Online Information Review
The international journal of digital information research and use

Number 6

733 Editorial advisory board

734 Protecting privacy on the Web: a study of HTTPS and Google Analytics implementation in academic library websites
Patrick O'Brien, Scott W.H. Young, Kenning Artisch and Karl Benedict

752 Image search and retrieval problems in web search engines: a case study of Persian language writing style challenges
Yaghoub Norouzi and Hoda Homavandi

768 Google as a political subject: the right to be forgotten debate 2014–2016
Linnea Lindskold

784 Social media users’ share intention and subjective well-being: an empirical study based on WeChat
Liang Ma, Xin Zhang and Xiao Yan Ding

802 Factors driving continued use of online health promotion competitions: evidence from an online weight-loss community
Xiaolong Song, Yi-Hung Liu, Jiahua Jin and Jianguo Zhao

821 Topical evolution patterns and temporal trends of microblogs on public health emergencies: an exploratory study of Ebola on Twitter and Weibo
Lu An, Chuanming Yu, Xia Lin, Tingyao Du, Liqin Zhou and Gang Li

847 Context congruity effects of online product recommendations: an eye-tracking study
Jing Luan, Zhong Yao, Yongchao Shen and Jie Xiao

864 Mood metadata on Chinese music websites: an exploratory study with user feedback
Xiao Hu, Christy W.L. Cheong, Siwei Zhang and J. Stephen Downie

880 Sexual health information-seeking behavior on a social media site: predictors of best answer selection
Yong Jeong Yi

917 Predicting the success of Twitter in healthcare: a synthesis of perceived quality, usefulness and flow experience by healthcare professionals
Mohamed Gamal Aboelmaged

923 What drives internet users’ willingness to provide personal information?
Ching-Hsuan Yeh, Yi-Shun Wang, Shin-Jeng Lin, Timmy H. Tseng, Chin-Hui Lin, Ying-Wei Shih and Yi-Hsuan La

940 Library marketing via social media: the relationships between Facebook content and user engagement in public libraries
Soohyung Joo, Namjoo Choi and Tae Hyun BaeK

956 Do corporate websites’ changes reflect firms’ survival?
Desamparados Blazquez, Josep Domenech and Ana Debon

971 A comparative study of disaster management information systems
Muhammad Massoud Rall, Taq Aziz and Sarosh Hashmat Lodi

999 Understanding social networking sites continuance: the perspectives of gratifications, interactivity and network externalities
Chun-Ming Chang
Protecting privacy on the web
A study of HTTPS and Google Analytics implementation in academic library websites

Patrick O’Brien, Scott W.H. Young and Kenning Arlitsch
Montana State University, Bozeman, Montana, USA, and
Karl Benedict
University of New Mexico, Albuquerque, New Mexico, USA

Abstract

Purpose – The purpose of this paper is to examine the extent to which HTTPS encryption and Google Analytics services have been implemented on academic library websites, and discuss the privacy implications of free services that introduce web tracking of users.

Design/methodology/approach – The home pages of 279 academic libraries were analyzed for the presence of HTTPS, Google Analytics services and privacy-protection features.

Findings – Results indicate that HTTPS implementation on library websites is not widespread, and many libraries continue to offer non-secured connections without an automatically enforced redirect to a secure connection. Furthermore, a large majority of library websites included in the study have implemented Google Analytics and/or Google Tag Manager, yet only very few connect securely to Google via HTTPS or have implemented Google Analytics IP anonymization.

Practical implications – Librarians are encouraged to increase awareness of this issue and take concerted and coherent action across five interrelated areas: implementing secure web protocols (HTTPS), user education, privacy policies, informed consent and risk/benefit analyses.

Originality/value – Third-party tracking of users is prevalent across the web, and yet few studies demonstrate its extent and consequences for academic library websites.

Keywords Web analytics, HTTPS, Third-party tracking, Web privacy

Paper type Research paper

Introduction

Third-party tracking can occur when web analytics services, such as Google Analytics, are utilized to measure visitation to websites. These services provide information about website use and user behavior, which can help libraries improve their online services. However, the analytics services operate sophisticated mechanisms through extensive networks to track users and their behavior across sites, acquiring user demographics and behavioral patterns. The detailed tracking enabled by Google Analytics is often performed without the fully informed consent of individual users of the website. The extent to which Google Analytics services have been implemented within the domain of library websites has been unknown prior to this study. Unknown, also, has been the extent to which available privacy-protecting features have been implemented on those websites.

The library profession has long supported the principles of privacy, but tracking used by analytics service providers has rendered those principles nearly untenable. For example, without proactive efforts to mitigate their impact, browser cookies set by Google Analytics act as beacons for collecting and sharing user data through a vast network of commercial trackers. By understanding the extent and significance of web tracking and the available...
privacy-protection mechanisms, libraries can begin to minimize their participation in third-party tracking on the web.

The results presented in this paper demonstrate conclusively that 279 academic libraries from around the world must do much more to ensure user privacy if they hope to maintain trust with their users. The principle of this trust is outlined in the privacy statements of the American Library Association (ALA), Coalition for Networked Information (CNI), National Information Standards Organization (NISO) and the International Federation of Library Associations and Institutions (IFLA).

In presenting our research, we first explain web tracking, web analytics and web privacy. We then detail our methods and results, followed by a discussion of the privacy implications of third-party web tracking. We conclude by offering recommendations for professional action and avenues for future research.

Literature review

Web tracking

The practice of third-party tracking on websites is widespread (Narayanan and Reisman, 2017), and has only increased in prevalence, variety and complexity over time (Lerner et al., 2016; Englehardt and Narayanan, 2016). One of the most common trackers found on the Web is produced by the Google Analytics web service, which is used to measure the visitation to a website (Lerner et al., 2016; Schelter and Kunegis, 2016). In exchange for this easy-to-implement and free-to-use analytics service, websites execute Google Analytics JavaScript code and pass user visit data to Google through browser cookies set by Google Analytics (Krishnamurthy and Wills, 2009). Such data are considered to be “leaked” if the user is unaware of its collection and does not consent to the data being shared with additional third parties (Sar and Al-Saggaf, 2013). An analysis of 1m websites found that nearly nine in ten websites leak user data to third parties without the user’s knowledge (Libert, 2015).

The Google Analytics tracker is not designed to leak user data across sites on its own, but its tracking capabilities are enhanced when combined with Google AdSense, Google’s popular cross-site advertising service that utilizes its Doubleclick tracker. When Google AdSense and Google Analytics have both been implemented in a website, the unique identifiers from each service can be linked by Google’s Doubleclick tracker such that Google can create browsing profiles that track users across sites (Roesner et al., 2012). Data leakage from Google Analytics can also occur when websites activate the additional Google tracking service known as Tag Manager, which allows for cross-site tracking and targeted advertising (Bashir et al., 2016). Under these expanded tracking conditions, third-party trackers can match user behavior data with user profiles, thereby allowing users to be tracked and targeted across the web (Olejnik et al., 2012; Falahrastegar et al., 2016; Kalavri et al., 2016). While data about Google Tag Manager and Google AdSense were collected during course of this study, full analysis is beyond the scope of this paper.

Data leakage and user profiling via web tracking represents a privacy issue for users because of a lack of transparency and the lack of opportunity for users to consent to the sharing of their tracked behavior. The following example illustrates this case:

A user logs into Gmail and then visits a library website that has implemented Google Analytics or Google Tag Manager. This user then searches for tax relief resources through the library website. Because Google 1) identifies and authenticates users via their Google IDs and passwords and 2) identifies and authenticates the library website through Google Analytics or Tag Manager, Google can link users’ library website activity to individual users’ Google profiles. Depending on the library’s Google implementation, this user activity may also be shared with Google’s advertising network, which targets users with personalized ads, such as credit cards or personal loan services, even after the user has left the library web site.
This style of tracking is pervasive; Google was shown to be capable of tracking users on nearly 80 percent of the top 1m websites (Libert, 2015). Websites that implement Google Analytics and other Google tracking services are participating in the extensive network of third-party trackers that are capable of sharing user data across sites. It appears that in most cases, the user has neither knowingly or explicitly given informed consent for this type of data sharing, nor does the website owner fully understand the capabilities and consequences of web analytics and other third-party trackers. While this is a common practice when interacting with many sites on the web, academic libraries using Google trackers without proactively enabling user privacy features may have unwittingly violated the principles of user privacy expressed by the ALA, CNI, NISO and IFLA privacy guidelines.

Web privacy

The library science professional literature includes many contributions that detail the implementation, application and justification of Google Analytics for the purposes of web traffic analysis and service improvement (Hess, 2012; Barba et al., 2013; Cohen and Thorpe, 2015; Fagan, 2014; Yang and Perrin, 2014; Conrad, 2015; Farney, 2016). User privacy is seldom mentioned in these articles and manuals. Yet, the user data collected by Google Analytics, such as search terms, user-agent software, geographical location, and time of day, can potentially be leaked to other third-parties via the network of web trackers. User privacy can be further undermined when third parties match behavior data with user profiles, thereby allowing users to be tracked and targeted across the web (Olejnik et al., 2012). Certain Google Analytics implementation methods can help reduce its data collection capability and reduce library participation in cross-site user tracking. These mitigating techniques include IP anonymization[1], opt-out mechanisms[2] and secure HTTP connections. A secure HTTP connection, also referred to as HTTPS, can be activated with a secure digital certificate and proper configuration of the host server (Naylor et al., 2014; Askey and Arlitsch, 2015). The use of HTTPS ensures that communication over the public internet is encrypted; and when the certificate is provided by a trusted certificate authority, it provides a verification mechanism to assure users that the website they are visiting belongs to the domain name owner and server they have requested. Without HTTPS protection in place, user activities over wired or wireless networks can be observed and retained.

Best practices for search engine optimization indicate that websites should automatically redirect non-secure URL user requests (HTTP) to secure versions of the URL (HTTPS) by way of a permanent webserver redirect (Arlitsch and O'Brien, 2013)[3]. These practices signal to users that site administrators are concerned with user privacy, thereby engendering trust.

Privacy has long been a concern of libraries (Milion and Fisher, 1986; Garoogian, 1991; Johnston, 2000; Nichols Hess et al., 2015), and defending privacy was much more attainable in the pre-digital world. Given the extent of third-party tracking on the internet, however, it is exceedingly difficult to implement analytics trackers like Google Analytics without compromising the privacy for users that libraries have championed. Library professional organizations have acknowledged the complexity of contemporary information privacy, and have modified privacy statements accordingly.

The ALA has published several statements and toolkits to help librarians achieve privacy for users. ALA describes its privacy and surveillance guidelines as an attempt “to balance the need to protect reader privacy with the needs of libraries to collect user data and provide personalized services[4].” In the Library Bill of Rights, the ALA offers a definition of privacy: “In a library (physical or virtual), the right to privacy is the right to open inquiry without having the subject of one’s interest examined or scrutinized by others[5].” The ALA continues in the Library Bill of Rights: “Libraries should not share personally identifiable user information with third parties or with vendors that provide resources and library services unless the library has obtained the permission of the user or has entered into a legal
agreement with the vendor.” ALA has offered additional calls-to-action through its Privacy Toolkit, which states, “For libraries to flourish as centers for uninhibited access to information, librarians must stand behind their users’ right to privacy and freedom of inquiry[6].”

A recent Executive Roundtable Report of the Coalition for Networked Information (CNI) notes: “Libraries collecting data using Google Analytics are realizing they may be violating the ALA Library Bill of Rights[…] this is but one example of how easily convenient web-based service offerings can come with unexpected consequences[7].”

The NISO has released a document that outlines 12 privacy principles for third-party e-resource systems[8]. The IFLA Statement on Privacy in the Library Environment recommends: “Library and information services should reject electronic surveillance and any type of illegitimate monitoring or collection of users’ personal data or information behavior that would compromise their privacy and affect their rights to seek, receive and impart information[9].” IFLA further identified that “the rapid advancement of technology has resulted in increasing privacy implications.” From within this context of networked complexity and third-party tracking, the Library Freedom Project has drafted the First Library Digital Privacy Pledge, which aims to increase the implementation of HTTPS on library websites, and has gained 21 endorsements from membership organizations, public and academic libraries, and vendors, as of this writing[10]. See Table I for a summary of privacy statements from professional library organizations.

A survey of librarians’ attitudes toward privacy found that 97 percent of respondents agree or strongly agree that libraries should never share personal information and circulation or internet records without authorization or a court order (Zimmer, 2014). In the same survey, 76 percent of respondents feel that libraries are doing all they can to prevent unauthorized access to individual’s personal information and circulation records; however, a different survey investigating the configuration of public internet terminals showed that many libraries have not installed ad-blocking and privacy-protecting features on web browsers, nor do they offer instruction to users regarding web privacy.

<table>
<thead>
<tr>
<th>Organization</th>
<th>Statement title</th>
<th>Statement excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Library Association (ALA)</td>
<td>Library Bill of Rights – Interpretation of Privacy</td>
<td>In a library (physical or virtual), the right to privacy is the right to open inquiry without having the subject of one’s interest examined or scrutinized by others</td>
</tr>
<tr>
<td>Coalition for Networked Information (CNI)</td>
<td>Privacy in the Age of Analytics</td>
<td>Libraries collecting data using Google Analytics are realizing they may be violating the ALA Library Bill of Rights…this is but one example of how easily convenient web-based service offerings can come with unexpected consequences</td>
</tr>
<tr>
<td>National Information Standards Organization (NISO)</td>
<td>NISO Privacy Principles</td>
<td>Libraries, publishers and software providers have a shared obligation to foster a digital environment that respects library users’ privacy as they search, discover and use those resources and services</td>
</tr>
<tr>
<td>International Federation of Library Associations and Institutions (IFLA)</td>
<td>Privacy Statement</td>
<td>Library and information services should reject electronic surveillance and any type of illegitimate monitoring or collection of users’ personal data or information behavior that would compromise their privacy and affect their rights to seek, receive and impart information</td>
</tr>
<tr>
<td>Library Freedom Project</td>
<td>Digital Privacy Pledge</td>
<td>Library services and resources should be delivered, whenever practical, over channels that are immune to eavesdropping</td>
</tr>
</tbody>
</table>

Table I. Privacy statements – professional library organizations
As other authors neatly summarize, “many websites use [Google Analytics and other click-tracking mechanisms], and their utility within library systems is an ongoing debate as we balance the needs of reliable metrics with patron privacy” (Caro and Markman, 2016). The use of Google Analytics on library websites is ubiquitous, yet the tension between web analytics and web privacy demands further investigation to ensure that libraries are in fact doing all we can to prevent unauthorized and unwanted data sharing.

In response to the perceived lack of professional knowledge regarding the extent of third-party web tracking on library websites, we have conducted a privacy audit that empirically measures the extent of Google Analytics services and related privacy-protection features on library websites. While many librarians agree that libraries should not share user data, their use of Google Analytics services without implementing available privacy-protection features signal that libraries are not doing all they can to prevent user data leakage. Understanding the prevalence of tracking and privacy infrastructure is a fundamental first step for taking concerted professional action that will benefit the privacy of users. As Lerner et al. (2016, p. 997) assert:

Measurement studies of web tracking are critical to provide transparency for users, technologists, policy-makers, and even those sites that include trackers, to help them understand how user data is collected and used, to enable informed decisions about privacy, and to incentivize companies to consider privacy.

To motivate the larger community of library professionals, we must first grasp the nature and extent of web tracking that occurs on library websites.

**Research questions**

**RQ1.** Do libraries implement HTTPS with proper redirect practices?

Does the library protect privacy with a secure connection (via HTTPS) between the user’s browser and the library’s website? Does the library use a permanent redirect to enforce the use of secure connections? Does the library redirect secure page requests to a non-secure version of the page in violation of recommended practice?

**RQ2.** Do libraries that use Google Analytics implement the available privacy-protection measures?

Does the library use Google Analytics? If the library is using Google Analytics, does it protect user privacy via a secure connection between the library website and Google’s servers? If the library is using Google Analytics, does it obfuscate individual user tracking using Google’s IP Anonymization feature?

**Methodology**

Webometrics is a subset of interrelated library and information science empirical research methodologies, whose relationship can be visualized as overlapping concentric circles beginning at the outer circle with informetrics and then moving inward toward bibliometrics, scientoetrics, cybermetrics and webometrics (Björneborn and Ingwersen, 2004). As a family of methodologies, informetrics and its subsets comprise a relatively small percentage of the published library and information science research (Togia and Malliari, 2017).

Webometrics was originally proposed as a research methodology in the late 1990s when it became apparent that longstanding informetric and bibliometric methods could be applied to the content and structure of the World-Wide-Web (Almind and Ingwersen, 1997).
Webometrics initially focused on statistical analyses of word and phrase frequencies, citations, characteristics of authors and publications, and rankings and impact factors, but the definition of the new methodology evolved quickly. “Link structures and search engines” were added (Björneborn and Ingwersen, 2001), and the definition was then further expanded to include “quantitative aspects of the construction and use of information resources, structures and technologies on the Web” (Björneborn and Ingwersen, 2004). It is this expanded definition that guides the current research, which fundamentally investigates the prevalence of security and privacy structures in academic library websites.

Research methods
Within the webometrics methodology, our data-gathering method can be classified as covert observation research, a social sciences research technique used to observe participant behavior without revealing the identity or presence of the researcher (Punch, 2014; Taylor et al., 2016). Covert observation has been used in many fields to gather both qualitative and quantitative data. In the health care field, for instance, it was used to discreetly observe behavior by participants in online communities that support or disparage eating disorders (Brotsky and Giles, 2007). It has also been used in quantitative marketing research, a business discipline that involves “collection of data gauging respondents’ reactions to stimuli and perceptions of a product” (Kurian, 2013), but its emphasis in this discipline is on the agreement reached by the buyer and the seller. Covert observational techniques can raise ethical concerns when the subjects are people (Hallenberg, O’Neill, and Tong, 2015; Stanley and McLaren, 2007). However, the observed subjects in the current research are the information structures publicly hosted on machines; specifically, we observe the presence or lack of HTTPS and the presence or lack of the Google Tracking Code. As such, there are no ethical concerns associated with this research.

Research design
Our research examines the website home pages of 279 US and international academic libraries. The study population included libraries with one or more memberships in the following organizations as of March 4 and 5, 2016:

- Association of Research Libraries (ARL)[11];
- OCLC Research Library Partnership (OCLC-RLP)[12]; and
- Digital Library Federation (DLF)[13].

These organizations were selected due to mission statements focused on “research” and “libraries”, and each organization has a membership exceeding 100 libraries[14]. The study population was audited on October 5, 2016 by requesting the publicly-available HTML pages listed on each organization’s membership page and logging the research library’s webserver response. The study includes 448 unique URLs from 279 libraries in 16 countries. Geographically, the data set represents 344 unique URLs published by 211 US libraries, 30 URLs by 16 Canadian libraries, 33 URLs by 22 UK libraries and 41 URLs from 30 libraries in other countries. The full data set is available through Zenodo[15]. The process used to generate this data set, including scripts and documentation, is available for verification and replication as open source code through GitHub[16].

The following procedures are presented as step-by-step outlines, organized by research question.

**RQ1: do libraries implement HTTPS with proper redirect practices?**

In order to answer this research question, we completed three main steps of analysis for the websites in our study population.
Step 1. We determined whether the library offers a secure connection (HTTPS) between the user’s browser and the library’s website:

(1) Check for a digital certificate (HTTPS). This test was accomplished by requesting a secure connection to each URL.

- For example, www.unm.edu/libraries.html is a unique URL that can be requested with a non-secure (http://) or secure (https://) connection. For this test, the secure connection, www.unm.edu/libraries.html was requested and the server response was logged.

- The test is “true” if the Library webserver resolved to a URL containing “https://.”

Step 2. We determined whether the library has implemented a permanent redirect to enforce the use of secure connections, i.e. connecting via HTTPS.

(1) Check if the server redirects non-secure requests to secure connections of the page requested:

- For example, when the non-secure URL http://scholarworks.montana.edu/ is requested, does the webserver respond with an HTTP 301 permanent redirect to the secure URL https://scholarworks.montana.edu/?

- This test is “true” if the library’s webserver uses an HTTP 301 redirect of non-secure URL requests (http://) to secure (https://) versions of the URL.

Step 3. We determined whether the library redirects secure page requests to a non-secure version of the page.

(1) Check if the server redirects secure page request to a non-secure connection:

- For example, if a secure URL is requested (e.g. https://scholarworks.montana.edu/) does the webserver redirect the user to a non-secure version of the page (e.g. http://scholarworks.montana.edu)?

- This test is “true” if the library’s webserver redirects the user to a non-secure page (http://) without first delivering an “HTTP 404 Page Not Found” message.

R2: do libraries that use Google Analytics implement the available privacy-protection measures?

In order to answer this research question, we completed three main steps of analysis for the websites in our study population.

Step 1. We determined whether the library website has implemented Google Analytics. As a corollary to this step, we also determined whether the library website has implemented Google Tag Manager. Below are the test requirements for this step:

(1) Each webpage was analyzed for the presence of the Google Analytics tracking code and, separately, for the Google Tag Manager tracking code, identified by unique markers that met the following criteria:

- for Google Analytics, the presence of a character string unique to and required by either Universal Analytics or Classic Analytics[17]; and

- for Google Tag Manager, the presence of a character string unique to and required by Google Tag Manager[18].

The use of Google Analytics was determined to be true if any of the library home pages contained the character strings unique to and required by Universal Analytics, Classic Analytics or Tag Manager.
Step 2. If a website tested positive for the tracking codes of either Google Analytics or Google Tag Manager, we then determined if the website was using available features to protect user privacy. First, we determined whether the library implemented a secure HTTPS connection between the library web server and Google’s web server. To complete this step, we analyzed each website’s source code for the presence of either “forceSSL”[19] or www.googletagmanager.com[20].

Step 3. If the website is using Google Analytics, we determined whether it has implemented the Google Analytics IP anonymization feature to obfuscate user tracking. To complete this step, we analyzed each website’s source code for the presence of “anonymizelp[21].”

We tested for the presence of these two elements – forceSSL and anonymizelp – because these are the only two Google Analytics privacy mechanisms that were observable at the time of this writing.

Results and discussion
Our study results indicate that libraries are not doing all they can to protect user privacy on the web. Analysis of the results of our testing follows.

Results summary
- Of the libraries in our study population (n = 279), 173 (62 percent) have implemented basic encryption technology via HTTPS (see Figure 1).
- Of the libraries that have implemented HTTPS (n = 173), 56 (32 percent) implemented a permanent redirect from HTTP to HTTPS to ensure that HTTPS is used at all times when communicating with users (see Figure 2).
- Of the websites in our study population (n = 279), we also found that 43 (15 percent) have implemented a redirect in the inverse direction: from HTTPS to HTTP, thus allowing user activity to occur over a non-secure connection without informed consent – even when the user specifically requests a secure connection.
- Of the websites in our study population (n = 279), 245 (88 percent) have implemented either Google Analytics or Google Tag Manager (see Figure 3).
- Of the websites that implemented Google tracking code (n = 245), 3 (1 percent) had implemented HTTPS connections between the library’s web server and Google’s

![Figure 1. HTTPS implementation for academic library home pages](image)
web servers, 34 (14 percent) had implemented Google’s feature to obfuscate user identification via IP anonymization and 0 (0 percent) implemented both of these available privacy features (see Figure 4).

RQ1: do libraries implement HTTPS with proper redirect practices?

The first test determined whether the library has provided the opportunity for a secure HTTPS connection between the user’s browser and the library’s website. If so, did the library automatically redirect non-secure URL requests (i.e. http://) to a secure version of the URL (HTTPS) by way of a permanent webserver (HTTP 301) redirect? If a user requested a secure connection and one is not available, did the library inform the user that a secure connection is not available or did the library redirect the user to a non-secure connection without informed consent?

