated the number of previous negative ratings and the number of previous positive ratings on each post at the time of the rating interaction. Along these same lines, there could have been a "recency" effect whereby contributions at the start of a discussion thread may have had a more or less likelihood of being rated positively relative to posts at the end of a discussion thread. To control for this effect, I calculated a simple ratio of the position of the post relative to the number of posts (at the time of the rating interaction) in a given discussion thread.

5 Results

No single unit of analysis exists with a positive or a negative rating because each rating is embedded in the context of a post, the giver of feedback (rater), the receiver of feedback (author of the post being rated), and the national and occupational culture of both the rater and the receiver of feedback. As such, I used a cross-nested three-level hierarchical linear model (HLM) to analyze these data instead of a single level OLS regression (Snijders & Bosker, 2012). Figure 2 displays the three-level cross-nested structure that I used to analyze these ratings data. Although displayed in Figure 2 as dotted lines, there is no community culture in the final cross-nested HLM models because I only analyzed a single website (n = 1). Therefore, all posts, raters, and receivers of feedback were embedded in the same culture-in-interaction.

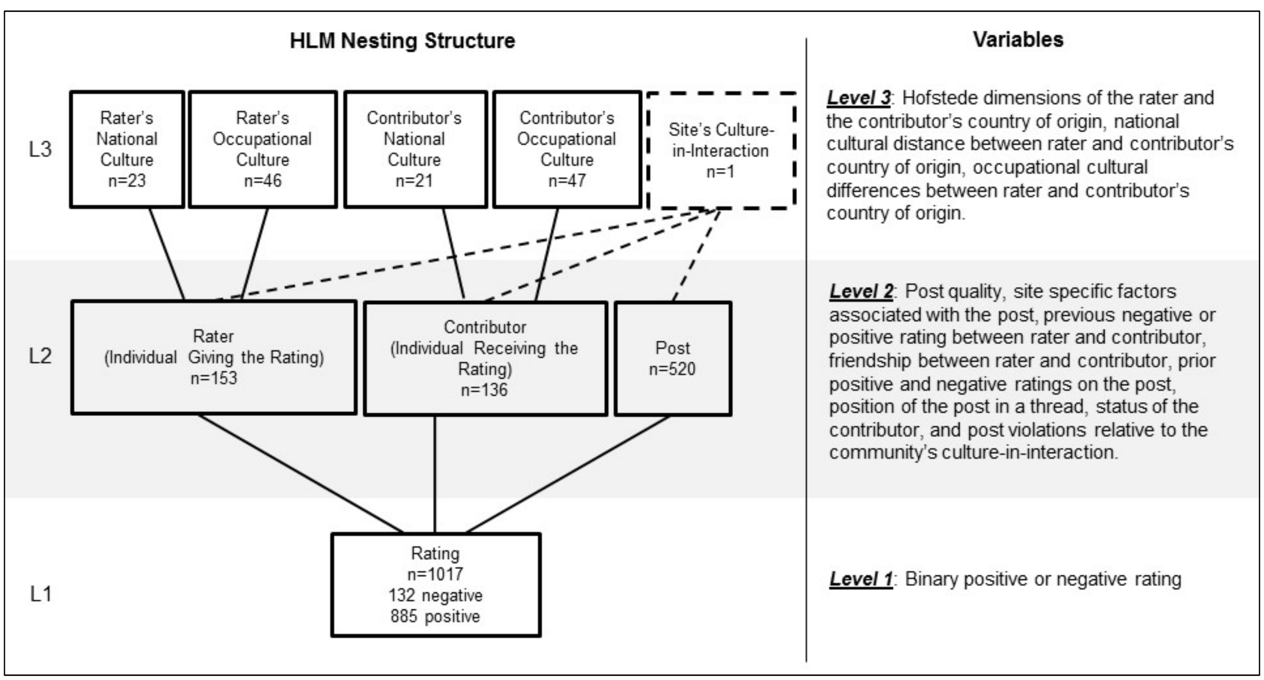

Figure 2. Three-level Cross-nested Structure

The dependent variable in all of the models was binary where a 0 represented a positive rating and a 1 represented a negative rating, so the dependent variable was the log odds of a negative rating (logit(P)). Therefore, I could not estimate the level-1 residual variances due to the binary nature of the dependent variable. However, I assumed that all of the level-2 and level-3 random effects (random intercepts only) had a multivariate normal distribution for each level-2 and level-3 factor (respectively), which is reasonable given the relatively large sample size and low standardized skew and kurtosis index values for the level-2 and level-3 continuous variables and random effects (Snijders & Bosker, 2012).

I used a residual pseudo-likelihood estimation technique to run each model in order to reduce the bias associated with determining fixed-effect estimates using maximum likelihood estimators. I also ran each model using a different variance component for each random effect and a between-within denominator degrees of freedom option, which splits the residual degrees of freedom into between-subject and within-subject portions. Finally, I used the iterative Newton-Raphson optimization technique to converge each model to a solution. This technique provides standard errors based on observed information as opposed to expected information in addition to using the shape of the likelihood function to find parameter values closer to the maximum (Snijders & Bosker, 2012).

Table 7 displays the descriptive statistics for the independent variables for the 1,017 rating interactions. Contribution quality and the status of the receiver of feedback had a low correlation, which means that members of all status groups contributed high-quality content in my sample. Therefore, any main effect of the contributor's status on the probability of a negative rating cannot simply be explained by simply concluding that higher status members of the community contribute "better" (higher-quality) contributions relative to lower status members of the community. The status of the receiver of feedback and culture-in-interaction violations also had a low correlation, which means high-status members of the community had a similar likelihood of violating the website's culture-in-interaction as low-status members of the community (in my sample). The Hofstede dimensions had the highest statistically significant correlations in my sample, but all variance inflation factors (VIFs) were lower than the common cutoff rules of thumb (O'Brien, 2007).

Table 7. Descriptive Statistics at the Rating Level

Variable (abbreviation)	Mean	Std. deviation
Culture-in-interaction violations	0.29	0.56
Status of the contributor	10.7	3.2
National culture distance	0.88	1.26
Occupational cultural differences	1.3	0.72
Friends[1]	0s = 836 1s = 181	
Prior positive rating relationship[1]	0s = 674 1s = 343	
Prior negative rating relationship[1]	0s = 946 1s = 71	
Post positioning	0.40	0.32
Number of prior positive ratings on the post	1.30	3.20
Number of prior negative ratings on the post	0.21	1.00
Site-specific factors	0.13	0.34
Contribution/Post quality	0.90	1.25
Power distance (PDI)[2]	0.13	7.71
Individualism-collectivism (IND)[2]	-4.91	14.5
Masculinity-femininity (MAS)[2]	-2.57	11.59
Uncertainty avoidance (UAI)[2]	-0.40	9.84
Long-term orientation (LTO)[2]	5.57	10.80
Indulgence (INDULG)[2]	-0.53	5.56

[1] For all binary variables, the 0s represent "no" and the 1s represent "yes".
[2] Mean centered variable so the descriptive statistics were for the mean-centered values.

Table 8 displays the solutions for all HLM regression models that I used to test my research hypotheses.

