Reputation, User Feedback, and Perceived Information Quality in Social Internet Media: An Empirical Study

Completed Research Paper

Dagmara Glogowska
Cork University Business School
University College Cork
College Road
Cork, Ireland
dagaglogowska@gmail.com

Csaba Csáki
Corvinus University of Budapest
Fővám tér 13-15, Sóház 225, Budapest
Hungary, H-1083
Csaki.Csaba@uni-corvinus.hu

Joseph Feller
Cork University Business School
University College Cork
College Road
Cork, Ireland
jfeller@ucc.ie

Rob Gleasure
Cork University Business School
University College Cork
College Road
Cork, Ireland
R.Gleasure@ucc.ie

Abstract

As the internet evolved into a more collaborative and interactive environment in which users both produce and consume content, related changes have occurred in terms of the factors affecting perceived information quality (PIQ). Two factors have emerged as being highly important in social internet media (SIM): community feedback and individual user reputation. This paper reports on a quantitative empirical study exploring the impact of reputation and user feedback on perceived information quality in SIM contexts. The paper presents a theoretical model derived from the literature, and the results of regression analysis of data gathered from a web-based knowledge sharing community. The findings support several prior studies but challenge others, and provide stimulus for fruitful avenues of further research. The study makes several contributions, including a refined conceptualization of PIQ and feedback and an improved understanding of the relationships between feedback, reputation, and PIQ.

Keywords: Perceived Information Quality, Social Media, User Feedback, Reputation, Community, Knowledge Sharing

Introduction

While the modus operandi of early web media (which we are labelling “Traditional Internet Media” or TIM in this paper) did not differ from the communication model of traditional publication contexts in the sense that there were discrete roles for content producers and consumers, over time the web has evolved into a highly interactive and collaborative community platform in which users both produce and consume content (which we are labelling “Social Internet Media” or SIM).
This evolution into a two-way communication channel has led to an increasing proliferation of user generated content and sharing of information. However, this growth did not come without issues and questions. One of the most important consequences of the appearance of social internet media relates to the quality of information shared this way, which raises the question how that quality is assessed in such environments. Literature of social internet media not only highlights the difficulties with information quality (Agichtein et al. 2008), but a few causes have also been identified. For example, issues with information quality may arise due to the lack of standards of information posted online, the vastly increased number of contributors, rapid growth of available content, and wide variety of information sources. Assessing information quality in Web contexts is thus more difficult as there are fewer overall quality control mechanisms such as editors, refereeing processes and reputational claims from established in publishing houses (Lee et al. 2002). Furthermore, people understand authority on the basis of personal experience with particular sources, reputations or other related factors (Wilson 1983), and use these factors consistently to reduce the amount of information they need to interact with. The issue here is that the Web has a dynamic nature, which means previous experiences and established authority structures are likely to be no longer directly relevant for this purpose.

Measuring the quality of information is not a straightforward task. Marchand (1990) discusses quality as “quality management” in library and information services. Levitin and Redman (1995) associate data quality with the accuracy of the information in products such as databases. Most commonly, the term “data quality” is described as data that is “fit-for-use” (Wang and Strong 1996) which implies that it is relative: data considered appropriate for one use may not possess sufficient attributes for another use (Tayi and Ballou 1998). From the point of view of the Internet, information quality (IQ) may also be defined in a user centric way when it is identified as the degree to which information is suitable for doing a specific task by a specific user, in a certain context (Emamjome et al. 2013). However, content perceived as excellent quality by some users might be considered as poor quality by others (Chai et al. 2009). One potential corresponding measure, therefore, is perceived information quality (PIQ) defined as the subjective judgements regarding the usefulness of the information in addressing the question posed. Early information systems research sets a foundation for PIQ definition as the user’s reaction to the characteristics of output information versus the user’s information requirements (Bailey and Pearson 1983). One judges information quality using the criteria of relevance, accessibility (validity), interpretability, and integrity (composed of accuracy and completeness) (Bovee 2004).

Another question relates to the factors describing IQ and PIQ. Over the years dozens of characteristics have been proposed and investigated including accuracy, authority, currency or novelty (Klobas 1995), but many other dimensions may be considered as well such as accessibility, actual value, completeness, credibility, flexibility, form, meaning over time, relevance, reliability, selectivity and validity (Olaisen 1990). Furthermore, assessment of PIQ may be impacted by several factors such as the source, the amount of data available to users, or the usefulness of the information.

Prior research suggests that dimensions of PIQ and its influencing factors have changed over time. While these attributes have been mostly preserved when moving from printed media to (traditional) Internet media (TIM), PIQ in the contemporary social internet (SIM) differs from earlier web environments in key ways. Two factors have emerged as characteristically different for collaborative media, that of community feedback and individual user reputation. Hence, the objective of this study is to explore the impact of reputation and user feedback on perceived information quality in social internet media contexts.