Our results indicate that HTTPS implementation on library websites is not widespread, and many libraries continue to offer non-secure connections without an automatically enforced redirect to a secure connection. In our study population, 62 percent had implemented HTTPS. Of those, only 32 percent automatically redirected non-secure
requests to secure requests. Furthermore, we found that 46 of the 106 websites without HTTPS available (43 percent) were redirecting secure requests made by users to non-secure URLs without notifying the user. This practice undermines privacy in a number of ways: first, by failing to inform the user that a secure HTTPS version of the page is not available; second, without obtaining informed consent before the user is offered non-secure web pages; and third, by increasing the risk of Man-In-The-Middle attacks of library website users connecting via non-secure Wi-Fi access points, such as coffee shops, or compromised wired network connections[22].

RQ2: do libraries that use Google Analytics implement the available privacy-protection measures?

We conducted three further tests: to measure the use of Google Analytics and Google Tag Manager across our study population; whether libraries had implemented a secure HTTPS connection between their websites and Google Analytics and/or Google Tag Manager; and whether libraries had activated Google Analytics IP anonymization. Our research results demonstrate that at a minimum 88 percent of the 279 academic library websites have implemented Google Analytics and/or Google Tag Manager, yet only 1 percent connect securely to Google via HTTPS, and only 14 percent have implemented Google Analytics IP anonymization. No library in our study activated both measures: HTTPS between their servers and Google’s servers, and IP anonymization.

In the face of these results, it is clear that libraries must take additional steps to ensure that our website practices are consistent with the professional library values of privacy and intellectual freedom as articulated through organizations such as the ALA, CNI, NISO, IFLA and the Library Freedom Project. Of the library websites in our study population, many offer secure web connections for users, but most do not enforce that secure connection, and some even force users into a non-secure connection without their informed consent. And of the many websites that have implemented Google Analytics, most have done so using non-secure web connections and without activating the available privacy-protection feature of IP anonymization.

Recommendations for practice

Major library organizations such as ALA, CNI, IFLA and NISO have articulated a professional set of principles that guide our work. These principles include privacy and
intellectual freedom. However, results from our study show that the web analytics practices of many academic libraries are not in line with our profession’s stated values. In order to realign toward privacy, we offer a set of practical recommendations for building more privacy-oriented library websites. Indeed, building websites that better protect user privacy can help libraries remain trusted sources of information and places of inquiry. We present our study results so that additional motivation can be generated toward building a more private web: by implementing HTTPS and automatically redirecting users to that secure URL, and by enhancing the privacy practices around Google Analytics and Google Tag Manager, which are so prevalent on library websites. We offer five recommended practices for enhancing user privacy on the web:

1. configure library web servers to use permanent redirects (301) to HTTPS using SSL certificates provided by trusted certificate authorities;
2. implement IP anonymization for Google Analytics;
3. provide user education related to online privacy;
4. obtain informed consent from users; and
5. conduct risk/benefit analyses when using third-party service providers.

These recommendations will be further explained in the following sections. The recommendations have been developed through a synthesis of relevant literature and our research results.

**Library webserver HTTPS**

HTTPS is a vital privacy-protecting mechanism. By providing secure web connections for our users, we can help protect their online behavior from data leakage and provide opportunity for informed consent before surveillance occurs. Implementing HTTPS can also assist with commercial search engine discovery, as Google announced in 2014 that it would begin to favor secure websites over those that are insecure (Askey and Arlitsch, 2015). Library-led efforts such as the Library Freedom Project and its Digital Privacy Pledge are raising awareness of web privacy and the value of HTTPS. The most effective method for protecting privacy calls for web servers to support ubiquitous encryption across all their domains, including all subdomains (Sivakorn et al., 2016). Tools to help implement and evaluate HTTPS include Open SSL and HTTPS Everywhere. Once HTTPS has been implemented, libraries can go further by ensuring that connections to their web servers occur securely via HTTPS, even when a connection is initiated through an insecure HTTP connection request. This can be accomplished by applying mechanisms that force secure connections via an HTTP 301 Redirect, which are available for WordPress, Apache, Drupal, and other major web publishing platforms (the precise process for implementing and enforcing HTTPS will vary according to web server platform and configurations). For the many libraries that use Google Analytics, forcing HTTPS provides the added privacy benefit of ensuring that user data transferred between the library’s servers and Google’s servers occurs over a secure connection.

**IP anonymization for Google Analytics**

Internet protocol Anonymization, also known as IP masking, is a customization to the Google Analytics tracking code that changes how Google uses and stores the IP addresses of website users. This setting gives website owners using Google Analytics the option to tell Google to use only a portion of a user’s IP address, thus allowing for geolocation without identification of individual users or their activity. In effect, anonymizing the IP address helps protect specific identification of library website users.
User education

The privacy landscape is shifting quickly as networked technologies experience widespread development and adoption. Users want more control over tracking, though they are often unsure how to protect themselves or are distrustful of readily-available tools (Melicher et al., 2015). As institutions for public good, libraries can help users understand the privacy implications of the contemporary web and, where possible, libraries can provide realistic means by which users can mitigate privacy threats. This includes informing users of privacy-based search engines such as StartPage and DuckDuckGo, IP address obfuscation through Tor relays (Acar et al., 2014; Macrina, 2015a; Huang and Bashir, 2016), library workshops that advocate and educate for privacy-related topics (Gressel, 2014; Macrina, 2015b), emerging standards such as the Tracking Preference Expression (Do Not Track – DNT)[30], and independent, third-party browser tools that can help mitigate tracking, such as Disconnect and Better[31]. With third-party tracking so prevalent and sophisticated, browser add-ons and extensions are imperfect tools for ensuring privacy (Libert, 2015; Merzdovnik et al., 2017; Starov and Nikiforakis, 2017), but they do add some measure of protection.

Informed consent

Informed consent increases transparency with users regarding library web tracking and privacy practices. When there is a lack of clear communication around web tracking, we can compromise the privacy of library users. Certain mechanisms currently exist to help educate users in context, such as cookie-consent notifications. This approach represents a feasible option, in that it can effectively inform users that traffic is being monitored by a third-party (Shih et al., 2015). A cookie-consent notification on library websites could include, for example, a set of call-to-action buttons for accepting or declining the presence of cookies, along with follow-up links to a library’s privacy policy page. Privacy policies are also useful tools for informing users about web tracking (Magi, 2007; Nichols Hess et al., 2015; Kritikos and Zimmer, 2017) and for building trust with users (Aïmeur et al., 2016). Privacy policies should be visible on the website and written in clear, specific language (Capistrano and Chen, 2015).

Risk/benefit analysis

Lastly, libraries and their users will benefit from a periodic risk/benefit analysis of third-party services, including analytics services such as Google Analytics. With continued projections for declining public funding for higher education, academic libraries face the pressure of demonstrating value by measuring and assessing the use of services and resources (Saunders, 2015; Alamuddin et al., 2016). Google Analytics is a powerful tool for assessment, yet its connection to the vast network of third-party web trackers threatens to compromise the core library value of intellectual freedom. As Zimmer (2013, p. 56) remarks, “Libraries should minimize the use of Web cookies, bugs, and other tracking technologies.” Other non-commercial web analytics services, such as Piwik, present alternatives to Google Analytics (Chandler and Wallace, 2016), although their use may be accompanied by increased administrative overhead to the library. Through a critical examination of the usage of these services, it is possible to balance the risks and the benefits to both the library and users.

Limitations and future directions

Our study design required explicit presence of the Google Analytics tracking script embedded into the homepage HTML code. This requirement produces conservative results; for instance, we discovered one case where the use of Google Tracking code was not obvious because the script was encapsulated in a separate file using a non-standard naming convention unique to
the website. This type of non-standard Google Tracking code implementation was not accounted for in our study results because our tests could not systematically identify it. Further studies should examine the prevalence, behavior and privacy impact of additional trackers present on library websites, including those from Google Tag Manager, Google AdSense and third-party library vendors. User expectation studies will help provide additional depth and context to this research (Anton et al., 2010; Lin et al., 2012).

Our current study was limited to libraries found to have membership in ARL, DLF and OCLC-RLP on March 4 and 5, 2016. Future studies into HTTPS and Google Analytics can expand the study population to include academic libraries outside those membership organizations, public libraries, tribal college libraries and special libraries. Ultimately, this research is just a starting point for understanding the breadth and depth of third-party analytics tracking on library websites. We expect future research to investigate this topic to a greater extent, both in terms of the study population and the technical analysis.

Conclusion
As a profession with a long-held value of intellectual freedom, libraries should act to protect privacy on the web for their users. The ALA, CNI, NISO, IFLA and the Library Freedom Project all champion web privacy, yet the actual practices of library websites are not in alignment with their stated values of privacy. Results from our empirical study indicate that many libraries undermine user privacy by not offering secure connections to their websites, and that most libraries have implemented the third-party analytics service Google Analytics—which potentially exposes users to data leakage—but have not activated the available privacy-protection features of this tool. We conclude by offering five practical recommendations for enhancing user privacy: HTTPS, IP anonymization, user education, informed consent and risk/benefit analysis.

Data set availability
Our research data set is based on information gathered from 448 unique URLs associated with 279 North American and international academic library organizations that have membership in one or more of the following organizations: ARL, DLF or the OCLC-RLP. The following Web server responses were recorded for each URL: error response, redirect response, resolved URL, HTML cache and request protocol (HTTP or HTTPS). The data set also includes results of analysis of each URL’s cached HTML used to locate and evaluate the JavaScript snippets associated with third-party web analytics trackers. The Python scripts used to make these analyses are included in the data set. The research data set and software is available through Zenodo at https://doi.org/10.5281/zenodo.1323403.

Acknowledgments
O’Brien conceived of the study, led the research design, collected data, performed the analyses, interpreted the results and critically revised the manuscript. Young participated in the research design, interpreted the results and drafted the manuscript. Arlitsch participated in the research design, interpreted the results and critically revised the manuscript. Benedict collected data, performed the analyses, interpreted the result, and critically revised the manuscript. All authors read and approved the final manuscript.

This research was conducted as part of a grant generously funded by the Institute of Museum and Library Services (IMLS). “Measuring Up: Assessing Accuracy of Reported Use and Impact of Digital Repositories” was a three-year research project led by Montana State University, in partnership with the Association of Research Libraries, OCLC Research, and the University of New Mexico (IMLS Log Number: LG-06-14-0090-14). The grant proposal narrative may be found at http://scholarworks.montana.edu/xmlui/handle/1/8924.
Notes

4. www.ala.org/advocacy/privacyconfidentiality
5. www.ala.org/advocacy/intfreedom/librarybill/interpretations/privacy
9. www.ifla.org/node/9803
11. www.arl.org/membership/list-of-arl-members
12. www.oclc.org/research/partnership/roster.html
13. www.diglib.org/members/
20. At the time of this writing, we could not locate any Google Tag Manager documentation describing features related to protecting patron privacy. Therefore, the presence of Google Tag Manager on a website indicates that no privacy-protection measures are in place
24. www.openssl.org; https://www.eff.org/https-everywhere
References


Corresponding author
Scott W.H. Young can be contacted at: swyoung@montana.edu

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Image search and retrieval problems in web search engines
A case study of Persian language writing style challenges

Yaghoub Norouzi
Department of Knowledge and Information Science, University of Qom, Qom, Islamic Republic of Iran, and

Hoda Homavandi
Department of Knowledge and Information Science, University of Tehran, Tehran, Islamic Republic of Iran

Abstract
Purpose – The purpose of this paper is to investigate image search and retrieval problems in selected search engines in relation to Persian writing style challenges.
Design/methodology/approach – This study is an applied one, and to answer the questions the authors used an evaluative research method. The aim of the research is to explore the morphological and semantic problems of Persian language in connection with image search and retrieval among the three major and widespread search engines: Google, Yahoo, and Bing. In order to collect the data, a checklist designed by the researcher was used and then the data were analyzed by descriptive and inferential statistics.
Findings – The results indicate that Google, Yahoo, and Bing search engines do not pay enough attention to morphological and semantic features of Persian language in image search and retrieval. This research reveals that six groups of Persian language features include derived words, derived/compound words, Persian and Arabic Plural words, use of dotted T and the use of spoken language and polysemy, which are the major problems in this area. In addition, the results suggest that Google is the best search engine of all in terms of compatibility with Persian language features.
Originality/value – This study investigated some new aspects of the above-mentioned subject through combining morphological and semantic aspects of Persian language with image search and retrieval. Therefore, this study is an interdisciplinary research, the results of which would help both to offer some solutions and to carry out similar research on this subject area. This study will also fill a gap in research studies conducted so far in this area in Farsi language, especially in image search and retrieval. Moreover, findings of this study can help to bridge the gap between the user’s questions and search engines (systems) retrievals. In addition, the methodology of this paper provides a framework for further research on image search and retrieval in databases and search engines.
Keywords Information retrieval, Image retrieval, Search engine, Persian language, Writing styles
Paper type Case study

1. Introduction
The web is increasingly becoming a storage for growing masses of information, and “it has become an ocean of all kinds of data, making any query into the huge information reservoir extremely difficult” (Isfandyari Moghaddam, 2007). Search engines organize huge amounts of incoming information into the web partly and provide them to users in different formats such as text, audio, picture, multimedia, etc. In this regard, pictures are very important as one picture can sometimes be more illustrative than 100 words. Therefore, “concurrent with the development of web and access to different types of Photography techniques like digital cameras and image scanners, the size of digital image collections is increasing” (Liu et al., 2007).

Many image retrieval approaches have been widely used to search for a great number of internet images; however, it is still difficult to retrieve images that satisfy the query intentions of users (Guo and et al., 2016). Many factors can overwhelm the searcher
who is trying to retrieve images. Consequently, searching for images is still a challenge for the majority of individuals, especially for images associated with texts written in different languages which are unfamiliar to users (Ménard, 2011). Moreover, many people need to do this for different purposes, ranging from a simple search for a group of images related to a general subject to a specific search which may be done by a scientist. “Efficient image search tools by Google (http://images.google.com), MSN (http://search.msn.com/images) and Yahoo (http://images.search.yahoo.com/) are clear evidence that image indexing and searching is common and widespread in today’s visual culture” (Neugebauer, 2010).

The “multilinguality” of web content provides opportunities for users to directly access and use previously incomprehensible sources of web information; nevertheless, web users find it difficult to take advantage of these opportunities when the online information access systems are monolingual (Bao and Chen, 2009). In recent years, the number of non-English resources on the web has been growing rapidly, and tools that can build specialized search engines in different languages are thus highly desired (Chau and et al., 2008). Based on the Internet World Stats’ (2016) latest estimates for internet users by language, 67.8 percent of users use English and 32.2 percent of them are Non-English-speaking users. This feature makes the web a multilingual and multicultural information space. Although there are several search engines for facilitating web search, it appears that their attention to non-English languages in comparison with English language is not enough (Lazarinis, 2007a). Persian users like Iranians are no exception to this rule. Based on Internet National Development Management Center’s (2015) latest estimates for internet penetration in Iran, 82.12 percent of the people use the internet to meet their information needs; therefore, their linguistic and writing problems must be taken into consideration. Persian language encompasses a broad range of speakers in Iran and some of its neighboring countries; consequently, codification of rules and criteria for Persian language is very important, especially when we take into account the increasing use of computer in Persian language and script area. Besides, concerns about the risk of dispersion and applying different and antithetical styles have been growing over the past years (Academy of Persian Language and Literature, 2010).

As it was mentioned above, the language of search is one of the most crucial factors whose significance in relation to the image search and retrieval is doubled because of the fact that one of the fundamental differences between textual and visual information is the nature of their retrieval process. Thus, the problem of image retrieval is an increasingly active area for research and development (Murthy and et al., 2010).

Retrieving of textual information is based on the context and usage whereas “Most of image search engines are relying on the text for indexing images, and this means that the quality of their results depends on the quality of the text associated with the image (such as file name, text next to the image, page title or HTML tag)” (TASI, 2008)[1]. Languages like Persian have some features and complexities that disregarding them could lead to problems in information searching and retrieving. As a result, the present study attempts to investigate image search and retrieval problems in connection with Persian writing styles in selected search engines, namely Google, Bing and Yahoo. These search engines were selected based on their ranking among the most common engines[2] and their backing of Persian language searching. We focused on images as there have been many conflicts over content-based retrieval systems that retrieve images based on visual features (like shape, size and color) and concept-based image retrieval systems for years (Lazarinis, 2008). So, automatically finding images relevant to a textual query remains a very challenging task (Meng, 2015). Selected search engines retrieve images based on textual information of them which is called concept-based image retrieving. So focus on image in this study could show search engines’ potentials (as an example of concept-based image retrieval systems) in relation to their approach to image storing.
(tagging and indexing) and retrieving which is completely different in comparison to their approach to retrieving texts. Besides, in the semantic part of the language, recording the image results is more illustrative than texts; for example, homonymous words like “شیر” with tree different meanings (milk, lion and tap) can be easily identified by image search. Also there are differences between webpages and image pages results, for example in the recognition of short vowels like (ﻗسکن-مسكن-فسکن), two spelling words such as (اژوقه-ازوقه) and loanwords and their equivalent like (کامیون-رایانه)

In order to identify major problems (semantic and morphological), we compared the selected search engines operations and finally determined the most efficient search engine among them. This study obviously aims to answer the following questions:

RQ1. What are the major problems of Persian language as far as writing features are concerned (morphological and semantic) in the selected words associated with image search and retrieval in search engines (Google, Yahoo and Bing)?

RQ2. What is the most efficient search engine in terms of compatibility with Persian language?

Thus, this paper can both test instances practically and propose some solutions to these problems based on research findings so as to help different groups involved in this process, including users, search engines and so on. In other words, the present paper shows some linguistic issues concerning user–computer interaction. Also, depending on information needs of users, it provides the possibility to compare linguistics features especially in relation to images from the viewpoint of morphologic and semantic problems.

2. Literature review
Persian is one of the oldest languages in the world. It is a criterion language whose history dates back to the sixth century BC. This language is a part of Indo–European languages. Ancient Persian, which was previously the language of the great empire of Persia, was widely used from Mediterranean to Indus. In the second century BC, Iranians managed to invent their unique alphabet called Pahlavi, which maintained its efficiency until the victory and conquest of Islam. After that, Arabic script with some extra characters was used for showing specific sounds. This language is called Farsi in Iran and Dari in Afghanistan. Also, a kind of Farsi which is called Tajik is common in Tajikistan (Katzner, 1997)

Today, Persian language is taught in more than 30 universities around the world.

Persian language has its own particular properties which originate from its nature. This language has been greatly influenced by Arabic and has also adopted some words from other languages like English, Turkish, etc. From experts’ point of view, there are some significant assortments and attributes in connection with writing features of Persian language, some of which are presented below and they are also appearing in the checklist.

Morphological and semantic features
Morphological features, in short, include the following:

- diversity of translating and recording foreign words, lack of integration in the use of imported words and their proposed equivalents by the Academy of Persian language and literature, coherence and discontinuity in writing derived and compound words, plural signs and multiplicity of plural signs (like: ها، ان، ات، ات، ون) and irregular plurals of Arabic words;
- using different forms of dotted “TA” (ﺓ);
- style of writing Hamza (沙特) at the beginning, middle and end of words with different seats (like: ﯾ، ﯽ) and so on, Different forms of writing limited and hidden “A” (الف),
use or non-use of punctuation in connection with the words which have similar writing forms and different pronunciations;

- application or removal of some Arabic signs (short vowels) (like: ً - ٌ - ٍ - ّ - َ - ُ - ِ) and mediator “y” (۶) can also cause some problems in information search and retrieval process;

- words with two spellings (the words with similar phonemes and different writing forms), using spoken language in texts, additional “e” (۶) and its substitutes in Farsi; and

- style of writing unpronounced “h” (۶), existence of innumerable dots and dents.

The above-mentioned problems make users confused in searching for information and in particular image searching and retrieving in search engines, and when they select and write special forms of words, they instinctively miss other forms of words.

Also, the two major problems in semantic communications are as follows:

- polysemy: when one word has different meanings; and

- synonymy: when different words have the same and unique meaning (Hosseini Beheshti, 2007).

Both of these phenomena disrupt identification relationship in information retrieval systems. Some of the semantic features include homonymous[5] (the words with similar morphological features), polysemy[6] and synonymous words. All of these cases can cause diversities in the semantic area of language, leading to some problems in searching words and sometimes search engines retrieve something different from what users have in mind. For example, this is what occurs in searching words like “شیر” (it means lion, milk and tap).

Overall, some problems such as writing styles of Latin words, different spellings of derived/compound words, plural signs, vowels, having a lot of loanwords from other languages like Arabic, English, Russian, Turkish, etc., are specific to Persian language. In this regard, many of Persian writing system features are created by its special alphabet, which is completely different from Latin ones because the majority of letters in a word connect to each other and as the script is cursive, the appearance of a letter changes depending on its position: isolated, initial (joined on the left), medial (joined on both sides) and final (joined on the right) of a word. The other feature has to do with the use of a group of vowels in Persian language as well as in modern Persian script. Historically, short vowels are usually not written, and only the historically long ones are represented in the text; words distinguished from each other only by short vowels are ambiguous in writing. These kinds of problems are related to the use of Arabic alphabet since modern Iranian Persian is written using a modified variant of the Arabic alphabet, but it uses different pronunciations and additional letters not found in Arabic. As a consequence, finding solutions to mitigate these issues seems necessary as many users may not be familiar enough with the English language or may be seeking information in other languages.

There are many studies that have evaluated search engines in terms of language criteria. Conducted surveys on the related literature have shown that in most cases researchers have considered morphological issues in connection with information retrieval in well-known search engines. Also, among Persian studies there is no research that has addressed both the morphological and semantic aspects of language related to image retrieving, and this is the distinction between this research and other studies.

Apparently the capabilities of local and general search engines for non-English languages are not satisfactory. Search engines do not take into account the morphology of the queries and retrieve different results in queries that differ in morphology, and thus users’ searches end up in failure. Lazarinis (2008) explored the behavior of the major search
engines (Google, Yahoo, MSN) in image retrieval using Greek text queries. The initial evaluation reveals that search engines retrieve different results in queries that differ in morphology or in grammar but still express exactly the same information need, and the omission of diacritics affects the retrieval negatively as well. As these kinds of language problems may appear in other parts of the web, Lazarinis (2007b) developed the domain of his research to commercial websites, and the evaluation showed that the local search engines do not take into account the morphology of the queries, which is significant in non-English and non-Latin languages like Greek, and thus, users’ searches prove to be a failure. Many of popular search engines claim that they can support several languages, but it has not been proven until now and in this study. For instance, Zhang and Lin (2007) in their research investigated the multiple language support features in internet search engines. The findings revealed that Google, EZ2 Find, and Online link are the search engines with the best multiple language support features in their categories, and multiple language support features in search engines remain at the lexical level. In this regard, Lewandowski (2008) carried out a study to test the ability of the major search engines, namely Google, Yahoo, MSN and Ask to distinguish between German and English language documents. According to this study, while none of the search engines encountered any problems providing results in the language of the interface used, both Google and MSN faced difficulties when the results were restricted to a foreign language. Namik and Bitirim (2015) illustrated the performance evaluation of the three-international web search engines including Google, Bing and Yahoo based on the three languages English, Arabic and Turkish in six various categories. The results indicated that the web search engines had a better performance in English document retrieval compared to Arabic and Turkish document retrieval.

The ability to distinguish features of different languages, especially languages with non-Latin roots, dramatically influences the process of searching and retrieving information. Bar-Ilan and Gutman (2005) conducted a study on the capabilities of search engines for non-English languages. As a test case, they examined four languages: Russian, French, Hungarian and Hebrew. Their results indicate that in the examined cases the general search engines ignore the special characteristics of non-English languages, and sometimes they do not even handle diacritics well. Also, AlSobh et al. (2010) built several tools to evaluate the effect of words and statements in several languages on the sequence or the ranking of the retrieved results in search engines. The results from their studies showed that there are some limitations and situations where search for results can be biased toward the popularity of the website language and/or location. In association with Arabic languages (which have similarities with Farsi language), Tawileh and et al. (2010) explored how Arab internet users can find the information in their mother tongue on the web. The core finding was that Google performed almost all the times better than the other engines. The other finding was the significant difference between search results and their descriptions for all tested engines. Some researchers try to cope with these problems in a different manner; for instance, Hammo (2009) did a study to enhance retrieval effectiveness of search engines for diacritised Arabic documents. This study showed that applying query extent techniques improve the results dramatically, and hence, using a stemmer and a thesaurus outperformed the original search using the full form of words.