Table 8. HLM Model Results

	Model 1 method=rspl type=vc ddfm=bw	**Model 2** method=rspl type=vc ddfm=bw	**Model 3** method=rspl type=vc ddfm=bw	**Model 4** method=rspl type=vc ddfm=bw	**Model 5** method=rspl type=vc ddfm=bw	**Model 6** method=rspl type=vc ddfm=bw	**Model 7** method=rspl type=vc ddfm=bw
Fixed effects	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
Intercept	-2.43***	-2.07***	0.42	-2.31*	-3.06*	-2.38*	-2.30*
Culture-in-interaction (CINI)				1.99***	3.19***	2.21***	2.03***
National cultural distance				0.35*	0.35*	0.33	0.33
Occupational cultural distance (OCD)				0.57*	1.02**	0.59*	0.58*
(CINI) * (OCD)					-0.82*		
(CINI) * (post quality)						-0.73*	
(Receiver status) * (PDI)							0.02*
Receiver status			-0.25**	-0.17*	-0.16*	-0.18*	-0.17*
Prior + ratings on the post			-0.15	-0.11	-0.09	-0.10	-0.10
Prior - ratings on the post			0.49**	0.36*	0.40*	0.34*	0.40*
Post quality		-0.83***	-0.80***	-0.70**	-0.70**	-0.40	-0.73**
Friends			-2.31**	-2.62**	-2.84**	-2.69**	-2.27**
Post position			-0.22	-0.24	-0.27	-0.13	-0.16
Prior + rating relationship			-0.48	-0.55	-0.09	-0.10	-0.75
Prior - rating relationship			1.83**	1.99**	0.40*	0.34*	1.96**
Website-specific factors			-0.007	0.002	-0.07	-0.02	-0.01
PDI			-0.11	-0.09	-0.09	-0.10	-0.28*
IND			-0.04	-0.04	-0.04	-0.05	-0.05
MAS			0.06	0.08*	0.07	0.08*	0.08*
UAI			0.04	0.05	0.05	0.04	0.04
LTO			-0.01	-0.01	-0.008	-0.01	-0.02
INDULG			-0.08	-0.02	-0.02	-0.02	0.004
* Significant at 0.05, ** Significant at 0.01, *** Significant at < 0.001							
Random effects	Variance component	Variance component	Variance component	Variance component	Variance component	Variance component	Variance component
L2 intercept (receiver)	1.59	1.41	0.46	0.12	0.17	0.11	0.12
L2 intercept (rater)	0.06	0.12	0.16	0.22	0.22	0.22	0.20
L2 intercept (post)	1.24	0.93	1.11	0.95	0.92	0.93	1.00
L3 intercept (rater national culture)	0.05	0.10	0.16	0.19	0.20	0.21	0.18
L3 intercept (rater occ. culture)	0.10	0.10	0.19	0.19	0.20	0.21	0.18
L3 intercept (receiver national culture)	0.44	0.57	0.29	0.18	0.11	0.19	0.20
L3 intercept (receiver occ. culture)	0.60	0.57	0.28	0.11	0.12	0.12	0.13
Model fit							
-2 residual log pseudo-likelihood	5,484.35	5,782.41	6,350.88	6,697.22	6,842.84	6,678.94	6,763.29
Explained variance[1]		10.4% 17.8%	25.1% 42.8%	33.7% 57.7%	34.0% 58.1%	33.4% 57.1%	33.1% 56.6%

[1] I calculated the top percentage estimate using the following formula proposed by Snijders and Bosker (2012): 1 - sum(variance components for the Model + (π^2)/3) / sum(variance components for the Null model + (π^2)/3). I calculated the bottom percentage by removing the (π^2)/3 level-one logit(P) estimator from the formula.

Due to the binary nature of the outcome variable, I calculated the nesting effect (a pseudo intra-class correlation coefficient (ICC)) by estimating the variance of the logistic function (level 1) using (π^2)/3 and using the squared covariance parameter estimates for levels 2 and 3 (Snijders & Bosker, 2012). Using this method with the unconstrained null model (Model 1), the level-1 rating (41 percent), the level-2 receiver of the rating (32 percent), and the level-2 post (19 percent) had the greatest ICCs. The overall level-2 ICC in the unconstrained model was roughly 51 percent and the overall level-3 ICC was 7 percent, which is evidence supporting the use of this nesting structure. Model 1 also indicated that the probability that a random post posted by a random member from a random occupational culture and random national culture would receive a negative rating (relative to a positive rating) from a random rater from a random occupational culture and random national culture was 8.1 percent[4].

In order to explain rating behaviors above, beyond, or instead of contribution quality using cultural variables at the national, occupational, and community levels, I first had to establish a baseline for the quality effect in rating practices. Model 2 provided this baseline by regressing logit(P) of negative ratings on just the post quality independent variable. Model 2 had a pseudo R^2 of approximately 10 percent. This result means that, in my sample of ratings, roughly 10 percent of the variance in logit(P) of negative ratings relative to positive ratings could be explained by just contribution quality. Model 2 also demonstrated empirically that the quality variable was statistically significant in the proper direction (i.e., low-quality posts had a higher probability of being negatively rated relative to high-quality posts). Therefore, quality posts had the desired effect, but the model only explained approximately 10 percent of the sample variance in rating practices, so there were many other factors involved in determining why a post received either a positive or a negative rating.

Model 3 contained the remainder of the control variables. This model had a pseudo R^2 of roughly 25 percent. Adding in the culture variables in Models 4-7 added between 8 and 9 percent of additional explained variance. Model 4 tested the H1, H2, and H3 main effects. For a post that did not violate the culture-in-interaction of the forum while holding all other covariates constant at 0, there was a 9 percent probability that the post would get negatively rated. However, a post that violated a single culture-in-interaction dimension (increasing the culture-in-interaction variable from 0 to 1) while holding all other covariates constant at 0 increased the probability of a negative rating from 9 to 42 percent. This probability increased to 97 percent when a post violated all three culture-in-interaction dimensions. Therefore, the Model 4 results supported the hypothesized main effect of culture-in-interaction violations on the probability of a negative rating interaction (H1).

Model 4 also supported the hypothesized effect of occupational cultural differences between the rater and the receiver of the rating on the probability of a negative rating (H2), but the effect was not as strong as with posts that violated the culture-in-interaction of the forum. When the rater and the receiver of the rating differed on a single occupational cultural dimension while holding all other covariates constant at 0, the probability of a negative rating increased from 9 to 14.9 percent. The probability of a negative rating increased to 35 percent when the rater and the receiver of the rating differed on all three occupational cultural dimensions (while holding all other covariates constant at 0).

Additionally, Model 4 supported the hypothesized effect of national cultural distance between the rater's national culture and the rating receiver's national culture on the probability of a negative rating (H3). Increasing the national cultural distance between the rater's and the rating receiver's national culture by one unit while holding all other covariates constant at 0 increased the probability of a negative rating from 9 to 12 percent. A two-unit increase in cultural distance increased the probability of a negative rating to 16.6 percent (while holding all other covariates constant at 0).