To achieve this, a quantitative study is performed on a web-based knowledge sharing community (the Q&A platform called The Workplace). A theoretical model derived from the literature and a systematic sampling method is used to gather necessary data. Empirical indicators for reputation, feedback, and perceived information quality are measured and eight hypotheses are tested using regression and other analysis techniques. While building the theoretical model it became clear that in a collaborative context it is necessary to differentiate between individual and group measures of PIQ and between the quantity and quality of user feedback. The findings support several prior studies but challenge others, and provide stimulus for fruitful avenues of further research. The study reveals that reputation does impact upon individual PIQ (the judgement of the answer seeker) but not on group PIQ (collective judgement at the community level). While feedback impacts both kinds of PIQ, the findings reveal it is the quality rather than quantity of feedback that is important. Finally, the study did not find evidence of a relationship.
between reputation and feedback, which is anticipated by prior research. The study thus makes several contributions, including a refined conceptualization of PIQ (individual vs group) and feedback (quality vs quantity) and an improved understanding of the relationships between feedback, reputation, and PIQ in social internet media contexts.

The remainder of the paper is structured as follows. The first section discusses the differences between traditional and social internet media. This is followed by a review of research of perceived information quality and its interpretation in both traditional and social internet contexts. The third section develops the proposed theoretical model. This is followed by a description of the study’s methodology. The fifth section presents the findings of the research. The paper closes with a discussion of the results, the contributions of the study, and suggestions for future research.

Comparison of Traditional vs. Social Internet Media

The web has evolved from a publication platform that initially resembled prior media in terms of discrete content producer/consumer roles, to a highly interactive and collaborative community platform in which users both produce and consume content. Web-based information in general potentially suffers from a relative absence of professional gatekeepers to monitor content, from content blending with advertising, from a lack of established reputation for information providers, and from its susceptibility to alteration, which may be difficult to detect (Alexander and Tate 1999; Flanagin and Metzger 2000; Flanagin and Metzger 2007; Johnson and Kaye 1998). Such factors put the responsibility of information assessment entirely on the media consumer, and put a question mark on the credibility of sources, the content posted by them, and the authority of the media that carry them (Flanagin and Metzger 2007; Johnson and Kaye 1998; Kim et al. 2000; Kiousis 2001; Schweiger 2000; Sundar and Nass 2000). Prior researchers argue that credibility is not an inherent property of the information or source, but is a property that is judged by the receiver of the information (Flanagin and Metzger 2007; Fogg et al. 2001; Freeman and Spyridakis 2004; Gunther 1992; Sundar 1998). Since the emergence of social internet media, the market space of traditional internet media has diminished. In the contemporary internet, it appears that most everything is shared, liked, commented, tweeted, and +1’d, and subscriptions and marketing are delivered more quickly and effectively to specific consumers (Kaplan and Haenlein 2010). SIM content can be quite different from TIM content in terms of style, quality, authorship, and explicit support for social connectivity and interaction. The explicit support for social interactions between users, such as posting comments, rating content, and responding to questions and comments makes the SIM unique and requires new techniques for analysing and retrieving relevant content (Bian et al. 2008). One of the key differences between TIM and SIM is that TIM content is largely static or at least more persistent. Traditional media content is researched, written, edited and then published; there are few or no changes post publication (Agichtein, et al. 2008). The typical range of information quality in TIM publishing is arguably narrower than in SIM. Thus the main challenge posed by content in SIM contexts is the fact that the distribution of quality has high variance, making the tasks of filtering and ranking in such systems more complex than in other domains (Agichtein et al. 2008). The two media models also differ in terms of frequency, that is the duration in which content is displayed is much longer in SIM than in TIM. A post on a website or online newspaper usually has a “shelf life” after which it is deleted or moved into archive. In SIM, posts are accessible for a much longer period, usually counted in years and people are able to join (possibly dormant) discussions at any stage. One of the problems shared by both TIM and SIM content is distinction between biased opinion and objective information. Additionally there is the issue of information revealing, as participants that know each other through previous online encounters or in real life might not express their full opinion due to personal relationships. Thus Nussbaum et al. (2002), argue that participants that know each other protect and nurture social relationships and show a general tendency to agree with one another to preserve harmonious social relations and to establish a common ground.