Finally, some studies have been conducted on Persian language. Norouzi and Homavandi (2014) in their study showed that the characteristics of Persian writing style significantly impact the number of retrieved results in IRANDOC(7) database. In a similar study, Akhshik and Fattahi (2012) analyzed problems with information storage and retrieval in two Persian databases (IRANDOC and RICEST(8)) with respect to writing different forms of some Persian words. Their results suggested that RICEST database is only searching the initial recorded form of words leading to retrieval of information and that
searching other forms had no result. Also, in IRANDOC database 58 percent of words were 
retrieved in spite of using different forms of words. These studies revealed that even local 
databases do not pay adequate attention to language features. Moreover, Abdolahi-Noorali 
and Jowkar (2009) in their research examined the problems with which Persian users are 
faced in searching different forms of words. Findings showed that major search engines do 
not consider different Persian writing styles in order to improve searching.

Overall, review of the related literature implies that despite the enormous acceptability 
and use of search engines and database for finding different kinds of information by users 
and the importance of linguistic factors and problems in this area, search engines ignore 
some of them. Although in many cases researchers studied those factors, as for Persian 
language, due to its unique nature and features as well as the increasing production of 
Persian webpages and blogs, more research still needs to be conducted. For example, as yet 
no research has been conducted in this area in which besides morphological features, 
semantic issues are considered, and therefore, these kinds of studies are likely to result in 
the improvement and creation of national search engines too; therefore, this study as the 
first work in the Persian language area only intends to investigate the problems of image 
retrieval specifically in the area of this language.

3. Methodology

This study is an applied research, and in order to answer the research questions, we made 
use of evaluative research method. First, after reviewing research studies, three search 
engines including Google, Yahoo and Bing were selected, which not only are among the 
most widely used search engines, but also enable their users to search images in Farsi. 
Then, to collect data, a researcher-made checklist was utilized. We examined the research 
checklists such as those made by Abdolahi-Noorali and Jowkar (2009) and Akhshik and 
Fattahi (2012) to focus especially on the morphological aspects. The checklist we designed 
includes 36 words, each of which represents one of the semantic and morphological 
characteristics and problems. These words were approved by the Professors of Persian 
Language and Literature, and indexing and retrieving information experts. This checklist 
has both semantic and morphological parts, with different forms of writing words given in 
morphology part. Altogether, it includes 31 words in morphology part and 5 words in 
semantic part. Some of the mentioned morphological problems such as abstract concepts or 
words, which do not have proper visual evidence for data analysis, could not be used 
as data collection tool. Thus, we selected features which include non-abstract and tangible 
concepts with clarity and lack of ambiguity in order to achieve acceptable results in data 
analysis stage[9].

Afterward, in order to collect data, on a specific date each of the variables (the words in 
the checklist) was run in image search of selected search engines separately (in August 
2013), then for each search engine and each word some tables were drawn for recording 
data. These tables in morphology part include different forms of each word and their 
number in 50 retrieved first results[10] (according to image tags). In semantic part and 
regarding words which do not have different writing forms, like "شیر" (milk, tap and lion) 
and "شور" (salty) we recorded the number of images that included different meanings of 
words in tables[11].

To carry out the above step, we recorded the number of images that had true match with 
keywords in tables. Then other forms of keywords were recorded and finally images with 
other tags (e.g. image with English tags) were counted. Concerning the semantic part, first 
the number of images which contained the same meaning with the run keyword was 
recorded by the researcher, and after that other meanings of keyword were counted and 
recorded in tables, and finally the number of images which have no related meaning to the 
desired keyword were counted. Based on this information, we used Friedman test to
determine the mean difference. The more differences in retrieved result we have, the more mean difference score is shown, and in this case, we concluded that search engines with higher score can identify retrieved different morphologic forms of every single word, and also can retrieve different meanings in semantic part. For example, when we searched “پرسنچه” in Google, if beside this keyword it retrieved images with “پرسنچه” so, it is difference score gets higher and therefore it was a good point for it. Likewise, in the meaning part if images for a keyword like “نشئ” show different meanings of it (images of lion, tap and milk), it is considered as a good point for search engine because in this way a user is made aware of these different forms and meanings and can refine her/his search.

Data analysis
In this study, data were analyzed in both descriptive and inferential statistical levels. To avoid presenting more than 30 tables (four different forms and meanings of keywords are shown in checklist), at first we ran every single keyword in three selected search engines separately and then counted and recorded the number of images with different tags or meanings in mentioned tables. Finally, total percentage frequency was calculated for each keyword and search engine separately. Based on this result in the next stage, $\chi^2$ test was conducted. We used the $\chi^2$ test to identify major problems of Persian language (morphological and semantic) in relation to image search and retrieval in selected search engines (Google, Yahoo and Bing). With this test, we try to identify whether there is a significant relation between Persian language’s features and image search and retrieval in selected search engines (the significance level of variables). Finally, the keywords which have significance levels under 0.05 are considered as major problems of Persian language with image search and retrieval in search engines. We also used Friedman test to identify the most efficient search engine in terms of compatibility with Persian language. Based on average rank of each keyword, we judge how search engines are compatible with Persian language. In relation to every keyword, search engines with the highest average rank among the three are considered more capable than others. To this end, the statistical software SPSS, version 16, was used and for the interpretation, we utilized Microsoft Word and Excel.

In addition to identifying major problems related to Persian language and its morphological and semantic features, we sought to recognize the differences between search engines in terms of image search and retrieval and their compatibility with Persian language. Finally, having analyzed the results, we selected the most efficient search engine from among them.

4. Findings

RQ1. What are the major problems of Persian language in regard to writing characteristics (morphological and semantic) in the selected words associated with image search and retrieval in search engines (Google, Yahoo and Bing)?

To answer this question, we ran keywords of the checklist into image search of the above-mentioned search engines and then results were counted and recorded. After statistical analysis and using $\chi^2$ test, the $\chi^2$ value, degrees of freedom, the significance level and $\phi$ coefficient in the morphology part were calculated for each keyword. The results for keywords as shown in Table I (each of these words represented one of the semantic and morphological features of Persian language) were obtained.

Different users select various forms of foreign words mostly without knowing and try other forms to improve their search results. These kinds of words, which originate from other languages such as Arabic, English and Turkish, are very common in Persian. As Table I indicates, in regard to words with significance level more than $\alpha = 0.05$ and with
0.95 reliability, there is no significant relationship between the two mentioned morphological features and image retrieval problems related to search engines. These kinds of derived and compound words in Persian are numerous, and usage of them depends on the users’ writing styles while they affect the search results. As Table II indicates, searching different forms of the words “Postman” and “Training college” (which have significance levels less than $\alpha = 0.05$) dramatically affects the retrieved results.

There are different signs to plural words, like Arabic signs or Persian ones, which causes problems in searching and missing some results as we can see about words “School” and “Books” with a significance level less than $\alpha = 0.05$. Also, according to Table III findings, using formal or colloquial forms of words influences retrieved images significantly.

Most of different forms of the words in Table IV are derived from a combination of Persian and Arabic letters and signs that lead to retrieving various results for one concept, like “Object”, which has a significance level less than $\alpha = 0.05$.

Writing or ignoring short vowels causes some problems in searching, especially for Persian language users, who do not usually use short vowels; this problem also results from Arabic words, as indicated in examples of Table V.

As it is evident in Tables I–V, as to words that have a significance level more than $\alpha = 0.05$ with confidence coefficient 0.95, there is not a significant relationship between morphological and semantic features and problems associated with image search and retrieval in search engines. But, significance level for different forms of keywords include

<table>
<thead>
<tr>
<th>Morphological problems</th>
<th>Keywords</th>
<th>English equivalent</th>
<th>$\chi^2$ value</th>
<th>df</th>
<th>Significance level</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing styles of Latin words</td>
<td>انفلوانزا - انفلوانزا</td>
<td>Influenza</td>
<td>3.67</td>
<td>4</td>
<td>0.452</td>
<td>Not significant</td>
</tr>
<tr>
<td>Loanwords and their equivalent</td>
<td>نیایوش - نیایوش</td>
<td>Titanium</td>
<td>0.946</td>
<td>2</td>
<td>0.623</td>
<td>Not significant</td>
</tr>
<tr>
<td></td>
<td>کامیورن - رایانه</td>
<td>Computer</td>
<td>1.13</td>
<td>2</td>
<td>0.567</td>
<td>Not significant</td>
</tr>
<tr>
<td></td>
<td>سیستم - نظام - سامان</td>
<td>System</td>
<td>1.074</td>
<td>4</td>
<td>0.89</td>
<td>Not significant</td>
</tr>
</tbody>
</table>

Table I. Results of $\chi^2$ test related to different forms of keywords in morphology part (foreign words)

<table>
<thead>
<tr>
<th>Morphological problems</th>
<th>Keywords</th>
<th>English equivalent</th>
<th>$\chi^2$ value</th>
<th>df</th>
<th>Significance level</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derived words</td>
<td>پستمن - یستم - چی</td>
<td>Postman</td>
<td>7.717</td>
<td>2</td>
<td>0.021</td>
<td>Significant</td>
</tr>
<tr>
<td>Compound words</td>
<td>کالجات - کالج</td>
<td>Library</td>
<td>1.19</td>
<td>2</td>
<td>0.551</td>
<td>Not significant</td>
</tr>
<tr>
<td>Derived/compound words</td>
<td>دانشسرا - دانش سرا</td>
<td>Training college</td>
<td>39.25</td>
<td>2</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td></td>
<td>قوانی - قن اوری</td>
<td>Technology</td>
<td>3.96</td>
<td>2</td>
<td>0.138</td>
<td>Not significant</td>
</tr>
</tbody>
</table>

Table II. Results of $\chi^2$ test related to different forms of keywords in morphology part (kinds of derived and compound words)

<table>
<thead>
<tr>
<th>Morphological problems</th>
<th>Keywords</th>
<th>English equivalent</th>
<th>$\chi^2$ value</th>
<th>df</th>
<th>Significance level</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plural signs</td>
<td>کل ها - کلیا</td>
<td>Flowers</td>
<td>5.41</td>
<td>2</td>
<td>0.067</td>
<td>Not significant</td>
</tr>
<tr>
<td>Persian and Arabic plural words</td>
<td>مدرس - مردیس</td>
<td>Schools</td>
<td>7.14</td>
<td>2</td>
<td>0.028</td>
<td>Significant</td>
</tr>
<tr>
<td></td>
<td>کتاب - کب</td>
<td>Books</td>
<td>12.11</td>
<td>2</td>
<td>0.002</td>
<td>Significant</td>
</tr>
<tr>
<td>Two spelling words</td>
<td>آدفک - آوه</td>
<td>Purveyance</td>
<td>1.94</td>
<td>2</td>
<td>0.327</td>
<td>Not significant</td>
</tr>
<tr>
<td></td>
<td>تهران - طهران</td>
<td>Tehran</td>
<td>0.918</td>
<td>2</td>
<td>0.63</td>
<td>Not significant</td>
</tr>
<tr>
<td>Use of spoken language</td>
<td>خانه - خو</td>
<td>Home</td>
<td>0.32</td>
<td>2</td>
<td>0.196</td>
<td>Significant</td>
</tr>
</tbody>
</table>

Table III. Results of $\chi^2$ test related to different forms of keywords in morphology part (different forms of plural words, words with two spellings and colloquial words)
OIR 42,6

Table IV.
Results of $\chi^2$ test related to different forms of keywords in morphology part (style of writing letters)

<table>
<thead>
<tr>
<th>Morphological problems</th>
<th>Keywords</th>
<th>English equivalent</th>
<th>$\chi^2$ value</th>
<th>df</th>
<th>Significance level</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manner of writing short A</td>
<td>Kasra - Kasra</td>
<td></td>
<td>0.009</td>
<td>2</td>
<td>0.96</td>
<td>Not significant</td>
</tr>
<tr>
<td>Use of dotted T</td>
<td>Mosques - Mosques</td>
<td></td>
<td>0.001</td>
<td>2</td>
<td>0.999</td>
<td>Not significant</td>
</tr>
<tr>
<td>Hamza (a) at the beginning, middle and end of words</td>
<td>Article - Article</td>
<td>25.57</td>
<td>2</td>
<td>0.000</td>
<td>Significant</td>
<td></td>
</tr>
<tr>
<td>Additional &quot;e&quot; () and its substitutes</td>
<td>Signature - Signature</td>
<td>3.03</td>
<td>2</td>
<td>0.129</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Use or non-use of &quot;y&quot; and Hamza relocation in Persian words</td>
<td>Boss - Boss</td>
<td>0.008</td>
<td>2</td>
<td>0.99</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>Style of writing un</td>
<td>Muezzen - Muezzen</td>
<td>0.057</td>
<td>2</td>
<td>1</td>
<td>Not significant</td>
<td></td>
</tr>
<tr>
<td>mediated &quot;y&quot; (o) and mediator &quot;y&quot; (y)</td>
<td>Officer - Officer</td>
<td>0.603</td>
<td>2</td>
<td>0.74</td>
<td>Not significant</td>
<td></td>
</tr>
</tbody>
</table>

Table V.
Results of $\chi^2$ test related to different forms of keywords in morphology part (short vowels)

<table>
<thead>
<tr>
<th>Morphological problems</th>
<th>Keywords</th>
<th>English equivalent</th>
<th>$\chi^2$ value</th>
<th>df</th>
<th>Significance level</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use or non-use of short vowels</td>
<td>Home, painkiller</td>
<td></td>
<td>0.038</td>
<td>2</td>
<td>0.999</td>
<td>Not significant</td>
</tr>
<tr>
<td>Use or non-use of short vowels</td>
<td>Butter, globes</td>
<td></td>
<td>2.72</td>
<td>4</td>
<td>0.605</td>
<td>Not significant</td>
</tr>
<tr>
<td>Use or non-use of short vowels</td>
<td>Emerald</td>
<td></td>
<td>1.65</td>
<td>2</td>
<td>0.431</td>
<td>Not significant</td>
</tr>
<tr>
<td>Use or non-use of short vowels</td>
<td>Scout</td>
<td></td>
<td>0.458</td>
<td>2</td>
<td>0.795</td>
<td>Not significant</td>
</tr>
</tbody>
</table>

“دشتم، بهشت” (derived words), “دانشسر داشت سر” (derived/compound words), “مدرس، کتاب” (Persian and Arabic plural words), “نام، مقاله” (use of dotted T) and “خانه، خودن” (use of spoken language) is less than $\alpha = 0.05$. Therefore, with coefficient 0.95 there is a significant relationship between morphological features and problems associated with image search and retrieval in search engines. On the whole, problems related to derived words, derived/compound words, Persian and Arabic plural words, use of dotted T and the use of spoken language are the major problems of image search and retrieval in search engines.

In Table VI, the amount of $\chi^2$, degrees of freedom and significance level for each of the words in morphology part were calculated. As it is evident in Table VI, as words that have significance level more than $\alpha = 0.05$ with confidence coefficient 0.95, there is not a significant relationship between morphological and semantic features and problems associated with image search and retrieval in search engines. However, the significance level for different meanings of “قلب” (polysemny) is less than $\alpha = 0.05$; thus, with coefficient 0.95 there is a significant relationship between semantic features and problems pertaining to image search and retrieval in search engines. Overall, the problems related to multiple meaning are among the main problems that affect image search and retrieval in search engines.

RQ2. What is the most efficient search engine in terms of compatibility with Persian language features?

To answer this question, we ran keywords of the checklist into the image search of the mentioned search engines and then results were counted and recorded. After statistical
analysis and using Friedman test, the results for keywords as shown in Table VII (each of these words represented one of semantic and morphological features of Persian language) were obtained. Retrieving different forms and meanings of each keyword among results was assumed as a reason for the search engine’s efficiency because in this way user becomes aware of these various morphologic and semantic forms.

As it can be observed in Table VII, as for different forms of writing foreign words, Yahoo and Google perform better than Bing, although they underperform when adjusting to these kinds of words.

According to Table VIII, as to different forms of derived and compound words, Yahoo and Google perform better than Bing, although none of them have any mechanism to inform users of trying other forms of the words.

In this regard, as Table IX findings illustrate, Google has the best performance in terms of covering different forms of plural words. Regarding the words with two spellings and use of spoken language, Bing and Google are better than Yahoo, but do not offer any substitute spellings, or do not inform users of formal forms of the word when it comes to colloquial words.

<table>
<thead>
<tr>
<th>Semantic problems</th>
<th>Keywords</th>
<th>English equivalent</th>
<th>( \chi^2 ) value</th>
<th>Significance level</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homonymous</td>
<td>شیر(خوان)، شیر (لبی)، شیر(…)</td>
<td>Milk, lion, tap</td>
<td>1.63</td>
<td>2</td>
<td>Not significant</td>
</tr>
<tr>
<td>Similar words with different meanings</td>
<td>شور(طبیعی)، شور(شبیه)</td>
<td>Passion, salty</td>
<td>1.04</td>
<td>2</td>
<td>Not significant</td>
</tr>
<tr>
<td>Polysemy</td>
<td>قلب(عصبی)، قلب (ختانی و صمیمی)، قلب (وزن، کردن، قلب (مرکزی)</td>
<td>Heart, mind, inverse, midst</td>
<td>6.37</td>
<td>2</td>
<td>Significant</td>
</tr>
<tr>
<td>Synonymy</td>
<td>نوک (نشانه)، نوک(پرست)</td>
<td>Top, beak</td>
<td>0.59</td>
<td>2</td>
<td>Not significant (foreign)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morphological problems</th>
<th>Keyword</th>
<th>English equivalent</th>
<th>Bing</th>
<th>Yahoo</th>
<th>Google</th>
<th>The most efficient search engine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing styles of Latin words</td>
<td>انفولوآن، انفولوآن</td>
<td>Influenza</td>
<td>23.02</td>
<td>31.87</td>
<td>31.47</td>
<td>Yahoo</td>
</tr>
<tr>
<td>Loanwords and their equivalent</td>
<td>تیتانیوم، تیتانیم</td>
<td>Titanium</td>
<td>32.12</td>
<td>36.4</td>
<td>33.25</td>
<td>Yahoo</td>
</tr>
<tr>
<td></td>
<td>کامپیوتر، رایانه</td>
<td>Computer</td>
<td>33.53</td>
<td>42.48</td>
<td>44.43</td>
<td>Google</td>
</tr>
<tr>
<td></td>
<td>سیستم - نظام سامانه</td>
<td>System</td>
<td>36.96</td>
<td>47.79</td>
<td>41.03</td>
<td>Yahoo</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morphological problems</th>
<th>Keyword</th>
<th>English equivalent</th>
<th>Bing</th>
<th>Yahoo</th>
<th>Google</th>
<th>The most efficient search engine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derived words</td>
<td>پستچی</td>
<td>Postman</td>
<td>64.32</td>
<td>24.76</td>
<td>34.82</td>
<td>Google</td>
</tr>
<tr>
<td>Compound words</td>
<td>کتابخانه - کتاب خانه</td>
<td>Library</td>
<td>36.35</td>
<td>37.81</td>
<td>41.88</td>
<td>Google</td>
</tr>
<tr>
<td></td>
<td>دانشگاه - دانشسرا</td>
<td>Training college</td>
<td>34.13</td>
<td>34.82</td>
<td>14.28</td>
<td>Yahoo</td>
</tr>
<tr>
<td></td>
<td>فناوری، فناوری اوری</td>
<td>Technology</td>
<td>40.26</td>
<td>43</td>
<td>33.59</td>
<td>Yahoo</td>
</tr>
</tbody>
</table>
As Table X indicates, as for the style of writing letters, each of the search engines is superior in some parts, but none of them have the mechanism to inform users of these various forms in order to improve the search.

According to Table XI, each search engine retrieved various images including different meanings for these words, but mostly they are unable to recognize short vowels, for example in the case of the word “ﺳﻤﮑﻨ” users (who search for the image of painkiller) may completely get confused when the image is retrieved because almost all of the images show home!
The average of words was calculated using Friedman test. The results in Tables VII–XI indicate that in morphology part, Google, Yahoo and Bing have, respectively, the most compatibility with Persian language.

As it is obvious in Table XII, the results show that in semantic part, Google, Bing and Yahoo have, respectively, the most compatibility with Persian language, and retrieve images with different meanings, but they do not have any specific mechanism for informing users or the improvement of results. Overall, findings of Tables VI–XII indicate that with respect to morphological and semantic features, among the selected search engines, Google has the highest degree of compatibility with the aforementioned features and after that Bing and Yahoo rank the same.

5. Conclusion
As the results of this study show, there are some detailed linguistics barriers in relation to image search and retrieval in Persian language. According to the findings which focused on image search and retrieval, there is a significant relationship among styles of writing derived and derived/compound words, Persian and Arabic plural words, use of dotted T and use of colloquial language in morphology part and also polysemy in semantic part and problems associated with image search and retrieval in search engines. For example, by selecting keywords like “شاپوریا” (Technology) and retrieving related images, users would miss images with “شاپوری” tag unconsciously in many cases and therefore considerable number of images would be lost. As a result, in addition to web-paged search and retrieval which mentioned in literature review, the issues mentioned above are considered as the major problems of Persian language in relation to writing characteristics (morphological and semantic) in the selected words associated with image search and retrieval in search engines (Google, Yahoo and Bing). The implications of the present study with respect to the features of the search engines could be generalized to other search engines and also other forms of information. It also extracted and considered most of Persian language features which were mentioned in previous works, and this point makes it more comprehensive in comparison to previous studies. The framework of this study could be applied to other research related to language especially non-Latin languages like Arabic which has major similarities with Farsi language. It helps system developers to improve image search and retrieval engines for non-English users which leads to users’ satisfaction. In addition, the results of this study are useful for policy setting and designing native search engines as it informs developers about weaknesses and strengths of their search engines and retrieval systems. It also identifies a set of criteria that might be useful for storing and retrieving

<table>
<thead>
<tr>
<th>Semantic problems</th>
<th>Keywords</th>
<th>English equivalent</th>
<th>Bing</th>
<th>Yahoo</th>
<th>Google</th>
<th>The most efficient search engine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homonymous</td>
<td>شیشه‌خوار: شیر (اسپاری: شیرا)</td>
<td>Milk, lion, tap</td>
<td>44.36</td>
<td>45.18</td>
<td>46.12</td>
<td>Google</td>
</tr>
<tr>
<td>Similar words with different meanings</td>
<td>شور (طعم: شور (اشیاء)</td>
<td>Passion, salty</td>
<td>16.81</td>
<td>17.4</td>
<td>19</td>
<td>Google</td>
</tr>
<tr>
<td>Multiple meaning</td>
<td>قلب (عضو بدن): قلب (خاطر و معنای) قلب (واژه‌های گردب: قلب (مرکز)</td>
<td>Heart, mind, inverse, midst</td>
<td>42.37</td>
<td>42.37</td>
<td>38.12</td>
<td>Yahoo and Bing</td>
</tr>
<tr>
<td>Synonymy</td>
<td>دریای خزر: دریای کاسپین: دریای مازندران</td>
<td>Top, beak</td>
<td>6.8</td>
<td>6.74</td>
<td>9.27</td>
<td>Google</td>
</tr>
</tbody>
</table>

Table XII. Results of Friedman test related to average of search engines in retrieving different forms of selected keywords in semantic part
images and determines some categories in morphological and semantic features of Persian language which are more frequent.