However, the occupational cultural differences variable and the culture-in-interaction violation variable were part of a significant higher-order interaction effect (H4) in Model 5 (see Figure 3). In Model 5, the effect of small occupational cultural differences on the probability of a negative rating was enhanced when the post violated one or more of the culture-in-interaction dimensions. For instance, when the occupational cultural differences between the rater and the receiver of the rating were along a single dimension (difference score of 1) while holding all other covariates constant at 0, the probability of a negative rating was 5.5 percent when the post did not violate any culture-in-interaction dimension. However, that 5.5 percent probability increased to 38 percent when the post violated a single culture-in-interaction dimension and 87 percent

[4] I calculated all reported probabilities in this section using both the intercept and the parameter estimates. I had to include the intercept because that provides the baseline when the variable of interest was 0. The formula is as follows: exp(intercept + (parameter estimate)*(value)) / (1+ exp(intercept + (parameter estimate)*(value))).

when the post violated two culture-in-interaction dimensions. In other words, small occupational cultural differences went relatively unnoticed until the post violated one or more of the culture-in-interaction dimensions. Therefore, Model 5 provided empirical support for the H4 moderating hypothesis.

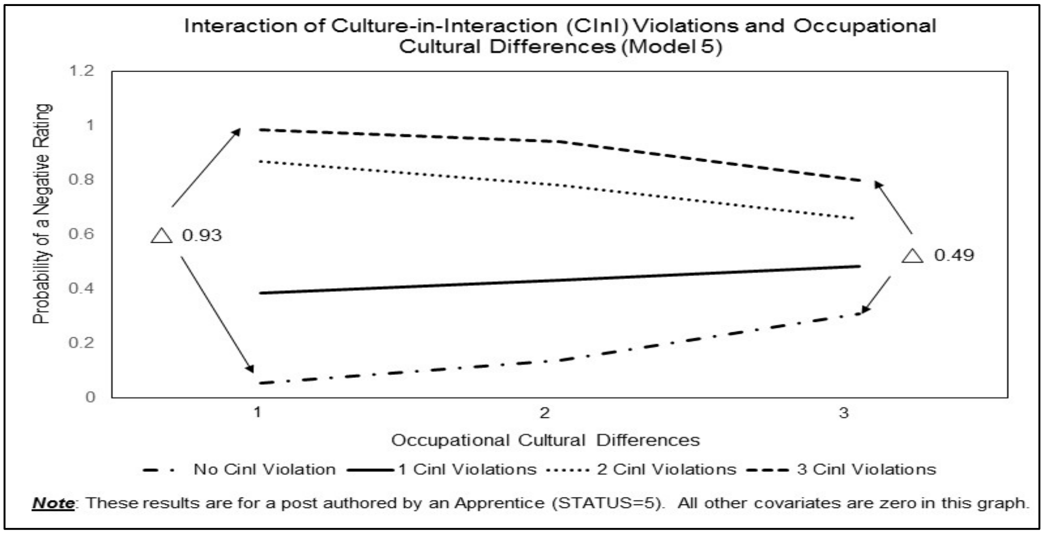

Figure 3. Interaction of Culture-In-Interaction Violations and Occupational Cultural Differences

Models 2-4 established the negative main effect of contribution quality on the probability of negative ratings, but Model 6 demonstrated that post quality was part of a statistically significant higher-order interaction effect with the culture-in-interaction violation variable (see Figure 4). In this model, low-quality posts (quality score of -2) had a 7.7 percent probability of a negative rating when the post did not violate the culture-in-interaction of the website (while holding all other covariates at 0). However, once that low-quality post (quality score of -2) also violated one of the culture-in-interaction dimensions (increasing the culture-in-interaction variable from 0 to 1), the probability of a negative rating increased to 76.6 percent (while holding all other covariates at 0). The effect was in the same direction for high-quality posts but the magnitude of the effect was less dramatic. For instance, a high-quality post (quality score of 2) without a culture-in-interaction violation had a 1.7 percent probability of a negative rating, but, when that same high-quality post violated all three culture-in-interaction dimensions, the probability of a negative rating increased to 13.8 percent (while holding all other covariates at 0). Therefore, Model 6 provided empirical support for the H5 moderating hypothesis.

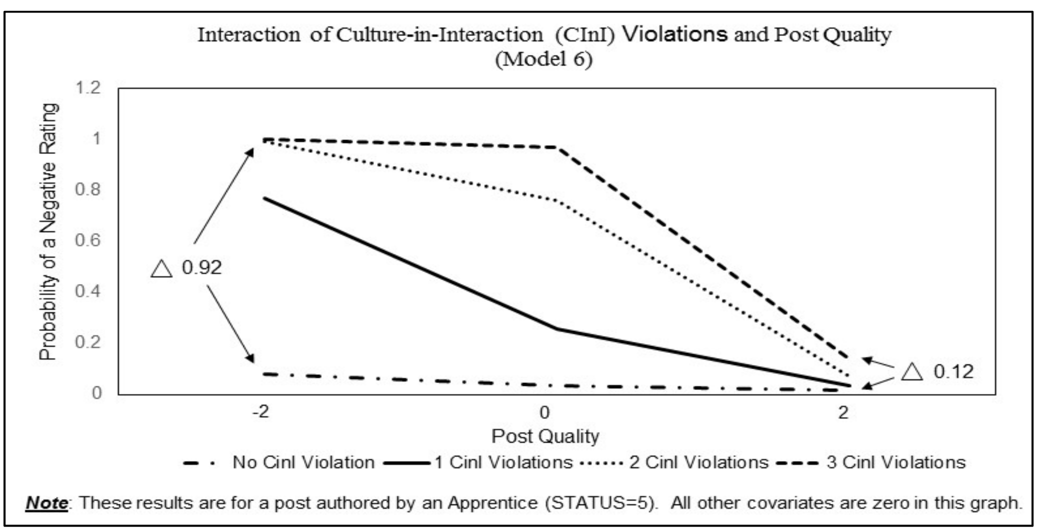

Figure 4. Interaction of Post Quality and Culture-in-Interaction Violations

Models 3 and 4 demonstrated that the rating receiver's status in TPC was negatively correlated with the probability of a negative rating. This negative correlation means that the higher the receiver's status in TPC, the less likely the receiver was to receive a negative rating even after controlling for contribution quality.

This finding provides strong evidence supporting the contributor anchor and resulting Matthew effect. Model 7, however, further demonstrated empirically that this Matthew effect was qualified by the power distance of the rater's national culture (see Figure 5). Raters from low power distance national cultures relative to high power distance national cultures had a greater probability of giving a negative rating even after controlling for the quality of the contribution, previous rating history, friendship between the two individuals involved in the rating interaction, and previous ratings on the contribution.

For instance, raters from low power distance national cultures had a 74 percent probability of giving a negative rating to a low-status content contributor (member of the disgraced status group), whereas raters from high power distance national cultures had less than a 1 percent probability of giving a post authored by that same member of the disgraced status group a negative rating (while keeping all other covariates constant at 0). For posts authored by high-status members of the community (member of the Stalwart group), raters from high power distance national cultures had almost a 0 percent probability of negatively rating them, whereas raters from low power distance national cultures had roughly a 2 percent probability of negatively rating them (while keeping all other covariates constant at 0). In these data, all raters showed a deference to high-status content contributors (even after controlling for content quality and the other control variables), but raters from low power distance national cultures showed less deference relative to raters from higher power distance national cultures. Therefore, the hypothesized moderating effect of power distance was greater for low-status content contributors (poor getting poorer) than for high-status content contributors (rich getting richer), but Model 7 still generally support the hypothesized effect.