Perceived Information Quality

The proliferation of user generated content in social internet media contexts raise questions of how the quality of information is assessed in such environments. Prior research suggests that Perceived Information Quality (PIQ) in the contemporary social web differs from PIQ in earlier web environments in key ways, and has theorised the saliency of user feedback and individual user reputation in terms of
PIQ. For this study, Information quality (IQ) dimensions recognized in both TIM and SIM research have been reviewed in three criteria classes: subject, process and object. The subject class includes dimensions that can only be determined by individual user based on their personal views, experience and background, (such as reputation, user feedback or relevancy). Object class dimensions can be determined by a careful analysis of information itself, such as completeness, timeliness, objectivity or amount of data. The process class consists of dimensions that are determined by the process of querying, and as such are objective, but temporary (such as response time, accuracy or accessibility).

**PIQ in Traditional Internet Media**

Existing research in the subject of assessing IQ within TIM is relatively mature and has led to the discovery and validation of numerous IQ dimensions and metrics. Many scholars have focused on dimensions that predate the internet era, and emerge media scholarship in 1970's and 1980's which began to theorise media as an extension of the human body and senses (McLuhan 1964). In reviewing the work of Taylor (1986) and of 13 other papers also reviewed by Parker et al. (2006), (Zeist and Hendricks, 1996; Strong et al., 1997; Alexander and Tate, 1999; Katerattanakul and Siau, 1999; Shanks and Corbitt, 1999; Naumann and Rolker, 2000; Zhu and Gauch, 2000; Dedek, 2000; Leung, 2001; Kahn et al., 2002; Eppler and Muenzenmayer, 2002; Klein, 2002; Liu and Huang, 2005), several dominant attributes of IQ are revealed. In this study, the objective information characteristics (object class) include (in order of frequency of mention in the literature): timeliness, completeness, objectivity, consistency, quality of source, and freedom-from-error (accuracy). The process characteristics (again in order of mention), include accessibility, accuracy of transmission, representation quality, speed of access, security, and ease of manipulation. Finally in terms of the subject class (which is the most relevant to this study given its focus on perceived rather than objective information quality, the literature highlights the importance of individual assessments of relevancy to self, believability, appropriateness, understandability), perceived reputation of provider, and the value-added by the content.

**PIQ in Social Internet Media**

As in traditional internet contexts, the assessment of IQ in SIM contexts depend upon a wide variety of characteristics at the subject, object, and process levels. However, the many differences between TIM and SIM discusses above mean that the rubrics for assessing IQ in the two contexts, are different in subtle ways. Additionally the subject class of characteristics plays an increasingly important role in SIM content perceived as being of excellent quality by some users might be considered to be poor quality by others (Chai et al. 2009). PIQ in SIM contexts is further complicated by the existence of a very large and widely varied field of information sources, which in addition to the nature of the content itself, plays an important role in finding and judging the quality of content. Finally, the diversity of platforms in the SIM domain (discussion forums, Q&A portals, peer-to-peer applications, review portals, weblogs, wikis and others) makes the ability to develop consistent dimensions for IQ very challenging. In reviewing the work of Nussbaum et al (2002), Chen and Xu (2005), Knight and Burn (2005), Han and Liu (2006), Jeon et al (2006), Chai et al (2007), Hu et al (2007), Klamma et al (2007), Lauw et al (2007), Wiertz and de Ruyter (2007), and Agichtein et al (2008), several dominant attributes of IQ in SIM contexts are revealed, as are some key difference between TIM and SIM environments.

In this literature, the objective information characteristics (object class) include In this literature, the objective information characteristics (object class) again include completeness, timeliness, and consisteny, but notably objectivity is mentioned more frequently (perhaps reflecting the proliferation of individual content providers) and “amount of data” is the most frequently mentioned, arguably in response to the information overload that characterises SIM. In terms of process, the accuracy, security and accessibility of transmission remain important. Finally, in terms of the subject class of characteristics, user feedback (unmentioned in the TIM PIQ literature) and reputation dominate, with feedback as the most frequently mentions and reputation jumping from one of the least frequently mentioned characteristics to the second most frequently mentioned. The characteristics of content appropriateness, understandability, value-added, and believability are also visible in the literature.
PIQ in Traditional vs. Social Internet Media

We would argue that the differences between TIM and SIM quality-assessment characteristics reflect the quality (and assessment) problems created through two-way communications and many-to-many mappings. The analysis of the literature suggests that some IQ dimensions that were important in TIM either do not exist in SIM, changed their meaning or decreased their value. Likewise new salient dimensions appear in the SIM literature. These differences suggest that users adapt to the changing internet world, and that different factors influence their perception of information quality in these different contexts. SIM environments move faster, exhibit greater interactivity, and a proliferation of content providers. Thus some dimensions lose value and are replaced by others that are more suitable. Most critical to this study is the clearly important role played by two subject class dimensions in the perception of quality in social internet media contexts: (1) the individual reputation of the content provider and (2) user feedback within the social space. These two constructs thus form the focus for this study and the theoretical development in the next section.