Overall, the findings of this study are consistent with the results of Abdolahi-Noorali and Jowkar (2009), who concluded that none of the search engines have attended to the challenges of Persian writing styles in order to improve search results. Moreover, an examination of the results shows their compatibility with the findings of Lazarinis (2007a), which suggested that search engines retrieved different results for different morphological features of words. According to the results, Arabic words in Persian language like various kinds of plural or writing styles of dotted T constitute a significant portion of these problems, and this matter requires some short- and long-term strategies. After Arabic words, writing continuous or separate forms of derived and derived/compound words is taken into account. In the semantic part, polysemy is a major problem like what users faced in searching "قلب". That is to say, in many cases users retrieve only images that show one meaning of this word and, as a result, they become confused. Solving these problems requires providing and implementing strategies for both search engines and efforts in order to integrate writing styles on the web. A survey of results in morphological and semantic parts indicates that Google has more advantages than Yahoo and Bing in many ways. The findings of this research are also consistent with the results of Zhang and Lin (2007) and Tawileh and et al. (2010) which suggested that Google is the best engine with multilingual support among the search engines that they had considered.

Overall, in the present study and other research studies such as Namk and Bitirim (2015), ALSoh and et al. (2010) and Han and Gutman (2005), the search engines under investigation do not pay adequate attention to morphological and semantic features of non-English languages like Persian. Due to the increasing number of internet users in Iran, Persian speakers need a native search engine which is designed based on their linguistic features. Moreover, Google, Bing and Yahoo, which are popular search engines, must try more to meet the needs of non-English users. As the results of this study indicate, ignoring or neglecting the characteristics of the users’ language can cause problems in search and information retrieval, and finally they might miss the relevant information or retrieve undesirable information.

The present study has focused on the weaknesses and strengths of major search engines in relation to Persian language, and the results can be used to improve their function and to design national search engines. Also, the findings of the present study can be employed in order to improve storing and retrieving image process in Persian language. Moreover, a wide range of Persian-speaking users who deal with information search and retrieval will benefit from the findings as they consider the morphological and semantic points of language and refine their searches. Thus, according to the findings, the following suggestions are offered:

- According to the results, it is suggested that Persian speakers consider the linguistic features in their information and especially image search for a successful retrieving.
- The results suggest that informing users of variations in writing, with emphasis on the major problems identified in this study, can help them achieve a more successful retrieval.
- According to the findings of this study, many of the morphological problems are related to different styles of writing a single concept; therefore, creating a mechanism for search engines in order to inform users during search, through showing synonyms, Persian equivalent of loanwords and also different spellings of derived and derived/compound words can help prevent user confusion.
- Search engines can help users with specific words which have several meanings by showing these meanings as users choose from among them.
In order to solve the identified semantic problems in this study, especially polysemy, showing a list of images including different meanings of keywords, and the simultaneous writing of keywords in a search box by users can help them select their desired meaning. For example, when users type the word "کر" in the image search box of a search engine, the list of images containing the different meanings of this word appears.

Notes
1. TASI (Technical Advisory Service for Images).
2. We used www.alexa.com/ranking.
3. Our mean of Persian language is the language which people write in Iran.
4. ISNA (Iranian Students’ News Agency).
5. Homonymous refers two words that either sound the same when spoken, have the same spelling or both and have different meanings and origins (Oxford, 2017a).
6. Polysemy is the coexistence of many possible meanings for a word or phrase (Oxford, 2017b). Polysemeous refers one word that can be used to express different meanings and the difference can be obvious or subtle.
7. Iranian Research Institute For Information Science and Technology.
8. Regional Information Center For Science and Technology.
9. In order to show major problems related to Persian language with search engines, our checklist was designed on the basis of past studies’ results such as Abdolahi-Noorali and Jowkar (2009), Akhshik and Fattahi (2012) and expert’s ideas regardless the kinds of problems and concepts at first, then we customized it for image search at the second stage.
10. Due to the impossibility of reviewing all the findings, and also considering the related studies which in many cases were conducted on 10 or 20 retrieved first result (Lazarinis, 2008, Namk and Bitirim, 2015), and also because of a more comprehensive study, eventually 50 retrieved first results for each keyword in any search engine were taken into consideration.
11. We did not judge relevance of results in order to determine which search engine has the best function in related Persian language.

References


Further reading


Corresponding author

Hoda Homavandi can be contacted at: H.Homavandi@gmail.com
Google as a political subject: the right to be forgotten debate 2014–2016

Linnéa Lindsköld
Department of Arts and Cultural Sciences, Lund University, Lund, Sweden

Abstract

Purpose – The purpose of this paper is to create knowledge on how Google and Google search are discursively constructed as a political subject suitable or not suitable for governing in the debate regarding the Right to be Forgotten ruling (RTBF).

Design/methodology/approach – A total of 28 texts are analysed using a Foucauldian discourse analysis focussing on political problematisations in the media and in blogs.

Findings – Google is conceptualised as a commercial company, a neutral facilitator of the world and as a judge of character. The discourse makes visible Google’s power over knowledge production. The individual being searched is constructed as a political object that is either guilty or innocent, invoking morality as a part of the policy. The ruling is framed as giving individuals power over companies, but the power still lies within Google’s technical framework.

Originality/value – The ruling opens up an empirical possibility to critically examine Google. The value of the study is the combination of focus on Google as a political subject and the individual being searched to understand how Google is constructed in the discourse.

Keywords EU, Discourse analysis, Information policy, Google search, Right to be forgotten

Paper type Research paper

1. Introduction

Google is an important subject in information politics, influencing how knowledge is constructed and distributed, thus being political (Introna and Nissenbaum, 2000; Vaidhyanathan, 2012). Google search is not giving an objective reflection of the world, rather it is producing search results through, among other things, tools like PageRank and personalised search (Hillis et al., 2013; Vaidhyanathan, 2012). In this context, the term “political” refers to the impact Google has on power relations in society, both on account of being a multinational company and the dominant search engine, influencing how knowledge is constructed and distributed (Vaidhyanathan, 2012). As a result of the latter, search engines have a profound effect on subject formation and identity when individuals are being searched (Lyon, 2001; Vaidhyanathan, 2012; van Zoonen, 2013). Researchers have problematized Google’s dominant role on the market during the last decade. The question of whether and how Google should be governed by nations or supranations has also been part of this discussion (Dwyer, 2016; Halavais, 2009; Hillis et al., 2013; Vaidhyanathan, 2012).

The latest addition to the governing of Google by the EU is the Right to be Forgotten ruling (RTBF) instigated by the European Court of Justice (CJEU). Since May 2014, EU citizens have been able to apply for search engine companies to remove search results of their name if the information is deemed to be irrelevant or in other ways inaccurate. This new policy opens up an empirical possibility for investigating different conceptualisations of Google in the media and how these conceptualisations make possible different actions towards the company. The aim of this paper is thus to create knowledge on how Google and Google search are discursively constructed in the debate as a political subject suitable or not suitable for governing.

This study was conducted within the project “Knowledge in a Digital World. Trust, Credibility and Relevance on the Web”, funded by a framework grant from the Swedish Research Council 2013-2016.
To accomplish this, I will identify and analyse the political problematisations, in media and blogs, of the “Right to be Forgotten” ruling, which emerged during the period 2014–2016. I will also discuss how these problematisations challenge or add to common conceptualisations of the company and its search engine.

The results will build towards a greater understanding of how large internet companies such as Google can be subject to governing.

This study concerns technologies that most of us use everyday and which are incorporated in our way of life. Few of us understand the technology or the policies behind such tools as Google search, which limits the way users can make informed decisions (Proferes, 2016). News articles and blog posts are important media for mediating and explaining information policy issues that affect the public. However, such reporting builds upon existing discourse and may reinforce power structures. Therefore, it is relevant to examine the framing of “problems” that a certain policy is set to resolve.

Focussing on the debates in media and on blogs on a policy proposal that affects Google is one way of rendering Google visible (Haider, 2014). The study focusses only on visible search results, not commodification of personal information, i.e., implicit or explicit harvesting of personal information that is being sold to advertisers.

First, the study is framed through a review of relevant literature on the politics of search engines and searching (Section 2), followed by a presentation of my methodological framework (Section 3). Using a Foucauldian discourse analysis, it is possible to analyse the different problematisations of the policy and how Google is conceptualised in these (Bacchi, 2009). Then, the results are presented, focussing on how Google the company and Google the search tool are politicised in the discourse, as well as how Google is conceptualised as a judge of moral character (Section 4). Thereafter, the results are discussed in relation to earlier research with regard to the framing of the ruling as a way of giving power to individuals over companies. This notion is challenged (Section 5) and finally, conclusions are drawn.

2. The politics of Google and the effects of search

Before describing the RTBF ruling in more detail, relevant research on the politics of Google and on the effects of search are discussed.

2.1 Information politics

From a social-constructionist perspective, search engines are never neutral, objective tools for information retrieval. Introna and Nissenbaum (2000) had already discussed that how biases in search engines might limit the web and its use in society, something they identify as a political problem. Halavais writes how search engines become invisible in our everyday life; it is something we use without reflecting on the limitations of these tools. He argues, as well as Vaidhyanathan (2012) and Hillis et al. (2013) that search engines, with Google as the major example, forms our understandings of knowledge. Halavais (2009, p. 1) writes that our culture has a common assumption of search engines: “[…] that a search engine will lead someone to a page that contains accurate information, and that questions are best directed first to a machine and only after that to other people”. Through biases in their infrastructure, search engines contribute to forming our understanding of identity, knowledge and the world. For example, a bias can be how search results are presented according to relevance, which affects what we perceive as important. In other words, search engines are political and the commercial companies behind them have great influence over how we perceive the world (Introna and Nissenbaum, 2000; Halavais, 2009; Vaidhyanathan, 2012).

Monica Horten (2016) writes that there are two different perspectives or narratives in an information policy discourse, highlighting the conflicts between various actors such as...
internet companies and governments. The first narrative is the “market-led perspective of policymaking”, where governments aim to construct policies that maintain a good relationship with the large internet corporations that use lobby organisations to put forward their views. The other is the user-empowerment narrative that is working for an open internet, lobbying for free speech (Horten, 2016). However, the user-empowerment narrative is complicated when taking account that users have difficulties in understanding the technologies and business models of platforms (Proferes, 2016). It is also problematic for a researcher to capture all the complexities of global social media platforms since Google is both a major political player and a technology that changes how we perceive the world and ourselves (Vaidhyanathan, 2012).

2.1.1 Google as a political subject. Google has both been criticised and praised for its dominant role. The company has been criticised for being too pliable to demands from the Chinese Government to filter search results, for distributing users’ search records to the US Government, for distributing content without the authors’ approval and for contributing to bias, where Anglo-Saxon culture is favoured through the PageRank system (Diaz, 2008; Halavais, 2009; Hillis et al., 2013; Vaidhyanathan, 2012). Vaidhyanathan (2012, p. 110) argues that rather than the content, it is Google’s distribution of information and the terms of access and use, the default settings, that is the dominant form of cultural imperialism.

News media play an important role in influencing public opinion as well as the policy agenda (Shahin, 2016). Google has been praised in the media and Shaker argues that trust in Google may be fostered through media reporting on the company’s fiscal success. Hoofnagle shows how the privacy rhetoric of Google representatives is vague in news articles. These findings suggest a discourse where Google is identified as a successful company in the press, thereby not relevant to criticise, along with a tendency to not discuss privacy issues in depth (Shaker, 2006; Hoofnagle, 2009). The RTBF may provide a discursive intervention in that it makes issues such as personal data and privacy “newsworthy” possible to discuss in the media, on a scale that has not been done before.

One of the key issues that affect how Google is discursively constructed is national regulatory frameworks and the way search engines can affect people’s lives. Historically, Europe and the USA have had different approaches to privacy legislation, where the EU approach is pre-emptive regulation and the US model is self-regulation (Zimmer, 2013). This makes for an antagonistic situation when European institutions make legislation that affects internet companies from the USA. The EU’s new general data protection legislation, in which the RTBF is included, is a relatively new attempt to regulate the internet on a European scale. However, Google has complied with the laws of other nations before, such as the national hate speech laws in Germany and France (Halavais, 2009; Hillis et al., 2013).

2.1.2 Implications of search on the individual. Another key issue that becomes visible when discussing Google search is the effect on the individual who is being searched and who might use the RTBF or, put another way, how Google’s role is understood in relation to the individuals being searched.

To understand the discourse on the RTBF ruling, it is important to consider the reasoning behind the policy. This study focusses on a limited kind of search, namely, searching a person’s name. Our digital identities are, according to Dwyer (2016, p. 78), “perhaps the most valuable assets we own”. This means that the search results for our names can have great effects on our lives. As personal information is stored for an indefinite future, it is almost impossible to allow the errors of youth to fade away (Halavais, 2009). Vaidhyanathan suggests that the notion of privacy in relation to Google search stands for a desire to have some control over our reputation online.

Privacy should not be seen as the nature of information shared, but the “terms of control over information” (Vaidhyanathan, 2012, p. 93). On a more abstract level, when
someone searches your name, he or she will retrieve search results that together build an understanding of who you are, what you have done and perhaps even what will become of you in the future. Many different search results from a whole lifetime, may be even regarding different persons, are melted into one body. For example, a search result also constructs identities through personalised searching, in ways that the person being searched for is unaware of since the results depend on the person doing the search (i.e. his/her previous searches). Google search can thus be seen as an unintentional tool for disciplining individuals.

Sociology professor David Lyon has written on the ethical aspects of modern technology and surveillance. His solution is an ethical stance for the researcher, who should always consider the material effects for individuals when discussing issues such as privacy and surveillance on the web. Discriminating structures for different groups exist both on and offline and may even be reinforced online (Lyon, 2001).

Google search opens up for horizontal surveillance, where peers are watching peers, and because this can have an impact on both social and work relations on an individual level and on a collective level, it can change the online behaviour of groups (Bennett and Parsons, 2013). Horizontal surveillance can also lead to public shaming of people who are (perceived as being) guilty of moral wrongs (Ronson, 2015). In the discussion on who is watching whom, with technology that is capable of total remembrance, another alternative is to be forgotten, which the RTBF ruling is aiming to achieve. Other suggestions have also been put forward, such as setting expiration dates on information stored in digital memory (Mayer-Schönberger, 2009).

Even though ethical aspects are not at the centre for this study, it is a relevant background for understanding the discussion of what is at stake, when concerning the right to be forgotten, namely people’s right to privacy in relation to freedom of expression.

2.2 The right to be forgotten

It was in the C-131/12 ruling that the RTBF was recorded in a legally binding document for the first time. Before being incorporated into policymaking, it has been a concept coined and discussed by academics, as Schiedermair (2015) shows in her informative overview of the law.

The CJEU instigated the RTBF ruling in May 2014 (C-131/12). It concerns the right of the individual in the EU to request removal of personal data from access via search engines; personal data is defined as “any information relating to an identified or identifiable natural person” (Regulation 2016/679, Council of the European Union, European parliament, 2016). Unless otherwise stated, the Right to be Forgotten and the abbreviation RTBF in this article refer to the C-131/12 ruling and not to the general principle of the right to be forgotten on the internet. The RTBF ruling means that search results can be deleted if they are deemed to be inaccurate but also “inadequate, irrelevant or excessive in relation to the purposes of the processing, that they are not kept up to date, or that they are kept for longer than is necessary unless they are required to be kept for historical, statistical or scientific purposes” (C-131/12, § 92). Any citizen of the EU can request that a search result for their name be removed, including personal websites and newspaper articles. It should be noted that no information is removed or erased through this policy, only it cannot be found through search engines when searching one’s name. However, the same source, for example, a news article, can still be found if other search terms are used. If a company does not follow the ruling, the EU will fine them. The ruling has been criticised for being ineffective and for destroying the internet, but it has also been praised for granting individuals their right to privacy.

Google has explicitly stated that the company is against the ruling since it is seen as being a threat to the company as well as the freedom of the internet, but it has still enforced
the ruling. A form is available through Google with which individuals can apply for the removal of search results; Google employees then assess the case based on the criteria made in the ruling. It is also possible to appeal against the decision. Public figures are exempted from the RTBF but the definition of a public figure according to Google is not transparent (Schiedermair, 2015).

At first, only the search results on European sites were removed; they still existed on Google’s non-European sites. But since early 2016, Google has started to block results from all of its domains if searches are initiated in Europe (Gibbs, 2016). Many lines of conflict are emerging in the ongoing debate on the RTBF, such as freedom of expression and freedom of the press vs the right to privacy, the differences in information policies between the EU and the USA, and the power of private companies such as Google vs national and international law. This is also visible in the media. In a study on media coverage of the RTBF in the USA and the UK, the ruling was discussed in five broad terms: USA vs EU, censorship, free market vs regulators, cause and practicality, meaning how the ruling should be implemented (Shahin, 2016). Shahin’s broad overview of the debate will be supplemented by my analysis, limiting and deepening the scope to Google as a political subject.

3. Problematisations as an analytical tool

A theoretical assumption for this study is that the RTBF ruling and the ensuing debate concern the limitations and possibilities for Google as a political subject to be governed.

As earlier research suggests, Google must be understood as both a global company in a political-economical context and as a technology that affects people’s identities. To understand Google theoretically through the RTBF ruling debate, I have used Foucault’s concept of problematisations, i.e., investigating how an issue becomes a problem that needs to be solved (or not) through a specific policy. By identifying and analysing problematisations, the researcher can make visible the conditions that make certain problematisations possible (Foucault, 1992). Problematisations are part of a discourse, for this study discourse refers to “socially produced forms of knowledge that sets limits upon what it is possible to think, write or speak about” (Bacchi, 2009, p. 35).

Political scientist Carol Bacchi (2009, 2012) has developed a methodology based on Foucault’s notion of problematisations, focussing on how governing (ruling) takes place. In a policy analysis, this means that a policy is not a solution to a given problem; instead, it is a way of representing a problem that can or cannot be solved by the policy. This is a broad approach to policy analysis and suitable for this study where the political aspects of large companies such as Google are seen as being important in order to understand power relations in society. In using this methodology, the aim for the researcher is not to assess if and how the RTBF is a correct solution to a problem, but to identify how and why the discourse makes different conceptualisations of Google possible.

Following Bacchi (2009), I am interested in how governing takes place by way of problematisations through two themes. The first is the assumptions underlying the problem representations in the discourse, which are not only those of the RTBF ruling, but also include the representation of the problem from the position that is against the ruling. The second theme concerns the effects that are produced by the problem representations in the discourse, focussing particularly on the ways in which subjects are constituted in the discourse (Bacchi, 2009). The intentions behind statements made by authors are irrelevant; it is the statements themselves that are studied as parts of a discourse. To concretise, in the debates following the RTBF ruling, statements were made from different subject positions made available through the discourse (Google employees, pro-privacy academics, etc.) which, in debate or opinion form, recommend how Google, individuals and the EU should act.
The analysis was done through close-reading of the material (presented in the next section) and identifying different representations of the problem that the RTBF will or will not solve in the debate. These problem representations structure the results section. I have also identified nodes, which means central concepts that are ascribed certain definitions or values by different subject positions. Global enforcement is one example of a node but the most evident one is Google that is conceptualised in different ways through different problem representations. Another task was to identify positions that were made possible in the discourse for different subjects. Two examples of subject positions that made statements in the discourse are Google employees and academics; two examples of positions that statements were made about are innocent and guilty bodies. Through the analysis, it will be possible to discern presuppositions and assumptions that are taken for granted in the discourse, and statements that are contested by different subject positions. The results will identify and analyse the struggle between different subject positions regarding how Google should or should not be governed.

3.1 Material

The empirical material consists of articles and opinion articles published in British and American newspapers, press releases from politicians in the European Parliament, and blog posts by academics, lobbyists and Google spokespeople published 2014–2016. A total of 15 articles have been retrieved, ten had been published in newspapers in the USA and five in the UK. Eight blog posts have been retrieved and four other text categories, a parliamentary report, press releases and the Google advisory council’s report have been included. The majority of the authors or people interviewed were academics affiliated with Google, different think tanks or NGOs, or universities. Sometimes the same person falls into all three categories. Politicians and journalists are also featured in the material.

The material consists of different genres published in different contexts and with different genre rules. However, there is genre overlap since many of them make reference to each other: a blog post can also be an opinion piece, a blog can be quoted in a news article and vice versa. One benefit of the variety of genres is that it becomes possible to follow how, for example, Google employees express themselves in different venues for different audiences. It can be assumed that Google employees would want to speak with one voice about the company and the search engine (compare Hoffmann et al., 2016, p. 3).

The aim is not to give a full picture of the debate but to discursively analyse a section of it, which focuses on the RTBF in a given period (2014–2016) and place (the USA, the UK and to some extent the EU). The material has been retrieved through keyword searching on Summon, Google and on websites such as that of the European Parliament, different party groups in the parliament and newspaper websites. The irony is not lost on the author that Google is needed as a tool for doing research on Google. In fact, it would be impossible not to use Google since many website search tools are provided by the company, as is the case with the British newspaper The Guardian. The keywords have been: right to be forgotten, right to be delisted and Google Spain judgment. Material has also been retrieved through references made in the texts. Only opinion pieces and longer articles with interviews have been included and therefore shorter news articles or news items only referring to the ruling without interviews or comments have been excluded. The blog posts selected for analysis were authored by either Google employees or academics specialised in the fields of information, law and privacy. Material has also been retrieved through references made in the texts. Only opinion pieces and longer articles with interviews have been included and therefore shorter news articles or news items only referring to the ruling without interviews or comments have been excluded. The blog posts selected for analysis were authored by either Google employees or academics specialised in the fields of information, law and privacy. Blog posts were included in the material so as not to exclude a large arena where the discussion on the ruling took place. For example, Google used its own blogs to make statements and the London School of Economics invited scholars to write guest posts on the issue (Fleischer, 2012; Goodman, 2015).

The texts chosen for analysis can all be described as “practical texts”, i.e., texts that instruct or argue for how, in this case, Google and/or the individuals affected by the ruling,
should act (Foucault, 1992). Some biases should be noted: first and foremost, only texts in English have been analysed but the RTBF has also been hotly debated in France and Germany. Germany is an interesting case with a history of being sceptical towards Google (Vaidhyanathan, 2012, p. 102). Another bias is that the majority of the authors are scholars, which frames the discussion in an academic setting. However, this is a result of the ruling being developed from an academic concept and discussed as such in the media. In order for other positions to be represented, interviews could have also been used as material.

4. Results

Four different problem representations have been identified in the material: In the first, the RTBF is understood as a solution that gives Google as a private company too much power. In the second representation, the RTBF is understood as an unnecessary solution to a problem not relevant to Google search, which is conceptualised as a neutral tool. Global enforcement is the third representation of the problem, where the question of whether laws of the EU should be applicable for American companies are discussed in relation to the politics of internet. In the final representation, the imagined users of the RTBF are central, and Google is conceptualised as a judge of character in the discourse. The results section discusses how Google is constructed as a political subject in the discourse through these representations. In the discussion, I will elaborate on what these representations mean for the conceptualisation of Google.

4.1 Google as a private company

In the discourse, Google is described as a private company and therefore not responsible for enacting the RTBF ruling; instead, governmental institutions should have that duty. The fact that Google is a company should come as no surprise to the reader, but the “Google as a company” rhetoric highlights a certain problematisation which does not emerge in other discourses on Google. In the material, both critics of the RTBF ruling and its proponents see a problem in that Google the company assesses removal requests and not a public institution. This is because a private company cannot and will not have the same kind of transparency as a government institution. Luciano Floridi, who was part of the Google advisory council, says in an interview in the New York Times: “If Europe really wanted to regain control over personal data, giving Google this type of power is an odd outcome” (Scott, 2016). A prominent subject position in the discourse consists of academics, active at universities, who work with questions on internet and privacy. This position also criticises Google for lack of transparency and demands that Google make the process more visible, both for the sake of the public and to be able to use the data for research (Kiss, 2015). Thus, Google is a suitable object for governing both by policymakers and by researchers. Google Advisory Council (2015) suggests in its report that RTBF is a misleading concept and instead suggests the right to be delisted, which is then used by Google. The renaming of the “right to be forgotten” to the “right to be delisted” can be seen as a way to distinguish the phrase from the academic concept of RTBF, and a way to downplay the impact of the action since delisted does not have the same connotation as forgotten.