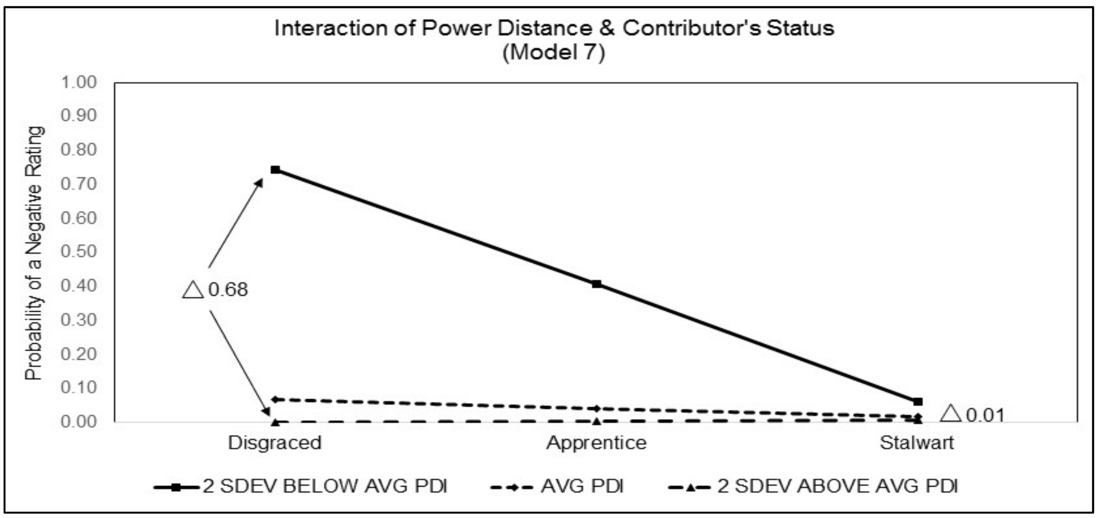

Figure 5. Interaction of Power Distance and Contributor's Status

Three of the control variables (besides contribution quality and receiver status) were consistently significant with meaningful effect sizes. First, friendship between the rater and the receiver of the rating had a strong negative effect. In Model 3, for instance, friends had roughly a 13 percent probability of negative rating each other but non-friends had close to a 60 percent probability of negatively rating each other (while keeping all other covariates constant at 0). Second, I found a strong positive effect when the rater and the receiver of the rating had a prior negative rating relationship. In Model 4, for instance, a prior negative rating relationship resulted in a 42 percent probability of a negative rating versus only a 9 percent probability when there was not a previous negative rating relationship (while keeping all other covariates constant at 0). Therefore, members used the feedback system to extract revenge from prior negative ratings even when controlling for contribution quality, but prior positive ratings did not have the reverse effect. Finally, all models showed statistically significant evidence of a negative contribution (post) anchoring effect. Prior negative ratings on a contribution (post) resulted in an increased probability of additional negative ratings on a contribution even after controlling for contribution quality. Table 9 shows the conclusions in relation to the stated hypotheses even when controlling for these robust set of alternative explanations.

Table 9. Hypothesis Conclusions

Hypothesis	Conclusion
H1: Culture-in-interaction violations (main effects)	Supported in Models 4-7 but the main effect was qualified by occupational cultural differences (Model 5).
H2: Occupational culture differences (main effects)	Supported in Models 4-7 but the main effect was qualified by culture-in-interaction violations (Model 5).
H3: Cultural distance between rater's and contributor's national culture	Supported in Models 4 and 5 at the 0.05 level and weakly supported at the 0.1 level in Models 6 & 7.
H4: Interaction of culture-in-interaction violations and occupational culture differences	Supported in Model 5.
H5: Interaction of culture-in-interaction violations and contribution quality	Supported in Model 6.
H6: Interaction of power distance and contributor's status	The moderating effect of power distance was greater for low-status content contributors (poor getting poorer) than for high-status content contributors (rich getting richer). Regardless, Model 7 generally supported this moderating hypothesis.

5.1 Alternative Interaction Effects

I tested several alternative interaction effects to rule out possible alternative explanations. Due to space limitations, Table 8 omits these models. First, it is possible that quality posts authored by high-status members of TPC were rewarded more than quality posts authored by low-status members of the community. If true, then it would qualify the conclusions that I made concerning the quality by culture-in-interaction violation moderating H5. To test this alternative explanation, I ran a model with all variables from Model 4 plus an interaction effect of contribution quality and the rating receiver's status. This model revealed a non-statistically significant interaction effect. This non-statistically significant interaction effect means that my sample of raters rewarded high-quality posts with a reduced probability of receiving a negative rating (relative to a positive rating) in a consistent manner for members across all status groups (from Exiled to Guru). Second, to further evaluate the robustness of the hypothesized culture-in-interaction by quality moderating effect, I tested a three-way interaction between contribution quality, the status of the content contributor, and culture-in-interaction violations with all of the variables from Model 4. This three-way interaction effect in this model was not statistically significant, which means that status was not a confounding factor to the hypothesized two-way interaction effect of contribution quality and culture-in-interaction violations.

Third, to evaluate whether contribution quality was a confounding factor for the power distance moderating hypothesis (H6), I tested a model with a three-way interaction effect (power distance, contributor status, and contribution quality) with all variables from Model 4. This model revealed a non-significant three-way interaction effect. Therefore, the significant interaction effect shown in Figure 5 (Model 7) cannot be explained by contribution quality differences between posts made by members of different status groups.

Finally, I tested several different interaction effects of each of the other Hofstede dimensions of national culture and contributor status or contribution quality in an exploratory manner. None of those two-way interaction effects (i.e., each Hoftsede dimension by quality and each Hoftstede dimension by contributor status interaction effects) were statistically significant. These results mean that power distance was the only significant national culture moderator in my sample of ratings.

6 Discussion and Conclusion

Feedback systems serve an important purpose for a wide variety of virtual environments from commercial to social to learning (Adler, Alfaro, Kulshreshtha, & Pye, 2011; Aral, 2014; Josang, Ismail, & Boyd, 2007). However, the usefulness of accumulated positive and negative ratings largely depends on how much noise (i.e., ratings not based on quality) the overall ratings contain (Chen et al., 2011; Khansa, Ma, Liginlal, & Kim, 2015; Poston & Speier, 2005). For instance, TripAdvisor, Yelp, or Amazon ratings that are based on personnel, social, or cultural factors and not on the quality of the product or service limit the usefulness of those reviews, which is problematic for both the buyer (not resolving seller uncertainty) and the seller (not being properly rewarded for having a superior product or service). In public electronic knowledge sharing communities, ratings that do not represent quality pose a significant risk for both the knowledge seeker and the knowledge contributor. The knowledge seeker may learn unsound practices concerning the skill-based craft due to the poor signal that the noisy ratings provide (Cheng & Vassileva, 2005), while highly knowledgeable content contributors may stop contributing due to their high-quality contributions' not being positively rated or even being negatively rated (Lampe & Resnick, 2004).

Yet, prior literature has reported minimal theoretical or empirical research explaining why something gets positively or negatively rated. In this paper, I suggest that cultural factors at a variety of levels are a significant source of rating's noise because different cultures may embrace or avoid conflict, may have different levels of respect for authority, and may have different attitudes towards publicly downgrading (negative ratings) or praising (positive ratings) other members in the feedback system. On Amazon, for instance, consumers from all over the world provide ratings, and, on www.codeproject.com, raters from many different national and occupational cultures rate each other. Given cultural differences on a variety of dimensions, it is not surprising that culture (at a variety of levels) may have an adverse direct and mitigating impact on rating practices.