Theoretical Development

Considering the changing emphasis on various IQ dimensions previously discussed, and in order to understand how information quality measures can be used to reduce information overload in SIM (so that people can benefit from good information shared online and make better decisions about information claims), it is necessary to understand the role of reputation and user feedback in the assessment of information quality. To achieve this objective, a theoretical model was developed built upon prior research, comprising three constructs and three propositions (see Figure 1). As these constructs and propositions represent the theoretical argument in a domain-agnostic manner, they are not directly testable. The methodology section describes their operationalization as testable hypotheses and empirical measures.

**Construct 1: Reputation (REP). The extent to which users can identify the standing of others, including themselves, in a social internet media setting.**

It has to be noted that reputation can have different meanings across SIM platforms. In most cases, reputation is a matter of trust, however, since information frameworks are not yet good at making scores of such highly qualitative criteria, SIM websites rely on tools that automatically aggregate user-generated information to determine trustworthiness. For instance, on Twitter, followers attest someone’s reputation, while on LinkedIn, reputation is built on endorsements from other users. However, in social internet media reputation refers not only to people, but also to content itself, which is often evaluated using content voting systems. On YouTube reputation might be based on “view counts” or “ratings” of the videos, while on Facebook this could be “likes” (Chai et al. 2007).
**Construct 2: User feedback (UF).** A specific type of user-generated content that is created as a response to other informational content.

UF can be either positive, negative or neutral. It can take the form of freeform text comments, emails, letters, ‘votes’, likes/dislikes, or more detailed survey-type information (c.f. Kaplan and Haenlein 2010).

**Construct 3: Perceived Information Quality (PIQ).** Subjective judgements as to the usefulness of the information in addressing the question posed.

PIQ can be understood as “the fitness for use” of a piece of content. Early information systems research sets a foundation for PIQ definition as the user’s reaction to the characteristics of output information versus the user’s information requirements (Bailey and Pearson 1983). One might further judge information quality using the criteria of relevance, accessibility (validity), interpretability, and integrity (composed of accuracy and completeness) (Bovee 2004).

**Proposition 1:** An increase in reputation (of the content provider) positively correlates with perceived information quality.

Many researchers have found that reputation is a factor of trust and is a strong predictor of attitude and behaviour (Jin et al. 2008; Klamma et al. 2007; Lee and Shavitt 2006; Purohit and Srivastava 2001). Reviews, ratings and other reputation systems already in place in SIM contexts allow users to obtain relevant information in order to perceive information quality. Through these measures, reputation can be assessed in order to reflect the collective opinions of users to rely on the diverse resources available online (Agichtein et al. 2008; Gregg 2009; Kim and Lennon 2013). Therefore, higher scores of reputation should show an increase in perceived information quality.

**Proposition 2:** An increase in user feedback positively correlates with perceived information quality.

Meaningful content promotes constructive discussion on a certain topic. Bringing positive contributions and different points of view to a conversation results in a richer exchange of opinions and leads to relevant additions (Agichtein et al. 2008; Wasko and Faraj 2005; Wiertz and de Ruyter 2007). Even when feedback occurs because an answer is controversial, rather than because it is ‘good’, such discussion may lead to arguments clarifying the disagreement eventually contributing to the answer (Ziegele, et al., 2014). Therefore, when users consider a piece of content worth reading, they are incentivized to participate in and generate new information from a healthy discussion (user feedback), which will correlate with higher perceived information quality.

**Proposition 3:** An increase in reputation (of the content provider) positively correlates with user feedback.

Prior research suggests that if a contributor adds high quality information on a certain topic, they will be considered as an expert in that field and their reputation will increase. Content added by this contributor is valued more highly by the community, as users will already have a certain level of trust based on previous encounters. Thus other users are more likely to engage in fruitful conversation (user feedback) with people they perceive more trustworthy and consistent (reputation) (Jeon et al. 2006; Nussbaum et al. 2002; Wiertz and de Ruyter 2007).

**Methodology**

**Site Selection and Data Gathering - Stack Exchange Network**

Data for this study were gathered from The Workplace (http://workplace.stackexchange.com/), a social Q&A site embedded within the larger Stack Exchange network. The Workplace allows users to pose questions about a range of workplace matters, such as standard workplace etiquette and common practices. Other users may provide answers, vote questions and answers up and down in terms of their quality, and comment on questions and answers. This site was selected for two reasons.