Earlier research has shown how Google is constructed by themselves and by users as being neutral, objective and understood as the act of searching for information in itself. An example of this, and of Google’s traits, which made it stand out from other search engines when it first entered the market, was the empty search box, a web page with no clutter. According to Hillis et al. (2013, p. 14) Google represents a kind of “magical empiricism”: “offering legitimisation through a matter-of-fact calculation […] which in turn draws on Enlightenment ideals of empiricism and its connection to ideals of progress”. But when the Google as a company-problematisation is used in the discourse, the neutral aspects of Google search become muddled.
The different responsibilities of commercial companies and governments are put forward as an argument both for and against the ruling (Goodman, 2015; Zittrain, 2014). Jimmy Wales, founder of Wikipedia and a member of the Google advisory council argues against the ruling in the council’s report: “I completely oppose the legal situation in which a commercial company is forced to become the judge of our most fundamental rights of expression and privacy, without allowing any appropriate procedure for appeal by publishers whose works are being suppressed” (Google Advisory Council, 2015, p. 27). Technology researchers Julia Powles & Enrique Chaparro also questions the role of Google, but argues for the ruling in The Guardian: “Google has acted as judge, jury and executioner in the wake of Europe’s RTBF. But what does society lose when a private corporation rules public information?” (Powles and Chaparro, 2015). In contrast, the EU frames the ruling as a user vs company conflict. Hannes Swoboda, President of the Socialists and Democrats Group in the European Parliament, said in a press release: “This court decision clearly demonstrates the spirit of European data handling: power to the user, not the company” (S&D, 2014). But Jonathan Zittrain, academic and critic of the RTBF ruling points out: “[…] here state power is being exercised without the involvement of the state” (Zittrain, 2014). Private vs public and company vs user are two dominant dichotomies in the discourse.

In this problem representation, the RTBF is not contested as such, rather it is the outcome of the ruling: that a private company has responsibility for enacting what is being criticised. Here, both Google employees and academics agree[1].

4.2 Google search as a neutral tool
In contrast to the private company Google, the search engine Google is described as being a neutral tool by employees and some scholars and journalists in the debate. On the Google Europe blog, Peter Fleischer, Google’s global privacy counsel, makes a distinction between “services that host content created by people”, exemplified by Facebook and YouTube, and “services that point people to content that exists elsewhere”, namely search engines (Fleischer, 2012). However, the RTBF ruling pushes Google search into a more visibly active role as gatekeeper. Larry Page, Google’s Chief Executive, commented on the result of the RTBF ruling in an interview with Farhad Manjoo:

To date, we’ve said we’ll try our best to represent the things that are out there on the Internet about you. It’s worked for 15-plus years. It makes a lot of sense. We’re a search engine. It seems like we should represent what’s in the world. So it was a pretty surprising ruling — it’s a different statement. You guys are now in charge of editing what’s out there in the world. In the past that’s not a responsibility we felt we had. (Manjoo, 2014)

In the quote, the responsibility of “editing” is described as something new, although Google has removed information such as social security numbers or copyright material for a long time. But even without taking into account these adjustments of search results, the technology of search engines themselves cannot provide “fair and unbiased results” (Lewandowski, 2017). Lewandowski (2017) shows how Google downplay the power the search engine has over search results, thus ignoring its monopoly status. As a result of the ruling, the editing role of search engines has become highlighted and debated in the media through the problematisation of the RTBF ruling in ways that have not existed before the ruling. Furthermore, there has been a shift in the problematisation of how threats against privacy are perceived since the ruling states that search engine providers, as well as website owners, become responsible for privacy violations (Schiedermair, 2015). In other words, it is no longer enough to claim that a search engine only represents what is in the world, but that the company also has responsibility for the results. This is not a preferred position for Google. Kent Walker, general counsel of Google, said in an interview: “We don’t create the information. We make it accessible. A decision like this, which makes us decide what
goes inside the card catalogue, forces us into a role we don’t want” (Toobin, 2014). When comparing Google to a neutral tool or a card catalogue, the ideological aspects of Google, and how the company affects knowledge production, become invisible. In this problem representation, it is the RTBF itself that is rejected and the focus lies on the function of Google search rather than the company.

4.3 Global enforcement

The third form of problematisation through which the problem of the RTBF is considered is whether the ruling should be applicable in countries other than the EU’s member states. Global enforcement is a node in the discourse, putting forward a trans-Atlantic conflict in the debate (Shahin, 2016). This also ties into the user-empowerment narrative discussed by Horten (2016). The two following quotes from a journalist at the Business Insider and the legal director of the Media Legal Defence Initiative are examples of an understanding of internet and Google search as being a free flow of information:

Understand that Google didn’t write or commission these posts. Google had nothing to do with them. All Google does is index them, along with the billions and billions of other pages on the web, so that people can find them when they search. (Rosoff, 2015)

Today the ECJ has identified the very function of the search engine as a serious threat to personal privacy and data protection [...] The real loser in this case: the public and the interest in society at large in the free flow of information and ideas. (Noorlander, 2014)

In other understandings of internet in the discourse, the internet was never free and is already under the control of different policies and corporations: “internet companies have been successful in making us believe that the internet is “public space”, when, in reality, it is just an algebraic representation of privately owned services” (Powles and Chaparro, 2015). Linked to the understanding of the web as a free flow of information is a fear that a local law with global reach will alter the ways in which the internet is used. Jennifer Granick, Stanford Law School, says in an interview that the RTBF is the beginning of the end of the global internet: “[...] and the beginning of an internet where there are national networks, where decisions by governments dictate which information people get access to” (Toobin, 2014).

Law professor Ellen Goodman writes:

The fear of EU over-reaching is that an aggrieved businessman in Milan can control what an interested citizen or journalist sees in New York. More than that, it’s the fear the EU over-reaching will spread to Pakistan and Russia, and that these countries’ strategies for suppressing information will be exported to the global Internet. (Goodman, 2015)

However, Powles (2015) argues that global enforcement of privacy laws is not a problem since copyright laws are already in place. The different views on the right to be forgotten, in the USA and in Europe, have been extensively commented upon, both in the debate and in research. The conflict between Europe and the USA can to some extent be explained by the countries’ different bodies of law, where the first amendment in the USA protects free speech, while the European Convention on Human Rights in Article 8 concerns the right to respect for privacy (Dwyer, 2016). In the US media, the ruling is understood as something unthinkable in their own national context: “The internet’s unregulated idyll seems to be coming to an end, at least in Europe” (Toobin, 2014). A note on this conflict is that it is usually described as one between the USA and Europe, ignoring the fact that views between European countries and the EU also differ. In the UK, the House of Lords argues in a report on the EU Data Protection Law that the RTBF is “misguided in principle and unworkable in practice” (House of Lords, 2014). Earlier attempts to govern internet privacy across borders have not been as successful, even though there is no evident “race to the bottom” (Zimmer, 2013).
However, the ongoing legislation work in the EU suggests that internet privacy for individuals is being put forward as a prioritised political issue.

The future of the internet and the politics of internet on a more abstract level are in focus in this problem representation. It is also a result of different national legislation in different countries.

4.4 Google as a judge of character

Search engines have political implications for how we perceive others and ourselves. Currently, there is a movement towards a fixation of single identities on the internet (van Zoonen, 2013). On the other hand, there is a need to acknowledge how people are affected by information politics. In the discourse, journalists, academics and Google employees build on real and imagined bodies that are affected by the ruling. There are many reasons for persons wanting to be forgotten by using the new ruling, but in the material, two reasons are identified: being guilty or innocent/a victim. Through this problem representation, Google is conceptualised as being a judge of character in the discourse.

After the ruling, a fear is expressed in the discourse that criminals will take the opportunity to hide their pasts: “convicted criminals are among those trying to hide links to stories from online search engines” (Harper and Owen, 2014). The Washington Post published a news story the day after the ruling with the heading: “Does a British paedophile deserve the ‘right to be forgotten’ by Google?” (McCoy, 2014). This framing of the RTBF ruling plays on a fear of not knowing the criminal pasts of people. David Drummond, Senior Vice President for Google, writes in an opinion piece that was also published on Google’s own blog that:

The possibilities for resistance are implicated in the quote, where the “elites” or criminals should be watchable by others. There is an underlying conflict over whether a criminal offence should always be part of a person’s identity or if it should be “forgotten”. Google search can be seen as an unintentional tool for disciplining individuals where the investigation and interrogation of an individual are always ongoing. But there are also academics in the debate, such as Eric Posner, who argue against it (Posner, 2014). The human mind and human societies can forget but the internet cannot and this is an ethic problem representation at the core of the RTBF ruling.

Bodies other than the criminal are used in the discourse to defend the RTBF ruling. Powles and Chaparro (2015) discuss search results: “They include prominent reminders that an individual was the victim of rape, assault or other criminal acts; that they were once an incidental witness to tragedy; that those close to them – a partner or a child – were murdered”. Other bodies that are mentioned as being in need of the ruling are people wrongly accused of crimes, poor single mothers, widows and people with HIV (Hakim, 2014; Tippmann and Powles, 2015; Powles, 2015). In particular, photos of dead children at the scene of an accident are mentioned (Toobin, 2014). Some examples are moral, for instance, a mother who wants to remove photos of her undressed teenage daughter or an ex-wife who wants to remove records of divorce (Hakim, 2014). While some of the personal information used as examples in the debate had been published by the individual, most had not. Most of the examples are public information that becomes more widely distributed than intended by the search engine society. Zimmer and Hoffman (2011, p. 177) call this “a leakage of personal data outside a particular informational context”. The subject position is constructed as being “ordinary people”, who suffer from horizontal surveillance. Google’s role as a judge of character through its personalised search is taken for granted and not questioned in the discourse.
The results show that there are conflicting problem representations in the discourse that affects the conceptualisation of Google and the possibilities for governing the company. This will be further discussed in this section, including the implications of the ruling for Google as a political subject.

4.5.1 How does the search engine become political in this particular discourse? In the discourse, there is a presupposition that Google should not be responsible for making the kind of assessments that the RTBF ruling entails, by both Google employees and critics of the company. At the same time, there is an assumption by some that Google search is a neutral tool and that the company is a defender of the free flow of information, so there is a conflict in the discourse between conceptualisations of Google the company and the search engine tool. This assumption can also be seen by some journalists and lobby organisations in the debate, where it is taken for granted that Google search represents a section of reality: “All Google do is index them” (Rosoff, 2015). However, as many researchers have argued, Google search is political through design since it produces search results, not mainly reflecting the world as it is (Hillis et al., 2013; Vaidhyanathan, 2012).

Horton’s (2016) two narratives in information policy discourse, the “market-led perspective of policymaking” and the “user-empowerment narrative”, are not enough to understand the conflict in this particular discourse. The different narratives or possible discursive formations are more complex since subject positions defending Google and the interests of internet corporations against the ruling use free speech for internet users and the possibility for resistance as an argument for their cause. The other position, defending the ruling, also speaks from a user perspective, evoking privacy as a right for the user.

Google has been accused of cultural imperialism and a bias towards Anglo-Saxon search results (Jeanneney, 2007). But following Vaidhyanathan (2012), the problem representation of the RTBF ruling is a way for the individual to take power over companies, and it constitutes a victory for the EU information policy. But since the assessment procedure is not transparent and it is carried out by Google, the power still lies within the technological framework.

The company has been successful in protecting and promoting its brand. Hillis et al. (2013, pp. 199-203) conclude their investigation of the culture of Google search by reflecting on the ideal of Google to be an all-encompassing “thinking machine”, giving answers that machines currently cannot provide, such as predicting future events. This conceptualisation of Google as an all-encompassing vision is challenged by the RTBF ruling through exposing the possibilities for editing and adjustments of search results. But Google can still uphold its “magical empiricism” by concealing its judgements and related data from researchers and the public.

The search results affected by the ruling are politicised in the discourse and have subjectification effects for the individual. With this problem representation, Google is constructed as a judge of character in the discourse by, producing search results of individuals, for example, evidence of criminal activities that needs to be public to keep us safe, meaning that those who want to have a search result forgotten are guilty, or search results that are described as being a result of grief or shame for an innocent person which that person wants to hide. As has already been mentioned, horizontal surveillance aspects of search are an assumption in the discussion on criminal bodies.

In texts defending the ruling, there are a lot of stories about material effects for individuals that have been bullied or in other ways exposed to negativity through search results. All the examples are anecdotic and do not take into account larger power structures in society. For example, women are more vulnerable to strategic attacks on the web, not least through “revenge porn”, which is now illegal in many countries.
Search engines make horizontal surveillance possible without social consequences for the searcher while the consequences for the individual whose personal information is made available online may be "[...] the inability to work, the persistent psychological fear of stalking, the dissemination of intimate facts that lead to harassment [...]" (Bennett and Parsons, 2013, p. 496). The fact that other laws cover discrimination against groups may be one reason why this is not mentioned in the discourse on the RTBF. The examples also focus on individuals and not groups.

An underlying but not explicitly stated assumption is that people who can benefit from the ruling are socially vulnerable due to poverty. An implication of the findings for online information is that the ruling can be seen as a way of giving power to individuals. But, as will be discussed further in the next section, the power still lies within the domains of Google. The ruling can also be used as an act of resistance towards the development of fixating single identities on the web (van Zoonen, 2013).

However, this is not done by challenging the presupposition that your search results are the real you or that Google search can be used as a judge of character. There is an implication in the discourse that one should not be judged by single errors made in youth, thereby building upon an idea that there exists an authentic identity or personality that can be put forward when removing false or embarrassing information (van Zoonen, 2013). This problem representation is uncontested in the discourse[2].

4.5.2 Implications for Google as a political subject. Due to the selection of material, there is a presumption in the debate that it is mainly articles in newspapers which individuals want to have removed from search hits. The fact that search engines and their algorithms have effects for media production is not new (Dwyer, 2016), but in the debate, many journalists argue that Google search produces a neutral results list and as such, the RTBF ruling can be understood as tampering with the freedom of the press. Content restriction on the internet, whether by companies or governments, is a highly ideological issue. The RTBF is an example of a safeguard measure but the power still lies within the internet companies. In the case of the RTBF ruling, both advocates and critics of the ruling have questioned Google’s mission to assess removal requests but it has not been discussed in the debate by representatives from the EU.

5. Conclusion
I have done an analysis of how the RTBF ruling has made possible different conceptualisations of Google.

To conclude, my findings indicate that Google is constructed as a political subject through the ruling as it makes Google’s power over knowledge production visible in the discourse. In the discourse, Google is conceptualised as a commercial company, a neutral facilitator of the world connected to the freedom of the internet and as a judge of character.

Critics of the RTBF describe the threat to Google as coming from policymakers; it can be seen as a reaction to political governance that is perceived as affecting the status quo of internet use. Thus, the critics do not take into account the fact that the internet has always been governed and that the practice of searching is dynamic. In many ways, Google and the technology of Google search have been constructed as neutral and objective in the discourse, but a consequence of the ruling is that Google’s advisors and executives are now pointing out that Google is a private company that is not required to be transparent and accountable in the way that governments must be. An interesting dynamic in the result is that Google is framed as both a commercial company and as a neutral facilitator of the world. These different conceptualisations have effects on how the internet should be governed, i.e., if Google is understood as a neutral tool, then no
governance is needed. If Google is understood as a commercial company, then governance may be needed of the company but, most importantly, the company itself should not be in charge of governing.

In the media reporting of the ruling, individuals affected by the ruling are constructed as either guilty or innocent. An assumption in the discourse is that Google search can and should reflect a single fixated identity and that the search engine provides a judge of character. Those who are for the ruling argue that the individual, with the aid of the RTBF, can manage his/her identity while those who are against the ruling argue that activities such as crime should be part of one’s digital identity, implying the possibilities for social control. Protecting privacy, through a right to be forgotten, is constructed as something for which governments should be responsible, not companies.

From a researcher’s viewpoint, the debate is an opportunity to deepen the layman’s understanding of Google. The ruling is framed as giving rights to individuals over companies but since the technological infrastructure of Google has not changed and since the assessments procedures of Google are not transparent, the company still holds the power. The implications for online information are that the hegemonic role of Google is, if not disturbed, at least coming into public view through the RTBF ruling. The ruling does give individuals the possibility of reputation management, while also opening up for a discussion on who has the right to be forgotten. Paradoxically, the ruling and the ongoing debate have made the politics of Google visible, as well as giving the company more power in judging knowledge dissemination.

Notes
1. For a discussion on different possibilities to be forgotten or deleted on the internet, see Mayer-Schönberger (2009).
2. For a discussion on ethical aspects of Google, see, for example, Diaz (2008), Hinman (2008), Hoffmann (2016).

References
C-131/12. Judgment of the court (grand chamber) of 13 May (2014), Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González, European Court of Justice.


Ronson, J. (2015), So you’ve been Publicly Shamed, Picador, London.


Material


Further reading

About the author
Linnéa Lindsköld is Senior Lecturer at the Swedish School of Library and Information Science, University of Borås, Sweden. During 2015-2016 Lindsköld participated in the research project Knowledge in a Digital World at Lund University. She is Researcher and Teacher in cultural-literature-and information policy. Her research interests include information policy, the concept of quality in cultural policy, literature policy, the politics of reading and Scandinavian radical right parties’ cultural policy. Linnéa Lindsköld can be contacted at: linnea.lindskold@hb.se

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Social media users’ share intention and subjective well-being
An empirical study based on WeChat

Liang Ma, Xin Zhang and Xiao Yan Ding
Shandong University of Finance and Economics, Jinan, China

Abstract
Purpose – The rise of social media has gained increasing attention in recent years; however, few studies have focused on social media users’ specific behavior and subjective well-being. To fill this research gap, the purpose of this paper is to develop an integrated model to investigate factors that affect social media user’s share intention and the relationship between user’s share intention and subjective well-being.

Design/methodology/approach – Structural equation model is used in this study. A field survey with 398 WeChat users is conducted to test the research model and hypotheses.

Findings – The empirical results show that: utilitarian value, hedonic value, user satisfaction and information source credibility are important factors affecting users’ share intention; users’ share intention positively affects user’s subjective well-being; moderating effects show that relative significance positively moderates the relationship between utilitarian value and users’ share intention; and users’ demographic characteristics differences actually exist in users’ share intention.

Originality/value – First, the authors clear that factors affect social media users’ share intention from the perspective of customer-perceived value. The results deepen our understanding about the factors that affect WeChat users’ share intention. Second, the authors focus on the effect of users’ specific behavior on users’ subjective well-being and found that users’ share intention is one of the important aspects that affect user’s subjective well-being. More importantly, the authors tested users’ characteristic differences in social media users’ share intention, which have previously received limited attention.

Keywords Perceived value, Subjective well-being, Information source credibility, Relative significance, Share intention

Paper type Research paper

Introduction
As the largest public service platform, WeChat has gained widespread attention in the academic and management circles (Che and Cao, 2014; Gan, 2017). WeChat is a Chinese social media (instant messaging, commerce and payment services) application developed by Tencent. The features of WeChat include messaging, official accounts, moments, WeChat pay payment services and so on. Specifically, WeChat official accounts, in this paper, refer to enterprise official accounts – a platform that can push messages and articles to users (customers) who follow these accounts. In a way, this feature establishes direct interaction between businesses and their follower through push notifications. Moments is WeChat’s brand name for its social newsfeed of friends’ updates (a mix of Facebook and Twitter), which allows users to post images, text, comment, share articles and “like” them along with status update on their own and friends’ pages. According to the newest “WeChat influence” report released by Tencent in 2016, the active number of WeChat users exceeded 889m, with the official account alone exceeding 8m, and advertising revenue being increased to RMB 3.679bn. With more and more users browsing articles of official accounts, sharing article information and commenting on WeChat, the application has become an important part of people’s lives in China.

Many scholars have studied WeChat users’ behavior, such as the posting “likes” behavior (Gan, 2017), word-of-mouth behavior (Che and Cao, 2014), satisfaction and stickiness behavior (Lien et al., 2016), continued usage intention (Zhang, Guo, Hu and Liu, 2016) and so on.
However, one of the most important aspects ignored by scholars is WeChat users’ shared intentions and behaviors in the Chinese context. When users share articles of official accounts to their moments, more and more people can see the articles being shared. At that time, the users who share such information become part of the medium for sharing information, which includes sharing of promotions and marketing of enterprise official accounts as well (Xu et al., 2015). Thus, factors that affect a user’s willingness to share is an important topic that deserves deep consideration. WeChat users often follow several official accounts to get information, such as those from newspapers and periodicals, information about enterprise management practice and so on. There are considerable amount of information pushed by official accounts, which might represent an overload. Therefore, we cannot help but wonder what factors will affect WeChat users’ share intention. Besides, previous studies have proven that different user characteristics such as sex, age and education show significant differences in the formation of behavioral intention (Mittal and Kamakura, 2001). It would be interesting and meaningful to find out if user characteristics such as sex, age and education lead to significant differences in the formation of users’ share intentions as well. Currently, the answer is uncertain.

Another important question is the relationship between social media users’ behavior and subjective well-being. On the one hand, as an important social medium in China, WeChat really brings convenience to users, in that they can browse information, communicate with friends or even work. On the other hand, more and more people are getting addicted to WeChat, even to the point of becoming obsessed. Although some scholars have begun to study the relationship between social media usage and users’ subjective well-being, no consistent conclusion has been reached (Gerson et al., 2016). Some studies report positive associations between social media use and subjective well-being (Grieve et al., 2013), while others report the opposite (Fox and Moreland, 2015). However, research on the effect of users’ specific behaviors on social media users’ subjective well-being is rare. Wasko and Faraj (2000) emphasize that some people enjoy sharing knowledge to help others, which brings in a kind of happiness. Thus, it would be interesting to find if there are associations between WeChat users’ share intentions and their subjective well-being. This paper therefore also addresses whether users with different characteristics will show differences in subjective well-being. The following research questions are then put forward:

**RQ1.** What factors affect WeChat users’ share intentions?

**RQ2.** Is there any relationship between WeChat users’ share intentions and subjective well-being?

Based on the above questions, this paper first developed an integrated model to study factors affecting WeChat users’ share intentions and the relationship between users’ share intentions and users’ subjective well-being. Then, we conducted an online survey and used structural equation modeling (SEM) to analyze the data in terms of the integrated model. These empirical results are presented and discussed in later part of the study.

This study offers several contributions. First, this paper has investigated factors affecting WeChat users’ share intentions from the perspective of customer-perceived value. The results deepen our understanding about factors affecting WeChat users’ share intentions, which differ from those for other social media. Second, the research reported in the current literature has ignored the effect of social media users’ specific behaviors on the users’ subjective well-being. This paper filled that gap and found that a user’s share intention is one of the important aspects affecting the user’s subjective well-being. In addition, this paper examined user characteristics in relation to social media users’ share intentions and subjective well-being, an aspect which has previously received limited attention.
This paper is organized as follows. The second section provides the literature review and hypotheses. The third section describes research methods, followed by data analysis in the fourth section. The fifth section discusses the research findings, followed by implications and limitations.