The good news is that my results demonstrate empirically that quality is an important predictor of the probability that a contribution will receive negative and positive ratings. Quality by itself explained roughly 10 percent of the variance in rating practices with a strong effect size in the expected direction (i.e., low-quality posts had a higher probability of receiving negative ratings). However, the bad news is that my results also demonstrate empirically that the impact of quality on the probability of negative and positive ratings was significantly moderated by culture-in-interaction violations. Posts that violated the culture-in-interaction of the forum removed most of the quality effect. Additionally, a greater percentage of the explained variance in the probability (Logit(P) to be more precise) of negative and positive ratings were based on occupational and national cultural similarities or differences between the rater and the receiver of the rating, previous interaction histories between dyadic pairs of members, adherence or a lack of adherence to the forum's culture-in-interaction, friendships, the social status of the receiver of the rating, and the power distance of the rater's country of origin. The cultural variables explained between 8 and 9 percent of the variability in rating practices in my sample, while the other social factors (i.e., friendship, prior rating history, Matthew effect of the contribution, and the Matthew effect of the contributor) explained close to 15 percent of the variance in the reported models. Therefore, my study demonstrates empirically several sources of ratings noise including a variety of cultural effects.

6.1 Theoretical Contributions and Future Research

Much of the literature on online feedback systems in a variety of contexts has focused on explaining the impact that ratings in these systems have on a variety of social and economic outcomes (Ba & Pavlou, 2002; Chen et al., 2011; Chevalier & Mayzlin, 2006; Dellarocas, 2006; Godes & Silva, 2012; Pavlou & Dimoka, 2006; Poston & Speier, 2005; Son et al., 2006). Far less literature has investigated why something gets positively or negatively rated online. My paper helps to close this gap by using a cultural theoretical lens to explain a portion of the ratings noise, specifically in public electronic knowledge sharing communities. In order to understand the meaning of accumulated ratings, my results suggest that information consumers need to understand the culture-in-interaction of the community, the occupational culture that the community is situated in, and the national cultures of the membership. Culture at each of these levels interact in unique ways to either directly impact or qualify the impact of other factors (quality and the Matthew effect) on rating practices. However, just because an electronic environment is culturally diverse does not necessarily mean that my proposed research model will be applicable. As such, a fruitful area of future research could investigate these cultural effects in other culturally diverse contexts (eCommerce, electronic knowledge sharing in other industries, or social sites such as Facebook or Baidu) in order to test the generalizability of my findings (Lee & Baskerville, 2003).

These cultural factors involved in ratings suggest that referring to online feedback systems as "reputation systems" may be a bit of a misnomer, particularly in culturally diverse settings. By definition, reputation is a measure of quality (Fombrun & Shanley, 1990; Washington & Zajac, 2005), but my study demonstrates empirically that ratings in these systems are a function of many additional social and cultural variables. This empirical finding may be one of the reasons why the prior literature reports inconsistent findings when empirically evaluating the reputation effect of online ratings in a variety of different contexts (Ba & Pavlou, 2002; Ghose et al., 2006; Kauffman & Wood, 2005). If, for instance, quality explains only 20 percent of the variance in a rating, then it is not surprising that previous literature has reported conflicting results because 80 percent of the rating relates to factors not associated with reputation. Instead of or in addition to reputation, these feedback systems may be measuring status, popularity, in-group versus out-group dynamics, social and cultural similarities and differences, or something else entirely. My results suggest future research that investigates this type of reputation effect must first statistically determine what percentage of the ratings are due to quality and then use that component in their predictive models.

6.2 Practical Implications

Feedback systems in public electronic knowledge sharing communities rarely have algorithms that filter ratings in or out by, for instance, accounting for the social or cultural context surrounding the rating interaction. Instead, typical feedback systems opt for an unfiltered, uncensored approach whereby all members may participate and all ratings count towards a member's total score. On the surface, this strategy seems like an effective one, but my study demonstrates empirically that many positive and negative ratings are noisy. To reduce this noise, feedback systems could incorporate social and occupational cultural factors into their design. For example, feedback systems can implement an algorithm that determines the probability of negative or positive ratings between dyadic pairs of members based on the practice-related data entered in each member's profile and previous rating history of all members (as I did in my study but in an automated manner). Based on these probabilities, which will inevitably vary from website to website and will change as new ratings are entered into the system, the algorithm may reduce noisy ratings by only counting a portion of those ratings (based on the probabilities) towards a member's total score.

The strong effect of culture-in-interaction violations provides website administrators with a bit of a conundrum. This finding may, on the one hand, be a positive appropriation of the feedback system because the negative ratings may ensure that new members learn the social dynamics, tone of the discussion forum, group boundaries, and overall assumptions related to participation in the particular group before posting in a given forum. This type of system usage may prevent certain disruptions to the dynamics of the electronic knowledge sharing community due to members' not understanding what it means to be a part of the group. On the other hand, however, punishing culture-in-interaction violations this severely may make the electronic knowledge sharing community less open to outsiders because outsiders may be afraid to participate for fear of being publicly downgraded in the feedback system. Consequently, this type of usage of the feedback system may promote lurking or hinder the growth of certain electronic social structures. Website administrators may want to explicitly state the expectations of member conduct in the specific forum in an about page so newcomers can learn the community's culture without having to lurk for an excessively long period of time before posting.

6.3 Limitations

Like all research, my empirical investigation has limitations. First, I theorize about how these types of systems are being used (i.e., actual ratings) and not system non-usage (i.e., members' choosing not to rate a specific post). Therefore, all reported probabilities were for a negative rating relative to a positive rating. A given member, however, may choose not to rate a post (third option) for a similar or a different set of factors than the factors that I propose in this paper. For example, occupational cultural similarities and differences may only become salient after the decision to rate has already been made. For non-raters, these differences may not be noticeable. Investigating non-raters would be an interesting future study that would certainly add to my proposed model.

Second, I make no claims that the three elements in my operational definition of an individual's programming occupational culture represent a complete list. Programming paradigm, ideology, and dominant programming language represent three measurable dimensions associated with an individual's programming occupational culture. Other elements would certainly include (among others) the type of company, the types of projects, the software development model (agile versus waterfall), and the structure of the development environment that a programmer works in. Testing the effect of more dimensions such that the range of difference scores between dyadic pairs of members increases from zero to three to, say, from zero to eight or 10 would be a worthwhile future study. Regardless, my study does provide empirical evidence that occupational cultural differences along at least these three dimensions are important factors in predicting ratings.

Third, my study is situated in a single, diverse field in a single occupational culture. TPC has forums that span the spectrum of programming paradigms and ideologies in the occupational culture of software development. Consequently, no dominant occupational group exists in this public electronic knowledge sharing community because members have many different occupational cultural backgrounds. As such, it is reasonable to posit that the impact that occupational cultural differences has on rating practices may be greater in a website such as Microsoft's Channel 9 or the Java OTN Forums. For example, a Java/Linux developer may be met with much less cordiality at Channel 9 than at TPC due to Channel 9's being presumably dominated by .NET/Microsoft developers. Future research could look to empirically test and possibly extend my proposed model by investigating an electronic knowledge sharing community with clearly defined dominant groups of members based on their occupational cultures.