First, we observed that while discussion on Stack Exchange was typically specialized and dominated by a small number of individuals, The Workplace was more broadly collaborative and inclusive. This allowed analysis to focus on a broader range of users, rather than a small minority of passionate contributors.
Second, (as of March 2015) The Workplace reported that 100% of questions (N = 6,505) asked across 324 discussed topics had received at least one answer (N > 21,000). This suggests a sufficiently suggested a broad range activity across a many different topics which would improve both reliability and generalizability.

A manual systematic sampling approach was adopted sampling every 100th question and removing those that appeared to be spam or otherwise not usable. The final usable sample contained 203 answers connected with 55 questions. Data was collected between the 14th and 16th of March 2015, All 55 questions selected by systematic sampling were answered and only one answer out of 203 came from a non-registered user. Any duplicate questions were avoided, as they could show invalid results. None of them appeared to be spam, or misplaced.

**Empirical Measures**

The theoretical model was operationalized through five measures and eight testable hypotheses based on the theoretical constructs and propositions presented earlier (see Figure 2).

The first six (H1.a-b and H2.a-d) test the impact of reputation (REP) and user feedback (UF) on perceived information quality (PIQ). The final two (H3.a and H3.b) test the impact of REP on UF. PIQ and UF were each measured two different. The empirical indicators are presented in Table 1 and discussed further below.
Table 1: Measures used in previous studies

<table>
<thead>
<tr>
<th>Answer Level</th>
<th>Prior Use in Literature</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reputation (REP)</td>
<td>RA (reputation of answerer)</td>
<td>Scale</td>
</tr>
<tr>
<td></td>
<td>(Agichtein et al. 2008; Hu et al. 2007; Klamma et al. 2007; Lim et al. 2006)</td>
<td></td>
</tr>
<tr>
<td>User Feedback (UF)</td>
<td>AC (no. of comments to the answer)</td>
<td>Scale</td>
</tr>
<tr>
<td></td>
<td>(Agichtein et al. 2008; Jeon et al. 2006; Kim et al. 2006; Klamma et al. 2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SC (score of the comments)</td>
<td>Scale</td>
</tr>
<tr>
<td></td>
<td>(Agichtein et al. 2008; Jeon et al. 2006)</td>
<td></td>
</tr>
<tr>
<td>Perceived Information Quality (PIQ)</td>
<td>AS (score of the answer)</td>
<td>Scale</td>
</tr>
<tr>
<td></td>
<td>(Agichtein et al. 2008; Jeon et al. 2006; Klamma et al. 2007; Wiertz and de Ruyter 2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BA (is the answer voted as best answer by question owner?)</td>
<td>Binary</td>
</tr>
<tr>
<td></td>
<td>(Agichtein et al. 2008; Jeon et al. 2006; Wiertz and de Ruyter 2007)</td>
<td></td>
</tr>
</tbody>
</table>

A single indicator was used to measure the reputation (REP) construct, i.e. the existing reputation of a The Workplace/Stack Exchange user in relation to the question topic (as one user may have more than one reputation score; one for each community the person is active in). This is a numerical score assigned by the Stack Exchange software based on a user’s prior behaviours and other characteristics. Users can earn reputation points and “badges” through active participation and by making valuable contributions to the community. Table 2 lists sample actions that can increase or decrease a user’s reputation score (see http://meta.stackexchange.com/questions/7237/how-does-reputation-work).

Table 2: Stack Exchange reputation award system

<table>
<thead>
<tr>
<th>Increase</th>
<th>Decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>You increase your reputation score when:</td>
<td>You decrease your score when:</td>
</tr>
<tr>
<td>One of your questions is voted up/useful</td>
<td>One of your questions is voted down/not useful</td>
</tr>
<tr>
<td>One of your answers is voted up/useful</td>
<td>One of your answers is voted down/not useful</td>
</tr>
<tr>
<td>One of your answers is awarded “best answer” by an asker</td>
<td>One of your answers loses “best answer” status</td>
</tr>
<tr>
<td>You suggest an edit to an answer and it is accepted</td>
<td>One of your posts receives six spam or offensive flags</td>
</tr>
<tr>
<td>You associate two or more Stack Exchange network site accounts, and at least one of those accounts already has 200 or more reputation</td>
<td>You unaccept an answer written by someone else to one of your own questions</td>
</tr>
<tr>
<td>One of your answers is awarded a bounty (user to user reputation transfer “prize”)</td>
<td>You award a bounty to another user</td>
</tr>
</tbody>
</table>
The user feedback (UF) construct is measured using two indicators available on The Workplace. The first one was the **quantity** of feedback received by an answer – measured by the number of comments attached to the answer (AC). The second was the **quality** of feedback received by an answer – measured by the sum of scores of those comments (SC), as rated by other users. While the correlation between these measures is not meaningless, it is below the threshold at which it undermines the tests performed and the stepwise regression is intended to provide transparency for those who are concerned about its impact (which does not appear problematic).