**Literature review and hypotheses**

**Customer-perceived value and share intention**

The theory of customer-perceived value is defined as the overall evaluation of benefits of a product or service that the customer can perceive and the trade-off between gain and cost of product or service. Research on the theory of customer-perceived value is mainly divided into two camps. One is from the perspective of customer perception, which means customers gain benefits from a product or service, the main concepts including customer value and experience value. The other perspective is that of the enterprise, which refers to the value of the customer to the enterprise, the main concepts including customer lifetime value and customer engagement value. On the dimensions of customer-perceived value, most scholars divide customer-perceived value into two dimensions: utilitarian value and hedonic value (Yu et al., 2013). In this paper, we study WeChat users’ value perception of official accounts, so we use the former meaning of customer value, that is, the benefit a user gains from the official account service. Furthermore, this study studied customer-perceived value according to the two dimensions of utilitarian value and hedonic value, which is accepted by most of scholars. Utilitarian value refers to the perceived usability of the WeChat official account, while hedonic value refers to the entertainment value of the WeChat official account. A large number of scholars have found that once a customer perceives good value, he/she will respond with a series of positive behaviors, such as a positive attitude, customer satisfaction and even knowledge sharing (Rho et al., 2014; Zhang, Li, Wu and Li, 2016). In this study, we will draw upon the customer-perceived value theory to explain how user’s value perception affects the user’s share intentions.

This paper defines share intention as a user’s intention to share articles from enterprise official accounts to the user’s social media (e.g. WeChat) moments. Users’ share intentions have been widely studied in the areas of virtual communities, online learning communities and so on (Yilmaz, 2016). In the context of WeChat, a user’s share intention and behavior has significance to both enterprises and users (Che and Cao, 2014). Therefore, it would be beneficial to establish a model to explore factors affecting WeChat users’ share intentions further. A plethora of scholars have initiated studies on social media users’ share behaviors. For example, Fu et al. (2017) investigated Facebook users’ content sharing behaviors and found psychological incentive, social capital focus and content type to be important factors affecting Facebook user behaviors. Moghavvemi et al. (2017) found that perceived enjoyment, perceived reciprocal benefits and knowledge power affect students’ knowledge sharing on Facebook. However, literature on WeChat users’ share intentions and behaviors is rare. To fill the gap, this paper investigates and analyzes factors affecting WeChat users’ share intentions. The theory of perceived value suggests that customer-perceived value will have an impact on users’ behavior intentions (Kim and Oh, 2011). When a user perceives value in information, she/he will generate a series of positive responses, such as share intention (Moghavvemi et al., 2017). And this paper also argues that the greater a user’s enjoyment and perceived usefulness of an article, the higher the possibility of willingness to share. Based on the above exploration, this paper puts forth the following hypotheses:

**H1.** Utilitarian value has a positive effect on users’ share intentions.

**H2.** Hedonic value has a positive effect on users’ share intentions.
Users’ satisfaction and share intention

Satisfaction is considered as response to consumer’s experience, and some scholars believe that satisfaction is a comprehensive evaluation based on the purchase and consumption of products or services (lazzi et al., 2016). This paper defines satisfaction as users’ satisfaction achieved through degree of information pushed by WeChat official account. Customer satisfaction theory suggests that when customers perceive satisfaction they will generate a series of behavioral intentions (Fu and Juan, 2016). A review of customer satisfaction theory highlighted some scholars’ assertion on the formation mechanism of customer satisfaction (Homburg et al., 2006), and other scholars’ establishment of the relationship between customer satisfaction and customers’ word-of-mouth behavior, increased loyalty and so on (Chang and Wang, 2011). Both parties have obtained a consistent conclusion – when customers are satisfied, they generate a series of positive behaviors, such as positive word-of-mouth behavior, consumer loyalty behavior, share intentions and so on (Ma and Agarwal, 2007; Hsu et al., 2014).

According to customer satisfaction theory, when users are satisfied, the series of behaviors generated are often beneficial to the company, and share intention is one of those favorable behaviors. For example, Hashim and Tan (2015) found that level of satisfaction toward the online community platform positively affects the user’s share intention. Consequently, in the field of social media, the current research also hypothesizes that when users are satisfied with the information pushed by WeChat official account, they will be more likely to have a willingness to share. Based on the above, this paper suggests the following hypothesis:

H3. Users’ satisfaction with the pushed information positively affects user’s share intention.

Information source credibility and share intention

The credibility of information sources is defined as the judgment made by information communicators to consider the credibility of the information disseminated (O’Keefe, 2002). For the purposes of the current research, credibility of information sources is defined as the judgment of WeChat users considering the credibility of WeChat official accounts. The three dimensions of the credibility of information sources are professional ability, ability to trust and reputation. As an important indicator, the credibility of information sources has been widely used to evaluate persuasive information. According to the elaboration likelihood model (ELM), source credibility can be a peripheral cue of how persuasive messages affect user behavior. In recent years, credibility of information sources has been gradually applied to the study of social media and social networking websites, and researchers have proven that source credibility influences users’ perceptions of the usefulness of information as well as their behavioral intentions (Chiou et al., 2013; Li, 2013; Hussain et al., 2017).

Compared with traditional offline communication, in the era of social media, because of the virtual characteristics of the network, the credibility of the information source becomes particularly important (Go et al., 2016). There is considerable information contained in social media. The user must effort to distinguish which information is true. Generally speaking, the perception of the credibility of information is affected by the users’ knowledge and the credibility of information sources. For example, users with certain knowledge usually have higher information screening capabilities than users with less knowledge. In addition, large research institutions have published reports that are more credible than those of small organizations. Studies have confirmed that the credibility of information sources will have an impact on users’ behavioral intentions, such as share intentions (Hu and Sundar, 2010; Li, 2013). The current research also holds that when the source credibility is high, the
willingness to share information is also increased since users tend to trust the information source. Based on the above, the current research proposes the following hypothesis:

\[ H4. \text{ Information source credibility positively affects a user's share intention.} \]

Relative significance

Relative significance refers to the extent to which the text or content is relevant to the user and the degree of importance (Bhattacherjee and Sanford, 2006). Relative significance in the current research refers to the degree of relevance and importance of the information contained in WeChat official accounts. Different from the perceived value, relative significance is more concerned with the relevance and importance of the information to the user, while the latter is more focused on the value of the article to the user. Reviews of previous studies show that various scholars have obtained a similar conclusion: under the same conditions, the stronger the relative significance, the stronger the attitude and behavioral intention (Chang et al., 2015; Winter et al., 2015).

In the field of social media, most scholars ignore the effect of relative significance on a user’s share intention. According to Chang et al. (2015), users’ share intentions will increase more dramatically when the users perceive the relative significance of the messages. In addition, according to the ELM, the effects of source credibility are moderated by a potential user’s motivation and ability to elaborate on informational messages. This study operationalizes the motivation dimension of elaboration as relative significance, defined as the message recipient’s perceived relevance of a pushed article to her/his knowledge system (Bhattacherjee and Sanford, 2006). Users are more likely to exhibit share intention when they perceive that the pushed information is more relative to their knowledge systems under the same level of source credibility. Thus, the current research also holds that users’ share intentions are more intense when associated with users’ perceived relative significance of the information in the WeChat official account. Based on the above, the following hypotheses are proposed:

\[ H5a. \text{ Relative significance positively moderates the relationship between utilitarian value and user's share intention.} \]

\[ H5b. \text{ Relative significance positively moderates the relationship between hedonic value and user's share intention.} \]

\[ H5c. \text{ Relative significance positively moderates the relationship between users' satisfaction with the article and user's share intention.} \]

\[ H5d. \text{ Relative significance positively moderates the relationship between source credibility of the official account and user's share intention.} \]

Share intention and subjective well-being

Subjective well-being is defined as “a broad category of phenomena that includes people’s emotional responses, domain satisfactions, and global judgments of life satisfaction” (Diener et al., 1999). Subjective well-being has gained increasing attention in recent years, and it has been studied in the fields of sociology, psychology and even social media (Knight et al., 2009; Wei and Gao, 2016). The concept of life satisfaction is frequently used as the measure of well-being in studies of social networking sites and their usage (Grieve et al., 2013; Gerson et al., 2016). Thus, the current research also uses the life satisfaction measure to assess users’ subjective well-being. Life satisfaction is a subjective evaluation of the quality of life and is based on the standard set by oneself. Empirical studies on the satisfaction of life began in the 1960s. These studies have either focused
mainly on the effect of marital status and the environment on the quality of life (Bailey, 2007), or on the study of adult and geriatric life satisfaction (Ghubach et al., 2010).

With the development of social media, more and more scholars have realized the important relationship between social media use and users’ subjective well-being (Gerson et al., 2016). However, no consistent conclusion has been reached. Some scholars hold that social media use is positively associated with life satisfaction (Grieve et al., 2013; Ishii, 2017), while others have found that social media use is negatively associated with life satisfaction (Satici and Uysal, 2015; Gerson et al., 2016). One of the possible reasons is that there are so many social media user behaviors; one of them may lead to a user’s life satisfaction, while another may lead to user’s dissatisfaction. However, research on the effect of a user’s specific behavior on the user’s life satisfaction is rare. Thus, to fill that gap, the current research considers a specific user behavior (share intention) and examines its effect on life satisfaction. Regarding the relationship between share intention and life satisfaction, Gagné (2009) suggests that knowledge sharing has a positive impact on an individual’s subjective well-being, which is supported by Jiang and Hu (2016). The current research proposes that when users demonstrate share intentions of information pushed by WeChat official accounts, those users are happy to help others, and their own social value rises, so correspondingly their life satisfaction increases. Based on the above, this research suggests the following hypothesis:

\[ H6. \text{ A user’s share intention positively influences her/his life satisfaction.} \]

User characteristic differences
The effect of user characteristics has attracted many researchers’ interest in further investigation (Lian and Yen, 2014; Pascual-Miguel et al., 2015). User characteristics include mainly sex, age, education, income and so on. Considerable research has emphasized user characteristic differences in users’ attitude, risk perception, emotion and behavior intention (Garbarino and Strahilevitz, 2004; Wu et al., 2017). However, little attention has been given to whether different user characteristics will result in differences in the formation of WeChat users’ share intentions and subjective well-being. Prior research has proven that users’ emotions and behavior intentions differ because of perceptions of risks, attitudes and emotions that vary between women and men (Dai et al., 2013; Lian and Yen, 2014), older and younger users (Liébana-Cabanillas et al., 2014), education qualifications (Forgas-Coll et al., 2013) and income (See-To et al., 2014). Thus, current research also holds that WeChat users’ share intentions and subjective well-being will vary according to users’ characteristics. In light of the aforementioned research, the following hypothesis is proposed:

\[ H7. \text{ WeChat users’ share intention and subjective well-being will vary according to user sex, age, education and income.} \]

Proposed model
Based on the above discussion, this paper puts forward the research model as shown in Figure 1. To sum up, utilitarian value, hedonic value, user satisfaction and source reliability positively affect users’ share intention. And users’ share intention positively affects users’ life satisfaction. In addition, relative significance acts in a positive moderating role between utilitarian value, hedonic value, user satisfaction, source reliability and users’ share intentions.

Research methodology
Measures
A questionnaire was used to validate the conceptual model. Items that measured each variable were developed from prior literature. Those measure and their factor loadings are
shown in Table I. The three items of utilitarian value and four items of hedonic value were adapted from Vries and Carlson (2014). The scale of user satisfaction was revised from Pavlou (2003). The scale of information source credibility was taken from Filieri (2015). The three items of share intention were adapted from Lee and Ma (2012). The relative significance measures were adapted from Bhattacharjee and Sanford (2006). Finally, the scale of life satisfaction was revised from Iii et al. (2015). All items were measured with a seven-point Likert scale ranging from 1 (strongly disagree/unlikely) to 7 (strongly agree/likely).

Figure 1. Research model

<table>
<thead>
<tr>
<th>Factor</th>
<th>Measure items</th>
<th>Factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utilitarian value (UV)</td>
<td>UV1. The article of the WeChat official account is helpful for me 0.907</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UV2. The article of the WeChat official account is useful for me 0.886</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UV3. The article of the WeChat official account is practical for me 0.848</td>
<td></td>
</tr>
<tr>
<td>Hedonic value (HV)</td>
<td>HV1. The article of the WeChat official account is fun 0.897</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HV2. The article of the WeChat official account is exciting 0.851</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HV3. The article of the WeChat official account is pleasant 0.861</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HV4. The article of the WeChat official account is entertaining 0.820</td>
<td></td>
</tr>
<tr>
<td>User satisfaction (US)</td>
<td>US1. I am satisfied with the information I have received from WeChat official account 0.906</td>
<td></td>
</tr>
<tr>
<td></td>
<td>US2. I am satisfied with my previous experiences with WeChat official account 0.876</td>
<td></td>
</tr>
<tr>
<td></td>
<td>US3. I am satisfied with articles I have received from WeChat official account 0.903</td>
<td></td>
</tr>
<tr>
<td>Source credibility (SC)</td>
<td>SC1. The WeChat official account is credible 0.884</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SC2. The WeChat official account is experienced 0.881</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SC3. The WeChat official account is trustworthy 0.852</td>
<td></td>
</tr>
<tr>
<td>Share intention (SI)</td>
<td>SI1. I intend to share the articles of WeChat official account in the future 0.931</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SI2. I expect to share the articles of WeChat official account 0.933</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SI3. I plan to share the articles of WeChat official account 0.923</td>
<td></td>
</tr>
<tr>
<td>Relative significance (RI)</td>
<td>RI1. I think the articles of WeChat official account are important to my daily life 0.868</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RI2. I think the articles of WeChat official account are relevant to knowledge in my life 0.860</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RI3. The articles of WeChat official account are considerable in my daily life 0.872</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RH4. I always read the articles of WeChat official account 0.859</td>
<td></td>
</tr>
<tr>
<td>Life satisfaction (LS)</td>
<td>LS1. Overall, my experience with WeChat official account was memorable having enriched my quality of life 0.895</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LS2. In general, I felt good about my life shortly after read the articles of WeChat official account 0.905</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LS3. After read the articles of WeChat official account I felt that I lead a meaningful and fulfilling life 0.895</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LS4. Overall, I felt happy upon read the articles of WeChat official account 0.896</td>
<td></td>
</tr>
</tbody>
</table>
Data collection
A small-scale pilot test with three professors and five post-graduate students was conducted to assess the questionnaire’s logical consistency, ease of understanding and contextual relevance. A total of 416 respondents participated in an online survey with the help of a professional questionnaire platform (Soujoup.com) in China. Reason behind larger sampled population is reasoned with the company’s provision of monetary reward cash (¥1) to each person who completed the questionnaire. The respondents were distributed across 26 provinces of China. Among the questionnaires received, 398 were valid and 18 were either incomplete or not answered. An additional questionnaire was administered to collect respondents’ demographic information, including sex, age, education and income. Furthermore, since the length of WeChat use, frequency of WeChat use per day and number of official account users followed may influence users’ life satisfaction, these three variables were added as control variables. The descriptive statistics of respondent characteristics and WeChat usage is summarized in Table II.

Data analysis
Partial least squares (PLS) SEM, a component-based estimation approach instead of covariance-based, was used to analyze the data. As a second-generation multivariate technique, PLS can simultaneously assess the measurement model and the structural model. Compared to the covariance-based SEM, PLS requires a relatively small sample size, does not rely on the assumption of a normal distribution and is more appropriate for exploratory analysis and for handling formative constructs (Chen et al., 2016). Thus, PLS was determined to be more suitable for this current research. Following the two-step analytical procedure (Hair et al., 2009), this research examines the measurement model and the structural model in that order.

Measurement model
The measurement model can be assessed by examining the reliability, convergent validity and discriminant validity. Specifically, reliability can be assessed by determining Cronbach’s $\alpha$.

<table>
<thead>
<tr>
<th>Demographic variable</th>
<th>Size</th>
<th>%</th>
<th>Demographic variable</th>
<th>Size</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>246</td>
<td>61.78</td>
<td>Senior middle school or below</td>
<td>21</td>
<td>5.33</td>
</tr>
<tr>
<td>Female</td>
<td>132</td>
<td>38.22</td>
<td>Junior college</td>
<td>81</td>
<td>20.44</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td>Bachelor’s degree</td>
<td>249</td>
<td>62.67</td>
</tr>
<tr>
<td>≤ 20</td>
<td>5</td>
<td>1.33</td>
<td>Master’s degree or above</td>
<td>46</td>
<td>11.56</td>
</tr>
<tr>
<td>21–30</td>
<td>209</td>
<td>52.44</td>
<td>Frequency use of WeChat per day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31–40</td>
<td>142</td>
<td>35.56</td>
<td>≤3 times</td>
<td>50</td>
<td>12.44</td>
</tr>
<tr>
<td>40–50</td>
<td>34</td>
<td>8.44</td>
<td>4–7 times</td>
<td>110</td>
<td>27.56</td>
</tr>
<tr>
<td>&gt; 50</td>
<td>9</td>
<td>2.22</td>
<td>8–11 times</td>
<td>94</td>
<td>23.56</td>
</tr>
<tr>
<td>Monthly personal income (RMB)</td>
<td></td>
<td></td>
<td>12–15 times</td>
<td>48</td>
<td>12.00</td>
</tr>
<tr>
<td>≤ 3,000</td>
<td>71</td>
<td>17.78</td>
<td>16–19 times</td>
<td>21</td>
<td>5.33</td>
</tr>
<tr>
<td>3,001–5,000</td>
<td>140</td>
<td>35.11</td>
<td>&gt; 20 times</td>
<td>76</td>
<td>19.11</td>
</tr>
<tr>
<td>5,001–8,000</td>
<td>104</td>
<td>26.22</td>
<td>Number of WeChat official account users follow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 8,001</td>
<td>83</td>
<td>20.89</td>
<td>≤10</td>
<td>111</td>
<td>28.00</td>
</tr>
<tr>
<td>The length of WeChat use (years)</td>
<td></td>
<td></td>
<td>10–20</td>
<td>120</td>
<td>30.22</td>
</tr>
<tr>
<td>≤1</td>
<td>28</td>
<td>7.11</td>
<td>10–30</td>
<td>106</td>
<td>26.67</td>
</tr>
<tr>
<td>1–3</td>
<td>207</td>
<td>52.00</td>
<td>30–40</td>
<td>32</td>
<td>8.00</td>
</tr>
<tr>
<td>3–5</td>
<td>117</td>
<td>29.33</td>
<td>&gt; 40</td>
<td>28</td>
<td>7.11</td>
</tr>
<tr>
<td>&gt; 5</td>
<td>46</td>
<td>11.56</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table II. Descriptive statistics of respondent characteristics
composite reliability (CR) and average variance extracted (AVE). According to the suggestions of Hair et al. (2009), Cronbach’s α is acceptable when it exceeds 0.7, CR is acceptable when it exceeds 0.7 and AVE is acceptable when it exceeds 0.5. As shown in Table III, Cronbach’s α for this study ranges from 0.843 to 0.921, all of which exceed the recommended value of 0.7. CR ranges from 0.905 to 0.950, which exceeds the recommended level of 0.70. AVE ranges from 0.735 to 0.863, which exceeds the recommended level of 0.50. Thus, the measurement constructs have high reliability. And convergent validity can be assessed by factor loading. As shown in Table I, all of the factor loadings were higher than 0.8 at the significant level of \(p < 0.01\), suggesting good convergent validity. Meanwhile, this paper examined the discriminant validity of the measurement items. As shown in Table III, the square root of each factor’s AVE is larger than its corresponding correlation coefficients with other factors, indicating the adequate discriminant validity.

**Structural model**

The PLS results of the structural model are shown in Figure 2. It can be seen in Figure 2 that both utilitarian value and hedonic value have a positive effect on the users’ share intention. Thus, \(H1\) and \(H2\) are supported. In other words, the users’ share intention will rise with the increase of perceived utilitarian value and hedonic value. Meanwhile, user satisfaction has a positive effect on their share intention. Thus, \(H3\) is supported.

<table>
<thead>
<tr>
<th>Items</th>
<th>AVE</th>
<th>Composite reliability</th>
<th>Cronbach’s α</th>
<th>UV</th>
<th>HV</th>
<th>CS</th>
<th>SC</th>
<th>RI</th>
<th>SI</th>
<th>LS</th>
</tr>
</thead>
<tbody>
<tr>
<td>UV</td>
<td>0.785</td>
<td>0.916</td>
<td>0.863</td>
<td>0.886</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HV</td>
<td>0.735</td>
<td>0.917</td>
<td>0.881</td>
<td>0.737</td>
<td>0.858</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>0.804</td>
<td>0.925</td>
<td>0.878</td>
<td>0.639</td>
<td>0.742</td>
<td>0.897</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>0.761</td>
<td>0.905</td>
<td>0.843</td>
<td>0.701</td>
<td>0.753</td>
<td>0.723</td>
<td>0.872</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RI</td>
<td>0.748</td>
<td>0.922</td>
<td>0.888</td>
<td>0.758</td>
<td>0.738</td>
<td>0.774</td>
<td>0.702</td>
<td>0.865</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>0.863</td>
<td>0.95</td>
<td>0.921</td>
<td>0.614</td>
<td>0.643</td>
<td>0.661</td>
<td>0.651</td>
<td>0.748</td>
<td>0.929</td>
<td></td>
</tr>
<tr>
<td>LS</td>
<td>0.806</td>
<td>0.943</td>
<td>0.920</td>
<td>0.733</td>
<td>0.746</td>
<td>0.739</td>
<td>0.712</td>
<td>0.775</td>
<td>0.759</td>
<td>0.898</td>
</tr>
</tbody>
</table>

Table III. Descriptive statistics and inter-construct correlations

**Notes:** UV, utilitarian value; HV, hedonic value; CS, customer satisfaction; SC, information source credibility; RI, relative significance; SI, share intention; LS, life satisfaction

**Figure 2.** Results of structure model analysis

**Notes:** *\(p<0.05\); **\(p<0.01\); ***\(p<0.001\)
This implies that when the users are satisfied with the article pushed by the WeChat official account, they will be more likely to have a willingness to share. Furthermore, information source credibility has a positive effect on the users’ share intention. Thus, $H_4$ is supported. This means that when the users share information, they will consider the source credibility of the information. Finally, the users’ share intention has a positive effect on their life satisfaction. Thus, $H_6$ is supported. This tells us that the level of the users’ life satisfaction will rise when they have a willingness to share. In addition, the results of the control variable show that only the number of WeChat official accounts that users follow has a positive effect on their life satisfaction, while the effects of the length of WeChat use and the frequency of WeChat use on the users’ life satisfaction are not significant. Finally, the $R^2$ square value is also calculated.

**Moderate effect**

The moderating role can be tested by accessing the differences in path coefficients for each subgroup (Hsieh et al., 2008). The users’ characteristic differences include the differences of their gender, age, education and income. To examine the differences in the users’ characteristics, path comparison testing was conducted between both groups. The standard used to dichotomize the groups is as follows: according to the suggestion of Liébana-Cabanillas et al. (2014), and considering the age status of WeChat, this paper defines the younger group (age $\leq$ 30, group 1) and older group (age $> 30$, group 2); according to the suggestion of Forgas-Coll et al. (2013), and considering the education status of China, this paper defines the lower education group (non-university degree or below, group 1) and higher education group (university degree or above, group 2); according to the suggestion of See-To et al. (2014), and considering the income status of China, this paper defines the lower income group (income $\leq$ RMB 5,000, group 1) and higher income group (income $> RMB 5,000$, group 2). First, this study tested $H_1$–$H_4$ and $H_6$ separately for both groups. We then compared the path coefficients based on the method developed by Keil and Wassenaar (2000), which is widely used in the IS context to test the moderating effects when the moderating variable is categorical (Chang et al., 2014; Guo et al., 2015). The results are shown in Table IV.

Based on the results of multi-group comparisons, $H_7$ was partially supported. The results show that:

1. **Gender differences** exist in the users’ share intention and subjective well-being.
   - The effect of utilitarian value on the users’ share intention was significantly higher for males than for females, which means that males were more likely to have a willingness to share when they have perceived utilitarian value. Meanwhile, gender differences were not found among the effect of hedonic value on the users’ share intention. In addition, the effect of user satisfaction on the users’ share intention was significantly higher for males than for females, which means that males were more likely to have a willingness to share when satisfied. The effect of information source credibility on the users’ share intention was significantly higher for females than for males, which means that females were more likely to have a willingness to share when they have perceived information source credibility. Finally, the effect of the users’ share intention on their life satisfaction was significantly higher for females than for males, which means that females were more likely to be satisfied with life when they have a willingness to share.