Fourth, my paper demonstrates that culture is a significant source of noise in ratings. Culture, however, is certainly not the only source of noise as demonstrated by the significant effects of several of the non-quality control variables. Future research could add to this model by investigating other social factors such as network structures and network positioning on the propensity to give and receive positive or negative ratings. Different network structures may promote the flow of information more efficiently than others (Reagans & McEvily, 2003; Reagans & Zuckerman, 2001), which may increase or decrease these cultural effects or may introduce additional main effects to this model. Future research could also further investigate culture-in-interaction violations by theorizing and testing the effects of each culture-in-interaction dimension separately instead of aggregating the violations to the culture-in-interaction level. These types of correlations may provide additional theoretical and empirical insights into the reasons why a contribution is rated the way that it is rated.

Finally, the website that I investigated did not allow either anonymous postings or anonymous ratings (at the time I conducted the study), but other websites may allow for anonymous posting and/or anonymous ratings. An interesting future study could investigate the potential mediating or moderating role that anonymity has on rating practices. For example, the impact of national cultural distance or occupational cultural differences on the probability of negative ratings may logically be mitigated by anonymity because not posting the names of the raters or the receivers of the ratings may reduce the impact that respect for authority or minimal group categorizations along occupational cultural dimensions has on rating practices. Furthermore, individuals from high power distance national cultures may be less worried about being deferential to high-status members in the community if the ratings were partially or completely anonymous. Investigating anonymity would certainly make for an interesting extension to my proposed cultural effects.

6.4 Conclusion

Why contributions get rated positively or negatively may often appear to be irrational and arbitrary, particularly in public electronic knowledge sharing communities. For example, consider the following exchange:

> Member 1: Posts a discussion topic or a question
> Member 2: Posts a well-articulated and insightful response
> Member 3: Posts a simple "I agree" with nothing new

In this exchange, member 2 may receive a combination of positive and negative ratings for their well-articulated and insightful contribution, while member 3 may receive mostly positive ratings for their rather bland contribution. Based on my theoretical analysis and empirical results, it could be that member 3 was rated more positively because the individuals who chose to rate member 3's contribution were occupationally similar to member 3, were from similar national countries, and/or were from high power distance national cultures. It is also possible that member 2's contribution violated the culture-in-interaction of the forum (even though it was a high-quality contribution), which may have resulted in a higher propensity for the contribution to receive negative as opposed to positive ratings. Furthermore, the members who rated member 2's contribution may have come from lower power distance cultures, which may have resulted in a greater likelihood that they would give negative ratings relative to positive ratings. My theoretical analysis and empirical results suggest that it is not enough to look at contribution quality in isolation when evaluating these ratings. It is further necessary to contextualize the rating in the context of social and cultural attributes of the rater and the receiver of the rating along with the culture-in-interaction of the forum.

References

Adler, B. T., Alfaro, L. D., Kulshreshtha, A., & Pye, I. (2011). Reputation systems for open collaboration. *Communications of the ACM, 54*(8), 81-87.

Allen, G. N., Burk, D. L., & Davis, G. B. (2006). Academic data collection in electronic environments: Defining acceptable use of Internet resources. *MIS Quarterly, 30*(3), 599-610.

Aral, S. (2014). The problem with online ratings. *MIT Sloan Management Review, 55*(2), 47-52.

Ardichvili, A., Maurer, M., Li, W., Wentling, T., & Stuedemann, R. (2006). Cultural influences on knowledge sharing through online communities of practice. *Journal of Knowledge Management, 10*(1), 94-107.

Arrigara, M., & Levina, N. (2008). *Social dynamics in online cultural fields*. Paper presented at the International Conference on Information Systems, Paris, France.

Ba, S., & Pavlou, P. A. (2002). Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly, 26*(3), 243-268.

Barnett, W. P., Goodman, S., & Stewart, D. (2012). *Ideological competition in status contests* (Stanford University Working Paper). Retrieved from http://faculty-gsb.stanford.edu/barnett/documents/IdeologicalCompetitioninStatusContestsJanuary2010v2.pdf

Blanchard, A. L. (2008). Testing a model of sense of virtual community. *Computers in Human Behavior, 24*(5), 2107-2123.

Bochner, S., & Hesketh, B. (1994). Power distance, individualism/collectivism, and job-related attitudes in a culturally diverse work group. *Journal of Cross-Cultural Psychology, 25*(2), 233-257.

Boland, R. J., Jr., & Tenkasi, R. V. (1995). Perspective making and perspective taking in communities of knowing. *Organization Science, 6*(4), 350-372.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*(2), 307-324.

Brewer, M. B., & Yuki, M. (2007). Culture and social identity. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 307-322). New York: Guilford Press.

Brockner, J. (2005). Unpacking country effects: On the need to operationalize the psychological determinants of cross-national differences. In B. M. Staw & R. L. Sutton (Eds.), *Research in organizational behavior* (pp. 335-369). Greenwich, CT: JAI Press.

Buchan, N. R., Croson, R. T. A., & Dawes, R. M. (2002). Swift neighbors and persistent strangers: A cross-cultural investigation of trust and reciprocity in social exchange. *American Journal of Sociology, 108*(1), 168-206.

Chen, J., Xu, H., & Whinston, A. B. (2011). Moderated online communities and quality of user-generated content. *Journal of Management Information Systems, 28*(2), 237-268.

Cheng, R., & Vassileva, J. (2005). Adaptive reward mechanism for sustainable online learning community. In *Proceedings of the Conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology* (pp. 152-159). Amsterdam, Netherlands: IOS Press.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345-354.

Chiou, J.-S., & Lee, J. (2008). What do they say about “friends”? A cross-cultural study on Internet discussion forum. *Computers in Human Behavior, 24*(3), 1179-1195.

Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science, 12*(4), 105-109.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591-621.

Clemons, E. K., Gao, G. G., & Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of Management Information Systems, 23*(2), 149-171.

De Cindio, F., Gentile, O., Grew, P., & Redolfi, D. (2003). Community networks: Rules of behavior and social structure. *Information Society, 19*(5), 395-406.

Dellarocas, C. (2006). How often should reputation mechanisms update a trader's reputation profile? *Information Systems Research, 17*(3), 271-285.

Dellarocas, C., & Wood, C. A. (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science, 54*(3), 460-476.

Dholakia, U. M., Bagozzi, R. P., & Pearo, L. K. (2004). A social influence model of consumer participation in network- and small-group-based virtual communities. *International Journal of Research in Marketing, 21*(3), 241-263.

Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European Review of Social Psychology, 1*(1), 263-292.

DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology, 32*(1), 271-297.

Earl, M. (2001). Knowledge management strategies: Toward a taxonomy. *Journal of Management Information Systems, 18*(1), 215-233.

Eliasoph, N., & Lichterman, P. (2003). Culture in interaction. *American Journal of Sociology, 108*(4), 735-794.

Erez, M., & Gati, E. (2004). A dynamic, multi-level model of culture: From the micro level of the individual to the macro level of a global culture. *Applied Psychology: An International Review, 53*(4), 583-598.

Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology, 113*(1), 1-40.

Fichman, P. (2011). A comparative assessment of answer quality on four question answering sites. *Journal of Information Science, 37*(5), 476-486.

Fleischer, A. (2009). Ambiguity and the equity of rating systems: United States brokerage firms, 1995-2000. *Administrative Science Quarterly, 54*(4), 555-574.

Fombrun, C., & Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *The Academy of Management Journal, 33*(2), 233-258.