Two indicators were also used to measure **perceived information quality** (PIQ). The first was score of the answer (AS), which is calculated by votes and represents the **community** perception of an answer’s quality. The second was the binary variable of “best answer” or not (BA). This decision is made by the author of the question, thus representing the **individual** perception of an answer’s quality.

### Data Analysis and Findings

Prior to hypothesis testing, basic descriptive analysis was performed to investigate the nature of the data. Five figure summaries are presented for each of the indicators in Table 3, suggesting a negative skew for each of the ordinal variables. Similarly, a frequency analysis of the binary variable (BA) suggests approximately 19% (N=203) were picked as best answers. The distribution of the AS dependent variable suggested a negative skew and positive kurtosis, thus a logarithmic transformation on these values (AS_L) was used as the dependent variable in its place. Such transformation makes it difficult to draw conclusions based on the size of specific relationships, however it allows for greater reliability in detecting the presence or absence of positive or negative relationships (Bartlett 1947; Osborne 2005). To make sure no illegal operations occur when using this new transformed variable (min. value of AS is -4) a constant of 4 is added to the function.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Name</th>
<th>Variable</th>
<th>Min.</th>
<th>1Q</th>
<th>2Q</th>
<th>3Q</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>Score of the answer</td>
<td>Dependent</td>
<td>-4</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>143</td>
</tr>
<tr>
<td>RA</td>
<td>Reputation of the answerer</td>
<td>Independent</td>
<td>0</td>
<td>448</td>
<td>2800</td>
<td>16900</td>
<td>68100</td>
</tr>
<tr>
<td>AC</td>
<td>No. of comments to the answer</td>
<td>Independent</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>SC</td>
<td>Score of the comments</td>
<td>Independent</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 3: Five-figure summaries for variables
No evidence for multicollinearity was observed above the .707 threshold (Straub et al. 2004). However, AC and SC correlated at almost 60% (see Table 4), thus stepwise hierarchical regressions were performed for hypotheses H1.a to H2.d as a precaution.

<table>
<thead>
<tr>
<th>Table 4: Correlation Matrix of independent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
</tr>
<tr>
<td>RA (Reputation of the answerer)</td>
</tr>
<tr>
<td>AC (No. of comments to the answer)</td>
</tr>
<tr>
<td>SC (Score of the comments)</td>
</tr>
<tr>
<td>NS = Not significant, ** = p&lt; .01 level (2-tailed)</td>
</tr>
</tbody>
</table>

Table 5 presents the results of hierarchical regression using log transformed score of the answer (AS_L) as a dependent variable. Results suggest that score of the comments (SC) has a strong positive correlation (Beta = .552***) and reputation has no significant correlation (Beta = .08 NS). The number of comments (AC) does not quite reach 95% confidence interval of statistical significance, however the results suggest a negative correlation may exist outside the threshold of sensitivity at the current sample size (Beta = -.14 NS, p = .067).

<table>
<thead>
<tr>
<th>Table 5: Results of Hierarchical Linear Regression on Answer Score (transformed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>RA</td>
</tr>
<tr>
<td>AC</td>
</tr>
<tr>
<td>SC</td>
</tr>
<tr>
<td>Nagelkerle R²</td>
</tr>
</tbody>
</table>

The data suggest that the reputation of the answerer (RA) is not a significant predictor of community-level PIQ. In model 3 the results reveals that quality of the comments (SC) is more significant than the quantity of the feedback (AC). More interestingly, when all three predictors are factored in there is some support for a negative relationship with AC. Thus, while community-generated PIQ may appear to increase with more comments, once the quality of these comments is factored in, more comments ultimately decrease PIQ. This suggests from a community-perspective, there is a limit on the useful amount of comments an answer can get, after which it becomes noise.

Table 6 presents the results of hierarchical binary logistic regression where the dependent variable is whether or not an answer was selected as the best by the information-seeker (BA). Results suggest that
both score of the comments (SC) and reputation (RA) have significant positive correlations with SC (Beta = .195*) and RA (Beta = .167*), while the number of comments to the answer (AC) has no significant correlation (Beta = .33NS). In contrast to AS, data show that the reputation of the answerer does indeed have a significant impact in predicting whether or not an answer is selected as the best.