2. **Age differences** were found in the users’ share intention and subjective well-being.
   - The effect of utilitarian value on the users’ share intention was significantly higher for the older group than for the younger group, which means that the older users were more likely to have a willingness to share when they have perceived utilitarian value.
value. Meanwhile, the effect of hedonic value on the users’ share intention was significantly higher for the older group than for the younger group, which means that older users were more likely to have a willingness to share when they have perceived hedonic value. The effect of user satisfaction on the users’ share intention was significantly higher for the younger group than for the older group, which means that younger users were more likely to have a willingness to share when satisfied. In addition, the effect of information source credibility on the users’ share intention was significantly higher for the younger group than for the older group, which means that younger users were more likely to have the willingness to share when they have perceived information source credibility. However, age differences were not found for the effect of the users’ share intention on their life satisfaction.

(3) Education differences were found in the users’ share intention and subjective well-being. The effect of utilitarian value on the users’ share intention was significantly higher for users with lower education than for users with a higher education, which means that users with a lower education were more likely to have a willingness to share when they have perceived utilitarian value. Meanwhile, the effect of hedonic value on the users’ share intention was significantly higher for users with a lower education than for users with a higher education, which means that users with a lower education were more likely to have a willingness to share when they have perceived hedonic value. The effect of user satisfaction on users’ share intention was significantly higher for users with a higher education than for users with a lower education, which means that users with a higher education were more likely to have a willingness to share when satisfied. In addition, the effect of information source credibility on the users’ share intention was significantly higher for users with a higher education than for users with a lower education, which means that users with a higher education were more likely to have a willingness to share when they have perceived information source credibility. However, age differences were not found for the effect of the users’ share intention on their life satisfaction.

<table>
<thead>
<tr>
<th>Items</th>
<th>Path</th>
<th>PC1</th>
<th>PC2</th>
<th>SE1</th>
<th>SE2</th>
<th>T</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (G1 = male) (G2 = female)</td>
<td>UV → SI</td>
<td>0.185**</td>
<td>0.166**</td>
<td>0.056</td>
<td>0.056</td>
<td>3.467</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>HV → SI</td>
<td>0.171*</td>
<td>0.170**</td>
<td>0.068</td>
<td>0.064</td>
<td>0.062</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>US → SI</td>
<td>0.324**</td>
<td>0.130</td>
<td>0.060</td>
<td>0.067</td>
<td>30.355</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>SC → SI</td>
<td>0.126</td>
<td>0.376**</td>
<td>0.074</td>
<td>0.063</td>
<td>35.373</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>SI → LS</td>
<td>0.758**</td>
<td>0.779**</td>
<td>0.027</td>
<td>0.021</td>
<td>8.730</td>
<td>O</td>
</tr>
<tr>
<td>Age (G1 = younger) (G2 = older)</td>
<td>UV → SI</td>
<td>0.173**</td>
<td>0.316**</td>
<td>0.059</td>
<td>0.043</td>
<td>27.595</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>HV → SI</td>
<td>0.141*</td>
<td>0.215**</td>
<td>0.060</td>
<td>0.060</td>
<td>12.261</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>US → SI</td>
<td>0.287**</td>
<td>0.217**</td>
<td>0.063</td>
<td>0.069</td>
<td>10.540</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>SC → SI</td>
<td>0.215**</td>
<td>0.066</td>
<td>0.072</td>
<td>0.062</td>
<td>22.109</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>SI → LS</td>
<td>0.765**</td>
<td>0.765**</td>
<td>0.024</td>
<td>0.023</td>
<td>0.307</td>
<td>X</td>
</tr>
<tr>
<td>Education (G1 = lower) (G2 = higher)</td>
<td>UV → SI</td>
<td>0.224**</td>
<td>0.140**</td>
<td>0.057</td>
<td>0.048</td>
<td>13.883</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>HV → SI</td>
<td>0.158*</td>
<td>0.141*</td>
<td>0.071</td>
<td>0.057</td>
<td>2.347</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>US → SI</td>
<td>0.208**</td>
<td>0.313**</td>
<td>0.079</td>
<td>0.057</td>
<td>13.731</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>SC → SI</td>
<td>0.156</td>
<td>0.243**</td>
<td>0.082</td>
<td>0.063</td>
<td>10.557</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>SI → LS</td>
<td>0.761**</td>
<td>0.768**</td>
<td>0.028</td>
<td>0.024</td>
<td>2.319</td>
<td>O</td>
</tr>
<tr>
<td>Income (G1 = lower) (G2 = higher)</td>
<td>UV → SI</td>
<td>0.152*</td>
<td>0.203**</td>
<td>0.072</td>
<td>0.038</td>
<td>8.863</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>HV → SI</td>
<td>0.024</td>
<td>0.288**</td>
<td>0.059</td>
<td>0.056</td>
<td>45.642</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>US → SI</td>
<td>0.286**</td>
<td>0.295**</td>
<td>0.079</td>
<td>0.057</td>
<td>13.731</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>SC → SI</td>
<td>0.323**</td>
<td>0.065</td>
<td>0.069</td>
<td>0.063</td>
<td>10.557</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>SI → LS</td>
<td>0.761**</td>
<td>0.768**</td>
<td>0.028</td>
<td>0.024</td>
<td>2.319</td>
<td>O</td>
</tr>
</tbody>
</table>

Notes: $S_{\text{pooled}} = \sqrt{\frac{\left( \frac{1}{N_1-1} + \frac{1}{N_1+N_2-2} \right) \times SE_{1}^{2} + \left( \frac{1}{N_1} + \frac{1}{N_1+N_2-2} \right) \times SE_{2}^{2} }{ \left( \frac{1}{N_1} + \frac{1}{N_2} \right) \times \left( \frac{N_1-1}{N_1} + \frac{1}{N_2} \right) }}, \hspace{1cm} t = (PC_1 - PC_2) / SE_{\text{pooled}}$, $N_1$, sample size of dataset for group $i$; SE, standard error of path in structural model of group $i$; PC, path coefficient in structural model of group $i$; O, support; X, not support. *$p < 0.05$; **$p < 0.01$
higher education were more likely to have a willingness to share when they have perceived information source credibility. Finally, the effect of users’ share intention on the users’ life satisfaction was significantly higher for users with a higher education than for users with a lower education, which means that more highly educated users were more likely to generate life satisfaction when they have a willingness to share.

(4) Income differences were found in the users’ share intention and subjective well-being. The effect of utilitarian value on the users’ share intention was significantly higher for higher income users than for lower income users, which means that higher income users were more likely to have a willingness to share when they have perceived utilitarian value. Meanwhile, the effect of hedonic value on the users’ share intention was significantly higher for higher income users than for lower income users, which means that higher income users were more likely to have a willingness to share when they have perceived hedonic value. However, income differences were not found among the effect of user satisfaction on the users’ share intention. In addition, the effect of information source credibility on the users’ share intention was significantly higher for lower income users than for higher income users, which means that lower income users were more likely to have a willingness to share when they have perceived information source credibility. Finally, the effect of the users’ share intention on the users’ life satisfaction was significantly higher for lower income users than for higher income users, which means that lower income users were more likely to generate life satisfaction when they have the willingness to share.

A multiple hierarchical regression method was used to test the moderating effect of relative significance (Keating et al., 2009). Before conducting multiple hierarchical regression analysis, the Anderson–Rubin method was used to calculate the factor scores of the constructs and make them center to zero. The results show that only H5α is supported ($β = 0.111, t = 2.189, p < 0.05, \text{one-tailed}$), as shown in Table V, which states that relative significance positively moderates the relationship between utilitarian value and users’ share intention. Figure 3 shows this moderating effect. As predicted, under a high level of relative significance, the user’s share intention increases considerably and more prominently as utilitarian value rises than under low levels of relative significance. However, relative significance did not act as a moderator in the other moderate hypotheses test.

**Discussion**

**Key findings**

The purposes of this study are to examine the main factors affecting WeChat users’ share intention and to examine the relationship between WeChat users’ share intention and subjective well-being. Based on these two purposes, this paper constructed an integrated model. In addition, utilitarian value, hedonic value, user satisfaction, information source credibility, and relative significance were considered. The results show that higher education users were more likely to have a willingness to share when they have perceived information source credibility. A multiple hierarchical regression method was used to test the moderating effect of relative significance. Table V shows the results of this test, which indicates that only H5α is supported ($β = 0.111, t = 2.189, p < 0.05, \text{one-tailed}$), as shown in Figure 3. The results show that relative significance positively moderates the relationship between utilitarian value and users’ share intention. As predicted, under a high level of relative significance, the user’s share intention increases considerably and more prominently as utilitarian value rises than under low levels of relative significance. However, relative significance did not act as a moderator in the other moderate hypotheses test.
credibility and relative significance are considered to be important factors that affect the users’ share intention and subjective well-being.

First, both utilitarian value and hedonic value increase WeChat users’ share intention, where the effect of utilitarian value on the users’ share intention is larger than the effect of hedonic value on the users’ share intention. This means that the more utilitarian or hedonic value of the information that the users perceived, the more likely they portray their willingness to share. Fu et al. (2017) investigated Facebook users’ share content behavior and found that psychological incentive, social capital focus and content type are important factors which affect such behavior. Our paper proved that perceived value, which includes utilitarian value and hedonic value, is an important factor that affects WeChat users’ share intention.

Second, WeChat users’ satisfaction with pushed information has a positive effect on the user’s share intention. This implies that when users are satisfied with the information pushed by WeChat official accounts, they are more likely to generate willingness to share. Hashim and Tan (2015) found that the level of satisfaction toward the online community platform positively affects the user’s share intention. Apart from the platform, this paper has proven that the users’ share intention can also be affected by their satisfaction with the article information.

Third, information source credibility is an important factor that affects WeChat users’ share intention. This means that before having a willingness to share, users consider the credibility of the information source. Generally speaking, the more credible the information source, the more likely the users project willingness to share. Prior studies have confirmed that the credibility of information sources has an impact on users’ behavioral intentions (Hu and Sundar, 2010; Li, 2013). This paper takes a further step and finds that credibility of information source positively affects WeChat users’ share intention.

Fourth, our result shows that WeChat users’ share intention affects their subjective well-being (life satisfaction). In other words, the users portray high subjective well-being when they have a willingness to share. One of the explanations behind this attitude is that when users have a willingness to share, they feel that the article they share will help others learn new knowledge, and their social value will be better reflected. Thus, they will feel a high subjective well-being.

Finally, the moderating effects show that, first, the users’ characteristic differences actually exist in their share intention and subjective well-being. Considerable research has emphasized user characteristic differences in users’ attitude, risk perception, emotion and behavior intention (Wu et al., 2017). This paper takes a further step ahead and finds that user characteristic differences also exist in the formation of WeChat users’ share intentions and subjective well-being. Second, relative significance acts as a positive moderating role between utilitarian value and users’ share intention, although other moderating effect is not found. According to Chang et al. (2015), users’ share intentions increase more dramatically
when they perceive the relative significance of messages. In line with Chang et al.’s findings, this study establishes such increase in share intention only under utilitarian value. There is no dramatic increase in share intention with the increase of relative significance under the same level of hedonic value, user satisfaction and information source credibility, nonetheless.

**Theoretical implications**

The theoretical implications mainly include the following three aspects: first, this paper contributes to customer-perceived value theory. Existing research on customer-perceived value theory has mainly focused on virtual community and enterprise organization (Dai et al., 2013; Chang et al., 2014). However, the application of customer-perceived value theory in the area of social media is rare. This paper takes a step ahead and applies customer-perceived value theory into the area of social media (WeChat) and finds users having a willingness to share once their requirements are satisfied, which has important theoretical implications.

Second, the current research on users’ subjective well-being has mainly been researched in the disciplines of sociology and psychology (Diener et al., 2002; Knight et al., 2009; Wei and Gao, 2016). Although some scholars have begun to study the relationship between social media use and users’ subjective well-being, no consistent conclusion has been reached (Gerson et al., 2016). Research on the effect of the users’ specific behavior on their subjective well-being is rare, wherein this paper filled the gap and has found that users’ share intention is an important factor affecting their subjective well-being.

Third, this study identifies new factors (e.g. information source credibility, relative significance) that affect the social users’ share intention in the Chinese context, which has rarely been researched. This study found that user satisfaction has the greatest impact on the users’ share intention, followed by information source credibility, utilitarian value and hedonic value. In addition, relative significance has a positive moderating role between utilitarian value and the users’ share intention. Most importantly, this paper tested the users’ characteristic differences in their share intention on social media, which has previously received limited attention. Our results show that the users’ characteristic differences actually exist in the social media users’ share intention and subjective well-being. The research results provide a micro knowledge base for the factors that influence WeChat users’ share intention, and also deepen the understanding of subjective well-being.

**Practical implications**

For practitioners the study results provide some insights for social media (WeChat) official account operators and could help them to attract more potential and existing users. First, according to our research, utilitarian value and hedonic value are two important factors that affect the users’ share intention. In particular, the effect of utilitarian value on the users’ share intention is greater than the effect of hedonic value on the users’ share intention. Thus, for social media official account operators, the primary task is to check the utilitarian value of the information pushed. Although hedonic information also needs pushing, it cannot occupy a significant portion.

Second, social media official account operators should pay attention to the users’ satisfaction and to the credibility of the information pushed. In the process of pushing information, it is advised not to promote information, recklessly; attention should be paid to establish interaction with the users, with the prerequisite to substantial understanding of the users’ needs and preferences. It is, therefore, wise to build an interactive community and communication with the users. In addition, content should be pushed to achieve excellence and the operators should strive to achieve customer satisfaction, and create a good impression and reputation, in turn. Furthermore, the official account operators should pay
attention to cultivating the credibility of the official account. Strict controls of the push contents are needed and the operators should strive to push information that is true and reliable because this will form a good impression of credibility in the minds of the users.

Third, according to the research conclusions of this paper, WeChat official account operators should attach great care to the text with relative importance and implement marketing strategies in accordance with the users’ characteristics. For example, if the user is from the older group, then the operator should push more information about health care, and so on. For social media users, this paper appeals them to share information which on the one hand can help others get adequate knowledge and information; and, on the other hand, their subjective well-being can also increase due to sharing of information.

Finally, the ultimate goal of our study is not restricted to investigating factors affecting the user’s share intention and subjective well-being, but is to also increase awareness of maintaining a long-term cooperative relationship with social media users. Suggestion is made to enable users become medium of communication. In such a scenario, our findings will be useful to social media official account operators who wish to develop procedures that attract and retain users.

Limitations and future research
Like other empirical studies, this study has experienced some limitations that should be acknowledged. The study only addressed social media users’ share intention rather than their actual share behavior and it did not consider the effects of other factors on the social media users’ share intention, such as costs. Thus, the actual share behavior and costs factors are recommended to be considered in a future study. Although several scholars assess subjective well-being through life satisfaction (Grieve et al., 2013; Gerson et al., 2016; Ishii, 2017), subjective well-being actually includes people’s emotional responses, domain satisfactions and global judgments of life satisfaction (Diener et al., 1999). Thus, recommendation is made to consider these three parts of subjective well-being in future research.

References


About the authors

Liang Ma is PhD student of Information Systems at the Shandong University of Finance and Economics, Jinan, China. His research focuses on international marketing, information systems and electronic commerce. He has published in *Total Quality Management & Business Excellence, International Journal of Market Research, Information Discovery and Delivery, Pacific Asia Conference on Information Systems* and others. Liang Ma is the corresponding author and can be contacted at: maliang1010@126.com

Xin Zhang is Professor of Information Systems at the Shandong University of Finance and Economics, Jinan, China. His research focuses on information systems, electronic commerce and electronic markets. He has published in *Journal of Intelligent & Fuzzy Systems, Applied Mathematics and Computation, Total Quality Management & Business Excellence* and others.

Xiao Yan Ding is PhD student of Information Systems at the Shandong University of Finance and Economics, Jinan, China. His research focuses on electronic commerce and electronic markets.

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com
Factors driving continued use of online health promotion competitions

Evidence from an online weight-loss community

Xiaolong Song
Department of Management Science and Engineering,
Dongbei University of Finance and Economics, Dalian, China
Yi-Hung Liu
Department of Business Administration, Zhejiang University of Technology,
Zhejiang, China
Jiahua Jin
Department of Economics and Management,
University of Science and Technology Beijing, Beijing, China, and
Jianguo Zhao
Department of Public Administration and Policy,
Dongbei University of Finance and Economics, Beijing, China

Abstract

Purpose – Gamification elements have been increasingly used in online weight-loss communities to help users lose weight. The purpose of this paper is to systematically examine whether and how social interactions influence users' continued participation in the context of online weight-loss competitions (OWCs).

Design/methodology/approach – This study empirically investigated sustained involvement in OWCs using a Cox proportional hazards model. Additionally, the research utilized a text-mining technique to identify various types of social support and explored their roles in sustaining participation behavior in OWCs.

Findings – Community response both within and outside OWCs positively influence users’ continued participation in OWCs. Moreover, whereas emotional support and companionship received within OWCs have a greater impact on users’ continued participation than informational support received within OWCs, informational support received outside OWCs has a greater impact on users’ continued participation than emotional support and companionship received outside OWCs.

Originality/value – This paper highlights users’ social needs in OWC engagement and provides empirical evidence on how different types and sources of social support influence continued participation behavior in OWCs. The research additionally provides management implications for online health community service providers.

Keywords Social support, Community response, Continued participation, Health promotion competition, Online weight-loss community

Paper type Research paper

Introduction

Obesity is a global health concern that not only threatens human health but also reduces the quality of life and results in a heavy economic burden on healthcare systems (World Health Organization, 2000). Various methods have been adopted for combatting obesity, but the long-term results are often inadequate (Mauro et al., 2008). Recently, increasing numbers of overweight and obese people have been turning to online communities that support weight loss. These communities offer social support and knowledge exchange regarding weight management, thereby empowering people to exert greater control over their weight (Ballantine and Stephenson, 2011).
Gamification utilizes game-design elements to engage users in pursuing their targeted goals (Liu et al., 2017). With the emergence of “gamification,” game-design elements have been increasingly used in online weight-loss contexts (Lister et al., 2014). An online weight-loss competition (OWC) is a type of online weight-control activity that draws users together and encourages them to attain their weight-loss goals through a virtual competition-based system. OWCs are not merely a new type of entertainment-oriented module, but a health-enhancement tool that may influence health behavior. For OWCs, motivational affordances such as goals and challenges are incorporated in certain tasks to encourage desirable weight-control behavior.

Despite the considerable interest in applying gamification designs to nongame contexts, the use of game-design elements to help users lose weight is still in its infancy. Whether and the extent to which OWCs benefit people seeking weight reduction remain inconclusive. A challenge for OWCs is that successful weight loss typically requires substantial and continual effort (Elfhag and Rössner, 2005), but most of OWCs are “one-off” events and have a short time limit. Moreover, the use of OWCs often sharply declines over time, and they exhibit high dropout rates, thus limiting their utility in weight control (Maloney et al., 2015). Keeping users engaged with OWCs is therefore crucial for increasing their effectiveness as weight-control aids. However, strategies to incentivize users’ continued participation in OWCs have received limited scholarly attention (Johnson et al., 2016; Looyestyn et al., 2017).

Social interaction has been identified as a key incentive to continued participation in online communities (Fang and Neufeld, 2009; Zhang et al., 2013; Joyce and Kraut, 2006; Wu and Liu, 2007); however, whether the roles of social interactions also apply to OWCs remains unclear. Online communities are mostly conversation based. By contrast, gamification designs are often used to engage users toward targeted goals, for which conversation may be regarded as a lower priority. For example, a study on online role-playing games suggested that sociality and interactivity have negligible effects on consumer loyalty (Huang and Hsieh, 2011). Some social activities such as out-game discussion that takes the time of a user might even take away their interest in gaming. Community response can be seen as one type of extrinsic reward for OWC users, which may undermine their intrinsic motivation (Deci et al., 1999). The effects of extrinsic incentives could decrease over time (Magni et al., 2010). This study systemically examined whether and how social interactions influence users’ continued participation in OWCs.

Social interaction elements are widely adopted in games (Liu et al., 2017). Players who do similar tasks help each other monitor progress, set expectations, and provide encouragement. Player–player interactions can create cohesive communities (Chang et al., 2008) and a sense of relatedness that can provide social reinforcement (Liu et al., 2017). In addition to OWCs, online weight-loss communities also have other social interaction elements such as group forums and general forums. Group forums are designed for members with common interests to interact with each other whereas general forums are created for people to ask questions or share information with a wider audience. These additional resources outside of OWCs enable users to hold out-game discussions. General forums and group forums allow us to distinguish between social interactions within OWCs and social interactions outside OWCs, which may have different strengths from a structural perspective (Burt, 1987; Centola and Macy, 2007; Susarla et al., 2012). Loose and weak relationships are often viewed as a vital source for obtaining information benefits (Granovetter, 1973; Centola and Macy, 2007). Social interactions outside OWCs are thus more likely to meet users’ information needs. In the online health context, people gain social support through communicating with others. Social support has various forms, such as sharing information and expressing concern, satisfying people’s different needs. Given the different sources and forms of social support, we thus quantified and compared the influences of different sources and forms of support on continued participation in OWCs.
By using a unique data set from a large online weight-loss community, we empirically investigated sustained involvement in OWCs using a Cox proportional hazards model. We considered a user’s social interaction within and outside OWCs in our proposed model and examined their effects on continued participation in OWCs. Furthermore, we utilized a text-mining technique to identify various types of social support and explored their roles in sustaining participation behavior in OWCs.

This study contributes to the gamification and continued participation literature by examining the effects of social interactions on continued participation in OWCs. We highlight the impact of social interactions may varies depending on the sources and contents of social interactions. Our findings expand the understanding of the motivational aspects of social interactions which can assist designers in identifying the social design elements that are appropriate for OWCs.

The remainder of this paper is organized as follows: the second section introduces related studies to provide the theoretical foundations of our study and develops the hypotheses. The third section presents the data and analytical strategies and then reports the results of our empirical analysis. Finally, we conclude with a discussion of the study’s implications and limitations in the fourth section.

Literature review and hypotheses

Health-related online communities

Health-related online communities are virtual communities in which people with common health-related interests or experiences gather to interact and exchange support (Eysenbach et al., 2004). These online communities offer a range of features designed to support healthcare and are becoming an important supplement to traditional health resources. The immense potential of online health communities to improve patient-centered care has attracted significant interest among scholars and practitioners seeking to understand how users benefit from their participation in online health communities (Swan, 2009; Fichman et al., 2011). Unconstrained by space and time, users freely learn from others in online health communities and change their conceptions of illness, healthy behavior and lifestyle (Merollì et al., 2013). Various types of social support exchanged in online health communities can help users to improve their health (Yan and Tan, 2014). Self-disclosing health information in online health communities can also be used to identify potential adverse drug events (Liu and Chen, 2015) by leveraging text-mining techniques. Literature has revealed that online health communities create social value by alleviating rural–urban health disparities (Goh et al., 2016).

Although the function and value of online health communities have been increasingly recognized, the extent to which people benefit from these communities depends on their use of online health communities (Merollì et al., 2013). Thus, another stream of online health community literature has primarily focused on investigating how to use these websites. Extant research has indicated that users are more likely to connect with others with similar health concerns (Yan et al., 2015). They exchange social support that may exhibit different roles and functions. For example, for mental patients, although informational support is the most prevalent type of social support exchanged, emotional support is often more effective in improving users’ health (Yan and Tan, 2014). Certain factors have been found to affect a user’s online behavior pattern. Perceived health status affects the diversity and frequency of online health information searches (Xiao et al., 2014). Information quality, source credibility and emotional support can affect users’ information adoption decisions in online healthcare communities (Jin et al., 2016). Social and self-regulatory motives are associated with habitual online weight-loss community use (Stragier et al., 2016).