Gallagher, S. E., & Savage, T. (2013). Cross-cultural analysis in online community research: A literature review. *Computers in Human Behavior, 29*(3), 1028-1038.

Gao, G. G., Greenwood, B. N., Agarwal, R., & McCullough, J. S. (2015). Vocal minority and silent majority: How do online ratings reflect population perceptions of quality? *MIS Quarterly, 39*(3), 565-589.

Geletkanycz, M. A. (1997). The salience of “culture's consequences”: The effects of cultural values on top executive commitment to the status quo. *Strategic Management Journal, 18*(8), 615-634.

Germonprez, M., & Hovorka, D. S. (2013). Member engagement within digitally enabled social network communities: New methodological considerations. *Information Systems Journal, 23*(6), 525-549.

Ghose, A., Smith, M. D., & Telang, R. (2006). Internet exchanges for used books: An empirical analysis for product cannibalization and social welfare. *Information Systems Research, 17*(1), 3-19.

Godes, D., & Silva, J. C. (2012). Sequential and temporal dynamics of online opinion. *Marketing Science, 31*(3), 448-473.

Gould, R. V. (2002). The origins of status hierarchies: A formal theory and empirical test. *American Journal of Sociology, 107*(5), 1143-1178.

Grace-Farfaglia, P., Dekkers, A., Sundararajan, B., Peters, L., & Park, S.-H. (2006). Multinational Web uses and gratifications: Measuring the social impact of online community participation across national boundaries. *Electronic Commerce Research, 6*(1), 75-101.

Hertel, G., & Kerr, N. L. (2001). Priming in-group favoritism: The impact of normative scripts in the minimal group paradigm. *Journal of Experimental Social Psychology, 37*(4), 316-324.

Hofstede, G. (1980). *Culture's consequences: International differences in work related values*. Thousand Oaks, CA: Sage.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage.

Hofstede, G., & Bond, M. H. (1988). The Confucius connection: From cultural roots to economic growth. *Organizational Dynamics, 16*(4), 4-21.

Hsu, G., Roberts, P. W., & Swaminathan, A. (2012). Evaluative schemas and the mediating role of critics. *Organization Science, 23*(1), 83-97.

Hu, N., Pavlou, P. A., & Zhang, J. (2009). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM, 52*(10), 144-147.

Jackson, L. A., & Wang, J.-L. (2013). Cultural differences in social networking site use: A comparative study of China and the United States. *Computers in Human Behavior, 29*, 910-921.

Jiang, Z., Heng, C.-S., & Choi, B. C. F. (2013). Privacy concerns and privacy-protective behavior in synchronous online social interactions. *Information Systems Research, 24*(3), 579-595.

Josang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems, 43*(2), 618-644.

Kappos, A., & Rivard, S. (2008). A three-perspective model of culture, information systems, and their development and use. *MIS Quarterly, 32*(3), 601-634.

Kauffman, R. J., & Wood, C. A. (2005). The effects of shilling on final bid prices in online auctions. *Electronic Commerce Research and Applications, 4*(1), 21-34.

Kayany, J. M. (1998). Contexts of unihibited online behavior: Flaming in social newsgroups on Usenet. *Journal of the American Society for Information Science, 49*(12), 1135-1141.

Khansa, L., Ma, X., Liginlal, D., & Kim, S. S. (2015). Understanding members' active participation in online question-and-answer communities: A theory and empirical analysis. *Journal of Management Information Systems, 32*(2), 162-203.

Kim, J. (2011). Two routes leading to conformity intention in computer-mediated groups: Matching versus mismatching virtual representations. *Journal of Computer -Mediated Communication, 16*(2), 271-287.

Kirkman, B. L., Lowe, K. B., & Gibson, C. B. (2006). A quarter century of “culture's consequences”: A review of empirical research. *Journal of International Business Studies, 37*(3), 285-320.

Kogut, B., & Singh, H. (1988). The effect of national culture on the choice of entry mode. *Journal of International Business Studies, 19*(3), 411-432.

Koh, B. J., Kim, Y.-G., Butler, B., & Bock, G.-W. (2007). Encouraging participation in virtual communities. *Communications of the ACM, 50*(2), 69-73.

Lampe, C., & Resnick, P. (2004). *Slash(dot) and burn: Distributed moderation in a large online conversation space*. Paper presented at the CHI '04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria.

Lea, M., O'Shea, T., Fung, P., & Spears, R. (1992). “Falming” in Computer-mediated communication: Observations, explanations, implications. In M. Lea (Ed.), *Contexts of Computer-mediated Communication* (pp. 89-112). New York: Harvester-Wheatsheaf.

Lee, A. S., & Baskerville, R. L. (2003). Generalizing generalizability in information systems research. *Information Systems Research, 14*(3), 221-243.

Leidner, D. E., & Kayworth, T. (2006). A review of culture in information systems research: Toward a theory of information technology culture conflict. *MIS Quarterly, 30*(2), 357-399.

Li, W. (2010). Virtual knowledge sharing in a cross-cultural context. *Journal of Knowledge Management, 14*(1), 38-50.

Li, X., & Hitt, L. M. (2008). Self-selection and information role of online product reviews. *Information Systems Research, 19*(4), 456-474.

Madupu, V., & Cooley, D. O. (2010). Cross-cultural differences in online brand communities: An exploratory study of Indian and American online brand communities. *Journal of International Consumer Marketing, 22*(4), 363-375.

Marshall, B. A., Cardon, P. W., Norris, D. T., Goreva, N., & D'Souza, R. (2008). Social networking websites in India and the United States: A cross-national comparison of online privacy and communication. *Issues in Information Systems, IX*(2), 87-94.

Massanari, A. (2013). *Playful participatory culture: Learning from Reddit*. Paper presented at the IR14: Annual Conference for the Association of Internet Researchers Denver, CO.

Matsumoto, D., & Hwang, H. S. (2011). Cooperation and competition in intercultural interactions. *International Journal of Intercultural Relations, 35*(5), 677-685.

McCoy, S., Galletta, D. F., & King, W. R. (2005). Integrating national culture into IS research: The need for current individual level measures. *Communications of the Association for Information Systems,* 15, 211-224.

McLaughlin, C., & Vitak, J. (2011). Norm evolution and violation on Facebook. *New Media & Society, 14*(2), 299-315.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56-63.

Mitchell, R. K., Smith, B., Seawright, K. W., & Morse, E. A. (2000). Cross-cultural cognitions and the venture creation decision. *Academy of Management Journal, 43*(5), 974-993.

Moe, W. W., & Trusov, M. (2011). The value of social dynamics in online product ratings forums. *Journal of Marketing Research, 48*(3), 444-456.

Mork, B. E., Aanestad, M., Hanseth, O., & Grisot, M. (2008). Conflicting epistemic cultures and obstacles for learning across communities of practice. *Knowledge and Process Management, 15*(1), 12-23.

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

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*(2), 103-122.

Murillo, E. (2011). Communities of practice in the business and organization studies literature. *Information Research, 16*(1), 1-46.

Murthy, D. (2013). *Twitter: Social communication in the digital age*. Cambridge: Polity Press.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673-690.

O'Reilly, C. A., III., Chatman, J., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal, 34*(3), 487-516.

Palich, L. E., Horn, P. W., & Griffeth, R. W. (1995). Managing in the international context: Testing cultural generality of sources of commitment to multinational enterprises. *Journal of Management, 21*(4), 671-690.

Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Information Systems Research, 17*(4), 392-414.

Pfeil, U., Zaphiris, P., & Ang, C. S. (2006). Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication, 12*(1), 88-113.

Pflug, J. (2011). Contextuality and computer-mediated communication: A cross cultural comparison. *Computers in Human Behavior, 27*(1), 131-137.

Posey, C., Lowry, P. B., Roberts, T. L., & Ellis, T. S. (2010). Proposing the online community self-disclosure model: The case of working professionals in France and the U.K. who use online communities. *European Journal of Information Systems, 19*(2), 181-195.

Postmes, T., Spears, R., & Lea, M. (2000). The formation of group norms in computer-mediated-communication. *Human Communication Research, 26*(3), 341-371.

Poston, R. S., & Speier, C. (2005). Effective use of knowledge management systems: A process model of content ratings and credibility indicators. *MIS Quarterly, 29*(2), 221-244.

Preece, J., & Shneiderman, B. (2009). The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction, 1*(1), 13-32.

Qiu, L., Lin, H., & Leung, A. K.-y. (2013). Cultural differences and switching of in-group sharing behavior between an American (Facebook) and a Chinese (Renren) social networking site. *Journal of Cross-Cultural Psychology, 44*(1), 106-121.

Rao, V. S., & Ramachandran, S. (2011). Occupational cultures of information systems personnel and managerial personnel: Potential conflicts. *Communications of the Association for Information Systems, 29*(Article 31), 581-604.

Raymond, E. S. (1999). The cathedral and the bazaar. *Knowledge, Technology, & Policy, 12*(3), 23-49.

Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly, 48*(2), 240-267.

Reagans, R., & Zuckerman, E. W. (2001). Networks, diversity, and productivity: the social capital of corporate R&D teams. *Organization Science, 12*(4), 502-517.

Resnick, P., Kuwabara, K., Zeckhauser, R., & Friedman, E. (2000). Reputation systems. *Communications of the ACM, 43*(12), 45-48.

Rui, J., & Stefanone, M. A. (2013). Strategic self-presentation online: A cross-cultural study. *Computers in Human Behavior, 29*(1), 110-118.

Schein, E. H. (2010). *Organizational culture and leadership*. San Francisco: Jossey-Bass.

Schlosser, A. E. (2005). Posting versus lurking: Communicating in a multiple audience context. *Journal of Consumer Research, 32*(2), 260-265.

Schultz, M., Mouritsen, J., & Gabrielsen, G. (2001). Sticky reputation: Analyzing a ranking system. *Corporate Reputation Review, 4*(1), 24-43.

Shackleton, V. J., & Ali, A. H. (1990). Work-related values of managers: A test of the Hofstede model. *Journal of Cross-Cultural Psychology, 21*(1), 109-118.

Shin, D.-H. (2010). Analysis of online social networks: A cross-national study. *Online Information Review, 34*(3), 473-495.

Siau, K., Erickson, J., & Nah, F. F.-H. (2010). Effects of national culture on types of knowledge sharing in virtual communities. *IEEE Transactions on Professional Communication, 53*(3), 278-292.

Sivakumar, K., & Nakata, C. (2001). The stampede toward Hofstede's framework: Avoiding the sample design pit in cross-cultural research. *Journal of International Business Studies, 32*(3), 555-574.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: SAGE Publications Ltd.

Son, J.-Y., Tu, L., & Benbasat, I. (2006). A descriptive content analysis of trust-building measures in B2B electronic marketplaces. *Communications of the Association for Information Systems, 18*.

Spinellis, D., Drossopoulou, S., & Eisenbach, S. (1994). Language and architecture paradigms as object classes: A unified approach towards multiparadigm programming. In J. Gutknecht (Ed.), *International Conference on Programming Languages and System Architectures* (pp. 191-207). Switzerland: Springer Verlag.

Srite, M., & Karahanna, E. (2006). The role of espoused national cultural values in technology acceptance. *MIS Quarterly, 30*(3), 679-704.

Stewart, D. (2005). Social status in an open-source community. *American Sociological Review, 70*(5), 823-842.

Stewart, K. J., & Gosain, S. (2006). The impact of ideology on effectiveness in open source development teams. *MIS Quarterly, 30*(2), 291-314.

Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage Publications.

Sun, M. (2012). How does the variance of product ratings matter? *Management Science, 58*(4), 696-707.

Surowiecki, J. (2004). *The wisdom of crowds*. New York: Doubleday.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*(2), 96-102.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology, 1*(2), 149-178.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp. 2-24). Chicago: Nelson-Hall.

Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. New York, NY: W. W. Norton & Company.

Triandis, H. C. (1972). *The analysis of subjective culture*. New York: Wiley-Interscience.

Triandis, H. C. (1994). *Culture and social behavior*. New York: McGraw-Hill, Inc.

Triandis, H. C. (2000). Culture and conflict. *International Journal of Psychology, 35*(2), 145-152.

Trice, H. M. (1993). *Occupational subcultures in the workplace*. Ithaca, NY: ILR Press.

Trice, H. M., & Beyer, J. M. (1993). *The cultures of work organizations*. Englewood Cliffs, NJ: Prentice Hall.

Turner, J. C. (1975). Social comparison and social identity: Some prospects for intergroup behaviour. *European Journal of Social Psychology, 5*(1), 1-34.

Van Roy, P., & Haridi, S. (2004). *Concepts, techniques, and models of computer programming*. Massachusetts: Massachusetts Institute of Technology.

Waguespack, D. M., & Sorenson, O. (2011). The ratings game: Asymmetry in classification. *Organization Science, 22*(3), 541-553.

Washington, M., & Zajac, E. J. (2005). Status evolution and competition: Theory and evidence. *Academy of Management Journal, 48*(2), 282-296.

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

Welser, H. T., Gleave, E., & Fisher, D. (2007). Visualizing the signatures of social roles in online discussion groups. *Journal of Social Structure, 8*(2), 1-32.

Wetherell, M. (1982). Cross-cultural studies of minimal groups: Implications for the social identity theory of intergroup relations. In H. C. Triandis (Ed.), *Social identity and intergroup relations* (pp. 207-240). Cambridge, UK: Cambridge University Press.

Wilder, D. A. (1986). Social categorization: Implications for creation and reduction of intergroup bias. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (vol. 19). Orlando, FL: Academic Press.

Yamagishi, T., Jin, N., & Kiyonari, T. (1999). Bounded generalized reciprocity: Ingroup boasting and ingroup favoritism. *Advances in Group Processes, 16*(1), 161-197.

Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & Merget, D. (2007). The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior, 10*(1), 115-121.

Yuki, M. (2003). Intergroup comparison versus intragroup relationships: A cross-cultural examination of social indentity theory in North American and East Asian cultural contexts. *Social Psychology Quarterly, 66*(2), 166-183.

Zhang, Y. L. (2013). Power distance in online learning: Experience of Chinese learners in US higher education. *International Review of Research in Open and Distance Learning, 14*(4), 238-254.

Zhao, W., & Zhou, X. (2011). Status inconsistency and product valuation in the California wine market. *Organization Science, 22*(6), 1435-1448.

About the Author

Tom Mattson has a PhD in International Management and is an Assistant Professor in the Robins School of Business at the University of Richmond. His research focuses on a variety of managerial and technological issues facing modern organizations.