Table 6: Results of Hierarchical Linear Regression on likelihood of Best Answers

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hypothesis</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>RA</td>
<td>H1.b</td>
<td>.176*</td>
<td>.184**</td>
<td>.167*</td>
</tr>
<tr>
<td>AC</td>
<td>H2.b</td>
<td>.147*</td>
<td></td>
<td>.033 NS</td>
</tr>
<tr>
<td>SC</td>
<td>H2.d</td>
<td></td>
<td></td>
<td>.195*</td>
</tr>
<tr>
<td>Nagelkerle R²</td>
<td></td>
<td>.031</td>
<td>.053</td>
<td>.078</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01, ***p<.001, NS = Not Significant

Table 6: Results of Hierarchical Linear Regression on likelihood of Best Answers

H3.a and H3.b concentrate on investigating whether the reputation of the answerer directly affects quantity of the user feedback and quality of the user feedback, respectively. In this case, RA remains as an independent variable and the measures of user feedback (AC and SC) act as dependent variables.

Both AC and SC demonstrate a negative skew and a high proportion of zero values. Thus, two sets of tests were performed to test the relationship between these variables. First, binary logistic regressions were performed to test whether RA predicted if an answer would receive comments or votes on the quality of those comments. The data suggest that there is no significant correlation between reputation (RA) and whether or not an answer received comments, Beta = <.001, p= .481, nor between reputation (RA) and whether or not the community voted on the quality of comments, Beta = <.001, p= .485.

Second, linear regressions were performed using only those answers that had received >0 comments and ratings on those comments. The results suggest that there is no significant correlation between reputation (RA) and quantity of user feedback (AC), Beta = -.050, p= .636, nor between reputation (RA) and quality of user feedback (SC), Beta = .149, p= .256.

Discussion

Summary of Findings

The data supported three out of eight hypotheses. H1. a and H1.b explored the relationships between the reputation of the answerer (RA) and two types of perceived information quality – community-perceived (H1.a) and individual-perceived (H1.b). The score of the answer (AS) measured community PIQ and individual PIQ was measured by whether an answer was selected as best answer or not (BA). H1.b was supported, suggesting the reputation of the answerer is an important predictor of individual PIQ. This is in line with previous research (e.g. Ulicny et al. 2007). However, H1.a was not supported, suggesting that reputation may not matter in predicting community PIQ. This finding contradicts a variety of existing studies (Han and Liu 2006, Jeon et al. 2006, Hu et al. 2007, Klamma et al. 2007, Agichtein et al. 2008). One possible interpretation for this difference is that the personal connection between information-seekers and an individual answerer is greater than that between different answerers. Thus, the weight of any particular answerer’s accumulated reputation is more likely to be salient at a one-to-one level, either from a trust perspective or simply because of a perceived pressure to agree with powerful individuals. An alternative explanation is that the information-seeker may not be as frequent a user as those community members casually browsing and voting on different questions relating to some topic. If this was the case, the information needs of information-seekers and other community members may be fundamentally
different, e.g. the level of nuance in some answer may be less valued the former, while simplicity may be less valued by the latter. We note that prior studies used different empirical measures of reputation, however the constructs in prior studies are theoretically consistent with our own. Thus, either the construct is affected in contrasting ways, or the construct is defined in a manner that makes possible inconsistent empirical instantiations. Arguably, each of these possibilities has similar implications – as a field, we don’t understand how reputation impacts community PIQ sufficiently well.

H2.c and H2.d explored the relationships between the quality of feedback (score of the comments of the answer (SC)) and both types of PIQ, the community-perceived AS (H2.c) and the individual-perceived BA (H2.d). Both of these relationships were supported by the data. This provides further support for a range of existing studies highlighting the importance of quality feedback (Nussbaum et al. 2002, Cheng and Vassileva 2006, Han and Liu 2006, Kim et al. 2006, Klamma et al. 2007, Lauw et al. 2007, Wiertz and de Ruyter 2007, Agichtein et al. 2008).

H2.a and H2.b explored the relationships between the quantity of feedback (the number of comments attached to the answer (AC)) and both types of PIQ, the community-perceived AS (H2.a) and the individual-perceived BA (H2.b). The data suggested no significant impact on seeker-perceived information quality and suggested a negative relationship between the number of comments and community-perceived information quality. Again, this contrasts with prior research (Price and Shanks 2005, Cheng and Vassileva 2006, Han and Liu 2006, Jeon et al. 2006, Kim et al. 2006, Ulicny et al. 2007). One explanation for this is that there is a limit on the number of useful comments for an answer, after which point, further commentary acts as noise diluting the actual information content of an answer. Thus, once the quality of commentary is considered, sheer quantity becomes detrimental. Further, it may be the case that certain types of comments provoke a significant number of responses, not because they add to the core discussion but because they are controversial, dogmatic, or even just humorous.

H3.a and H3.b explored the relationship between the reputation of the answerer (RA) and both dimensions of user feedback, quantity (H 3.a) and quality (H3.b). Neither hypotheses were supported by the data, again contradicting existing research (Cheng and Vassileva 2006, Han and Liu 2006, Jeon et al. 2006, Lim et al. 2006, Hu et al. 2007, Lauw et al. 2007, Ulicny et al. 2007, Agichtein et al. 2008). This further suggests a dilution of individual reputation as interactions become less one-to-one in nature and more decentralized and collaborative.

**Contributions and Model Refinement**

Previous research suggests that reputation and user feedback and interaction is an important predictor of perceived information quality. This paper adds nuance to that body of research by suggesting the impact of reputation may be limited to individual-level perceptions, rather than those of the broader community. Further, this study did not find evidence of any relationship linking reputation and user feedback (quality or quantity). This study also shows that while the quality of user feedback is an important predictor of perceived information quality, the quantity of feedback has either no effect (from an individual perspective) or possibly a negative effect (from a community perspective).

These findings provide three key contributions to the scholarship on PIQ in social internet contexts.

The first contribution of this study is the additional empirical evidence linking both reputation and user feedback to perceptions of information quality in SIM. This adds to the body of work performed in this area (e.g. Han and Liu 2006, Kim et al. 2006, Ulicny et al. 2007) and lends some support to prior theoretical work. Additionally, specific findings linking the quality of user feedback to perceived information quality replicate specific findings from previous studies (Nussbaum et al. 2002, Price and Shanks 2005, Cheng and Vassileva 2006, Jeon et al. 2006, Lauw et al. 2007).

The second contribution is in challenging other findings from prior research. Data from this study did not support the relationship between reputation and user feedback, in contrast to the findings of previous studies (Han and Liu 2006, Hu et al. 2007). Neither did the data support the relationship between reputation and community-perceived information quality, challenging the theoretical work of (Suryanarayana and Taylor 2004, Adler and De Alfaro 2006). Moreover, the data suggested there may be a negative relationship between the quantity of feedback and community-perceived information quality, which has not been identified in previous research (Jeon et al. 2006, Agichtein et al. 2008).
The third contribution is the reconciliation of these similarities and differences in the form of the refined model presented in Figure 3. This refined model separates two of the initial proposition-level constructs, for whom sub-components were interacting differently with other theoretical entities, into separate hypothesis-level constructs that can be predicted more reliably. Thus, building on the findings and analysis presented in the previous section (i) perceived information quality (PIQ) is separated into community-perceived information quality and seeker-perceived information quality (ii) user feedback is separated in quantity of user feedback and quality of user feedback.

**Figure 3: Revised Theoretical Model**

**Suggestions for Future Research**

This study suggests multiple opportunities for future research.

First, the generalizability of the study reported is limited by the fact that the study concentrated on a single discussion board (The Workplace), with a particular subject, on a single social internet media platform. This suggests a need to replicate the study in other discussion boards and/or on other social internet media platforms.

Second, the explanatory power of the study is limited by the use of a single quantitative method. While the study has identified a number of interesting relationships, the data is limited in explaining the phenomenon, and thus a deep qualitative study is needed to investigate how people feel about perceiving information quality and what features of an answer (or of the community or platform) act to engage them.

Third, there is the possibility that alternative empirical measures for the theoretical constructs explored will reveal new insights. For example, the measure used for reputation (a system generated value representing positive community participation in the past) was deemed appropriate for this research, however alternative measures that more holistically capture a content provider’s history and standing in the community may yield new results. Additionally we would call for research exploring the possibility of individuals gaming the reputation and Q&A systems.

Fourth, our work suggests that feedback should not be treated as a single concept. Rather, this study suggests the need for future research that directly compares both quantity and quality of user feedback. In particular research is needed to further explore the suggested negative relationship between quantity of feedback and PIQ.

Fifth, our work suggests that perceived information quality also needs to be divided into two concepts: community-perceived and individual-perceived quality. Thus, this study suggests the need for future research that explores these two different areas of PIQ, and critically the relationship between the collective intelligence of community-perceived and individual judgements of quality.
Finally, while reputation and user feedback explain some of the variance in PIQ, it certainly does not explain all. Presumably, unexplained variance is accounted for in terms of timeliness, completeness, understandability, or other traditional (or perhaps previously unidentified) factors in the information quality literature. Thus, we'd call for research into the relationships between the constructs included in this study, and those presented in other internet information quality studies but not included here.

Acknowledgements

This research was funded in part by the Lewis Charitable Foundation (USA) via the TOTO research project, Cork University Business School, University College Cork, Ireland.

References


Han, J., and Liu, Y. 2006. "Dubious feedback: fair or not?" in Proceedings of the 1st International Conference on Scalable Information Systems, ACM.