Online health communities typically support members to better engage in their health self-management process. Online health communities are thus specifically suited to chronic diseases like obesity, which require substantial self-care (Ballantine and Stephenson, 2011).
Although studies have revealed a positive association between weight-loss outcomes and online weight-loss communities (Neve et al., 2011), the effects of internet-based weight-control interventions are often modest and short lived because of the difficulty of health behavior change (Krebs et al., 2010). The success of online weight-loss communities in supporting weight control largely depends on the extent to which users access these communities in their daily lives (Taiminen, 2016). Thus, enhancing engagement is critical for internet-based weight-control interventions. Weight-control tasks are commonly viewed as laborious and boring. The application of game-design elements may increase users’ motivation (Lumsden et al., 2016). Playfulness has been found to be critical for online community users’ satisfaction and intention to continue (Chiu et al., 2011; Chiang, 2013). A recent study suggested that social comparison in online platforms is more effective at increasing physical activity than social support is (Zhang et al., 2016). OWCs have been found to increase the empowerment for and amount of weight loss (Allam et al., 2015). Because weight loss is often a long-term process that requires individuals’ persistent efforts, retaining users in OWCs is crucial for increasing their effectiveness. Thus, the motivating factors that drive users’ continued participation in OWCs must be investigated.

Continued participation theory
The antecedents of continued use of information technology (IT) differ from those of its initial adoption, because the factors that emerge after the initial use may influence subsequent decisions of whether to continue use (Karahanna et al., 1999). Numerous studies have explained the continued use of information systems (ISs) by applying expectation–confirmation theory, which states that users’ decisions regarding their continued use of an IS depend on their user experience (Karahanna et al., 1999; Bhattacharjee, 2001). The continued use decision can change over time as users gain experience with IS usage (Bhattacherjee and Premkumar, 2004). Specifically, users form their perceived value of an IS by assessing their experience with it. Their level of satisfaction with the IS then determines their continued usage behavior. Many factors that can influence users’ continued participation behavior in online communities have been identified. For instance, it has been suggested that user satisfaction and a sense of belonging have combined positive effects on continuance intention for Facebook; when focused on the quality of user experience, it has been argued that an optimal experience leads to customer loyalty to online games; and it has also been suggested that whether an online weight-loss community can provide users with an enjoyable online experience is crucial for sustained online weight-loss community use.

Studies of online communities have frequently addressed the effects of social interactions on members’ continued participation (Fang and Neufeld, 2009; Zhang et al., 2013; Joyce and Kraut, 2006). As the foundation of an online community, social interaction enhances the perceived enjoyment of IT use (Chen et al., 2016). Social interactions can also foster members’ trust in potential benefits, which encourages members to form attachments to each other (Sassenberg, 2002). Through social interactions, members obtain their desired informational and emotional benefits (Sangwan, 2005), which strengthens their intentions to continue participation (Chen, 2007). Members who post messages often expect some type of response (Ridings et al., 2002); consequently, a lack of community response can diminish member satisfaction.

In the online health context, social interactions are also an essential antecedent for continued participation in online health communities. For example, research suggests that dropout users in online health communities are often those who do not receive adequate responses from other users (Zhang and Elhadad, 2016). Both informational and emotional interactions are critical for retaining users (Zhang, 2016). Researchers believe that users obtain social support through social interactions in online health communities. This support includes informational support, emotional support and companionship (Bambina, 2007).
The types and amounts of social support received are often associated with the length of a user’s participation in online health communities. A study by examined the roles of different types of social support in patients’ online engagement, and found that emotional support had a greater effect on their continuance behavior than did informational support. However, another study suggested that, although emotional support plays a critical role in retaining users, too much informational support may cause users to leave online health communities (Wang et al., 2015). Studies have also highlighted the importance of social aspects in gamification. Social factors play a crucial role in people’s willingness to continue using exercise-related gamification services (Hamari and Koivisto, 2015). Interaction is considered a critical aspect related to the positive experience of engaging in specific online activities. For example, demonstrated that pleasurable social interaction with other users can contribute to optimal experiences with online games. Gamification with social support can result in increased physical activity and empowerment (Allam et al., 2015). Thus, the effects of social interactions on continued participation in OWCs need to be investigated.

Hypotheses development

The literature suggests that social interactions and the benefits obtained from social interactions play vital roles in motivating users’ continued participation in online communities (Lamb and Kling, 2003; Zhang et al., 2013; Zhang and Elhadad, 2016). Engagement in OWCs is fundamentally a type of social activity. Player–player interaction is a crucial aspect of the participative experience in online games (Choi and Kim, 2004; Liu and Peng, 2009; Cole and Griffiths, 2007); therefore, an optimal experience is attained only if the player experiences effective or pleasant interactions with other players (Choi and Kim, 2004). Social interaction in OWCs can create a sense of relatedness that can improve users’ engagement (Liu et al., 2017).

In addition to OWCs, online weight-loss communities typically have other social networking features (e.g. groups and forums in our study) that are beyond the boundary of OWCs. OWC users can also interact with other members within these social networking features and receive social support, which we characterized as social interaction outside OWCs. Social interaction outside OWCs may facilitate community identification (Hars and Ou, 2002), which enables users to identify themselves as part of the community and motivates them to continue participation (Chuang and Yang, 2014). These social networking features could also assist users in obtaining information to meet their information needs. These benefits help enable users to feel positive about the experience of community involvement and improve their loyalty to the community, which is related to their perceived value of and trust in the platform (De Ruyter et al., 1998). Loyalty to the community encourages members to spend more time engaged in online activities in the community. Their faith in the virtual community may further translate into user loyalty to other community features (e.g. OWCs). For example, suggested that social capital in an online gaming community can affect community trust and perceived social value, leading to users developing continuance intentions toward massively multiplayer online games. For OWC users, the perceived value of community responses gained outside OWCs can foster trust, thus providing members with greater confidence in the potential of OWCs.

Thus, we propose the following hypotheses:

H1. Community responses received within or outside OWCs will positively affect users’ continued participation.

H1a. Community responses received within OWCs will positively affect users’ continued participation.

H1b. Community responses received outside OWCs will positively affect users’ continued participation in OWCs.
Researchers believe that users in online health communities can obtain social support through social interactions, including informational support, emotional support and companionship (Bambina, 2007). When participating in OWCs, people may be confused as to the exact requirements of an OWC and how to achieve them, which causes them to seek informational support from other users. If their information demands are satisfied, they are more likely to perform effectively and have a positive experience during their engagement. Moreover, in-game chatting can provide emotional support, such as encouragement, understanding and patience. For instance, users may receive encouragement from their peers regarding maintaining involvement when they experience setbacks performing a particular activity. In addition, companionship enables users to know that they are not striving alone, thus motivating them to continue because their progress is being observed.

All the three types of social support may influence loyalty, but one or some of them may have more weight than others. Studies have documented that local ties and short ties, those that are stronger in tie strength, often have less informational advantage but greater relational advantage (Granovetter, 1973; Centola and Macy, 2007). Thus, the relational roles of social interactions within OWCs are crucial for driving continued participation in OWCs. Emotional support and companionship received within OWCs can foster a sense of social identity. The stronger sense of social identity promotes cohesion, which enhances localized conformity and reinforcement (Bikhchandani et al., 1992; Centola, 2010), thereby promoting users’ continued participation.

The nature of interpersonal influence outside OWCs may differ from that within OWCs. Granovetter (1973) suggested that loose and long ties are often seen as a source of diverse and novel information. Because long ties are weaker in the relational sense but stronger in informational sense (Granovetter, 1973; Centola and Macy, 2007), the informational role of social interactions outside OWCs may have a greater influence on users’ continued participation in OWCs than does their relational role. Therefore, we propose the following hypotheses:

- **H2.** Informational support, emotional support, and companionship received within or outside OWCs will positively affect users’ continued participation in OWCs.
- **H2a.** Emotional support and companionship received within OWCs will have a greater impact on users’ continued participation than does informational support received within OWCs.
- **H2b.** Informational support received outside OWCs will have a greater impact on users’ continued participation than does emotional support and companionship received outside OWCs.

**Method**

**Research approach**

To evaluate our research hypotheses, we examined the effects of social interaction on continued participation behavior using an empirical model. We primarily focused on whether and for how long a user participated in OWCs, which requires analysis of duration. Because the duration of continued participation is positive and the distribution of event timing is typically far from normal, linear ordinary least squares regressions are not appropriate models (Kalbfleisch and Prentice, 2002). Survival analysis has been broadly used to describe the relationship of factors of interest to the time to an event (e.g. death, divorce or disease recurrence). Survival times refer to the time to the event of interest. For example, if the event of interest is failure of smoking cessation, then the survival time can be considered as the time in years until a person starts smoking again.
The Cox proportional hazards model (David, 1972) is the most widely used statistical model for the analysis of survival data involving times to an event of interest. The hazard function gives the probability of surviving up to the specified time per time unit. The model links the survival time of an individual to a set of predictor variables of interests and provides an estimate of the treatment effect on survival from the effects of other variables. The Cox proportional hazards model has been employed in various fields. For example, used this method to explore the relationship between specialization, diversification, and rate of survival in the digital publishing industry.

Hence, we employed the Cox proportional hazards model to identify the factors that influenced the timing of the continued participation event. The mathematical form of the model is as follows:

\[ h(t, x) = h_0(t) e^{\sum_{i=1}^{n} \beta_i x_i}, \]  

where \( h(t, x) \) is the hazard rate, which is the likelihood of OWC abandonment at time \( t \); \( h_0(t) \) is the baseline hazard function with no explanatory variables; \( x_1, \ldots, x_n \) represent the explanatory variables; \( \beta_1, \ldots, \beta_n \) are coefficients to be estimated, which indicate how the explanatory variables affect the hazard rate; \( \beta_1 = 0 \) indicates that no association exists between explanatory variable \( x_i \) and the hazard rate \( h(t, x) \); \( \beta_1 < 0 \) indicates that explanatory variable \( x_i \) reduces the hazard rate \( h(t, x) \) and increases the survival time and vice versa; and \( \exp(\beta_i) \) represents the hazard ratio of the explanatory variables. In this study, the event was a user ceasing to use the OWC feature in the online weight-loss community. Correspondingly, the timing of the event was the survival time of a user’s continued engagement in OWCs, which was measured as the difference between a user’s first and last appearance in OWCs.

**Data**

The data used in the present study were taken from a large online weight-loss community designed to assist users with weight loss. The platform offers a number of features, including a general forum that enables members to interact, interest groups for members to engage with others with similar health interests, and a weight recording function. Specifically, the community provides an OWC feature called “challenges.” In a challenge, the administrator posts a task (e.g. engaging in a specific sport for 30 min every day) and specifies various parameters, such as the start time and the duration. Once the task is posted, interested participants can join in the challenge and perform the task in the real world for a fixed period and then report their weight change. Participants can observe each other’s weight change performance via leaderboards. This setting provides a suitable context for investigating continued usage of OWCs.

We collected detailed data from the focal community for a period of 281 weeks from August 2, 2008 to December 20, 2013. The data set included the online activities and user profile information during this period. Self-reported weights may be inaccurate because of data entry errors (Hwang et al., 2013). Moreover, Tell et al. (1987) demonstrated that people are likely to overstate their weight losses by 2–3 kg. To reduce inaccuracies in self-reported weight, we excluded users whose starting weight or goal weight was 0 kg, users whose goal weight was higher than their starting weight, and users for whom the difference between their starting and goal weights was less than 3 kg.

Because the data were right censored and the period of observation ended on December 20, 2013, determining whether a user intended to continue participation in OWCs was difficult if the end date of our observed period was used as the judgment time. For example, a user may disappear in the final week and return after a short period. Therefore, ensuring
that individuals had ceased appearing in OWCs over a long period was necessary for judging events. We assumed that if a user had not participated in any OWC during the final 12 weeks, the event had occurred. To avoid selection bias, we also selected 4 weeks and 20 weeks before the end date as judgment times.

To further reduce the effects of right censoring, we excluded users who first used the OWC feature during the final 20 weeks. Ultimately, we obtained a data set of 4054 challenges with 27,378 users. Figure 1 illustrates the ratio of users’ continued participation over time. The ratio declined sharply between the 5th and 15th weeks. Long-term users composed only a minimal portion of the total number of users. Less than 20 percent of users continued to engage in OWCs after the 20th week.

Measures
According to whether and when users used OWCs, we captured the survival time of users’ engagement in OWCs through the dependent variable Time of Continued Participation, which was measured as the length of the time interval from a user’s first appearance to their last appearance during the observation period. We established our independent and control variables as follows.

Independent variables. The act of receiving a response indicated a one-way communication from a commenter to a recipient. As mentioned, we considered the effects of two types of community response: community responses received in OWCs and community responses received outside OWCs. These types refer to the explanatory variables OWC Response, which captured the number of replies a user received in OWCs in the focal week, and Forum Response and Group Response, which represented the number of replies a user received in general and group forums, respectively, in the focal week.

Control variables. In online weight-loss settings, weight change reflects personal healthcare performances. We averaged a user’s reported weights over a one-month period to obtain their monthly weight. Because users have different body weights, absolute change in body weight is not appropriate for directly measuring weight-loss outcomes. To control for baseline differences in individual weight, we computed the ratio of monthly variation in weight as the indicator of the control variable Weight Change, consistent with prior weight-loss treatment.
outcome studies (Kolotkin et al., 2001; Khan et al., 2006). Different OWCs may lead participants to have divergent experiences, which consequently affect their intentions to continue participation. Thus, we included factors related to the OWC that may affect users’ continued participation intention as controls. Studies have suggested that group (community) size affects the behavior of users. For example, Zhang and Zhu (2011) determined that group size was related to social benefits, which positively influenced users’ contribution intentions. However, noted that the influence may be simultaneously positive and negative. Another study indicated that community size is also a critical factor for continuance (Zhang et al., 2013).

From our research perspective, challenges with more effective content designs are more likely to be popular and produce satisfying experiences. Hence, we adopted the control variable Challenge Size, which was measured by averaging the number of participants across all challenges that a user joined in a one-month period. OWCs also vary in duration, which affects the time a user must remain to complete the competitions. Some users may prefer remaining in an OWC until it has been completed. We used the control variable Challenge Length, which was the average length of time across all challenges, to control for this potential effect.

Table I lists the descriptive summary statistics and pairwise correlations. All the pairwise correlations between the explanatory variables and control variables were less than 0.20. We calculated the variance inflation factors (VIFs) for all such variables. The highest VIF was 1.03, which was markedly lower than 10, indicating that multicollinearity did not affect our estimation (Belsley et al., 2005; Kutner et al., 2004). Due to the large variance in predictor variables, the distribution of variables could not be normal. To reduce the skewness of the data, we performed log transformations on all the predictor variables before including them in our estimation.

Empirical analysis
Table II presents the results of the survival analysis. To avoid bias in the judgment time selection, we selected 4, 12 and 20 weeks before the end date of our observed period as the judgment time. We reported corresponding parameter estimates and hazard ratios for the three settings in Table II. A predictor variable with a negative coefficient indicated a
reduction in the hazard rate of loss, increasing the time of continued participation in OWCs and vice versa.

For the judgment time of 12 weeks, the significant negative coefficient of OWC Response (−2.248, \( p < 0.001 \)) suggests that community responses received from OWCs positively influenced continued participation behavior, indicating support for \( H1a \). Users who had more communication with others in OWCs tended to receive more social support and were thus more likely to maintain continued participation. As shown in Figure 2, the coefficients of Forum Response (−0.721, \( p < 0.001 \)) and Group Response (−0.522, \( p < 0.001 \)) were negative and significant. The significant negative effects of Forum Response and Group Response on the attrition rate confirmed the positive relationship between community response gained outside OWCs and the duration of sustained participation. Community response gained outside OWCs can incentivize members to remain in the community and increase their trust in the platform, which increases the probability of engaging in OWCs. Therefore, \( H1b \) is supported. The findings related to \( H1a \) and \( H1b \) indicate support for \( H1 \) and highlight the significant role social interaction plays in continued participation, which is consistent with previous studies (Fang and Neufeld, 2009; Zhang et al., 2013; Joyce and Kraut, 2006). Although Forum Response and Group Response had positive impacts on users’ continued participation in OWCs, the effects of Forum Response and Group Response on the hazard rate were considerably smaller than that of OWC Response.

The coefficient of Weight Change was significant and negative, suggesting that weight-loss outcomes motivated users to continue OWC participation. The results also suggested that the heterogeneity of OWCs is a critical factor that should be considered. Users were more likely to continue their participation when they joined a more popular or longer-term OWC. The estimated results for the three settings were qualitatively similar, indicating that our results were robust for each selected judgment time.

To further establish how different types of social support influence users’ engagement in OWCs, we extracted the instances of social support from online communication and examined them. A relevant study of online health communities suggested that three forms of social support exist, namely informational support, emotional support and companionship (Bambina, 2007). Whereas informational support refers to the provision of information intended to help people cope with difficulties, emotional support refers to the expression of emotional concerns. Companionship supports people by fostering a sense of social belonging.

```
| Community responses received within OWCs | \( H1a (−2.245^{***}) \) |
| Community responses received outside OWCs | \( −0.721^{***} \) |
| Forum responses | \( H1b \) |
| Group responses | \( −0.522^{***} \) |

Note: \( ^{***} p < 0.001 \)
```

Figure 2. Results of community response effects
We utilized a text-mining technique to detect types of social support. Following the tagging classification criterion of other studies (Bambina, 2007; Yan and Tan, 2014), as listed in Table III, we performed a text classification by using LingPipe. Text classification typically assigns a document to one or more categories according to their content by automated or human means. LingPipe is a natural language processing tool kit for processing text using computational linguistics (http://alias-i.com/lingpipe/index.html). The tool kit offers text classification as one of its features and has been widely used for text mining in academic research (e.g. Lu et al., 2013; Yan and Tan, 2014), thus meeting our requirements. We identified the social support type of a reply comment by comparing the probability that it belonged in a specific social support category. With the tagging classification criterion, we calculated probabilities of each category for a reply comment. The higher the probability was, the closer the match was. We then selected the category with the highest probability as the type of social support.

Table IV lists the total numbers of posts for each type of social support. Typically, people mostly exchange informational support in online health communities (Yan and Tan, 2014). However, Table IV reveals that, whereas informational support was the most frequently exchanged type of social support in public forums and groups, emotional support was the most common social support type in OWCs.

To identify the roles these supports play in sustaining participation in OWCs, we further used the numbers of posts for different types of social support from different sources as independent variables to repeat our models separately. Figure 3 and Table V reveal that all the types of social support gained within OWCs were able to increase users’

<table>
<thead>
<tr>
<th>Social support categories</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Informational support</strong></td>
<td></td>
</tr>
<tr>
<td>Advice</td>
<td>Advice is an opinion or recommendation offered as a guide to enable people to cope with their weight-related problems</td>
</tr>
<tr>
<td>Referral</td>
<td>Referrals provide the recipient with a source of experts or resources</td>
</tr>
<tr>
<td>Teaching</td>
<td>Teaching provides factual information about weight-related issues or the skills needed to deal with these issues</td>
</tr>
<tr>
<td>Information broadcasting</td>
<td>Information broadcasting is a non-targeted communication activity that requires no reaction from others</td>
</tr>
<tr>
<td>Personal experience</td>
<td>Personal experience refers to users’ own experiences regarding their weight control</td>
</tr>
<tr>
<td><strong>Emotional support</strong></td>
<td></td>
</tr>
<tr>
<td>Understanding/empathy</td>
<td>Understanding/empathy indicates a sympathetic or empathic awareness of the recipient’s feelings</td>
</tr>
<tr>
<td>Encouragement</td>
<td>Encouragements provide recipients with hope for and confidence in achieving the goal of weight reduction</td>
</tr>
<tr>
<td>Affirmation/validation</td>
<td>Affirmation/validation refers emotional support that expresses agreement with the recipient’s actions or opinions on the situation</td>
</tr>
<tr>
<td>Sympathy</td>
<td>Sympathy acknowledges the recipient’s emotional hardships associated with weight problems and provides comfort and support</td>
</tr>
<tr>
<td>Caring/concern</td>
<td>Caring/concern refers to sensitivity to the health and needs of the recipient</td>
</tr>
<tr>
<td><strong>Companionship</strong></td>
<td></td>
</tr>
<tr>
<td>Chatting</td>
<td>Chatting represents interactive, spontaneous communication among users</td>
</tr>
<tr>
<td>Humor/teasing</td>
<td>Humor/teasing makes people laugh and feel comfortable</td>
</tr>
<tr>
<td>Groupness</td>
<td>Groupness is a feeling of common identity that helps individuals to ease their loneliness and helplessness</td>
</tr>
<tr>
<td>Buddy/friend seeking</td>
<td>Buddy/friend seeking is the tendency to form social ties with others</td>
</tr>
<tr>
<td>Self-introduction</td>
<td>Self-introduction represents individual identity to help others to know him or her</td>
</tr>
</tbody>
</table>

Table III. Classification criteria for tagging
continued participation time. The magnitude of the impact of emotional support and companionship received within OWCs was stronger than that of informational support received within OWCs. The results suggest that for social interactions within OWCs, emotional support and companionship are more important in driving continued participation. H2a is thus supported.

By contrast, for social interactions outside OWCs, the coefficients of companionship were not significant, suggesting that only companionship provided within OWCs can facilitate continued participation in OWCs. The coefficients of informational and emotional support

<table>
<thead>
<tr>
<th>Social support categories</th>
<th>Number of posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenge Info</td>
<td>8,698</td>
</tr>
<tr>
<td>Challenge Emo</td>
<td>11,837</td>
</tr>
<tr>
<td>Challenge Comp</td>
<td>1,825</td>
</tr>
<tr>
<td>Forum Info</td>
<td>20,971</td>
</tr>
<tr>
<td>Forum Emo</td>
<td>12,327</td>
</tr>
<tr>
<td>Forum Comp</td>
<td>2,960</td>
</tr>
<tr>
<td>Group Info</td>
<td>14,897</td>
</tr>
<tr>
<td>Group Emo</td>
<td>13,919</td>
</tr>
<tr>
<td>Group Comp</td>
<td>1,928</td>
</tr>
</tbody>
</table>

Table IV. Numbers of posts for each type of social support

Notes: *p<0.05; **p<0.01; ***p<0.001

Figure 3. Results of the effects of different types of social support from different sources

Online health promotion competitions

813
received outside OWCs were negative and significant. Therefore, H2 is partially supported. The magnitude of the impact of informational support was greater, consistent with H2b. Informational support received outside OWCs to facilitate continued participation was more important than emotional support and companionship received outside OWCs. In addition, social support gained within OWCs played a greater role than social support gained outside OWCs in promoting sustained participation in OWCs.

Discussion and conclusion
This study examined the influence of social interactions on users’ continued participation in OWCs. Moreover, the study went beyond community response and explored how different types and sources of social support influence continued participation behavior. Specifically, our survival analysis revealed that both community response within and outside OWCs positively influence users’ continued participation in OWCs. The type and source of social support were also critical factors affecting the relationship between social interactions and continued participation behavior. Several theoretical and practical implications are worthy of discussion.

First, our study contributes to the gamification literature and continued participation literature by investigating how social interactions keep users engaged in OWCs. Although prior research has focused on the benefits of gamified systems and sustained participation in an online community, the continued use of gamified systems has received scant attention until now (Liu et al., 2017), especially in an online health context. Our research seeks to address this gap by quantifying the impact of different social support on continued participation. Our results provided evidence that social interaction can encourage users to continue use of OWCs. Moreover, we observed stark differences in the effects of types of social support gained within and outside OWCs on continued participation behavior. Whereas emotional support and companionship received within OWCs have a greater impact on users’ continued participation than informational support received within OWCs, informational support received outside OWCs has a greater impact on users’ continued participation than emotional support and companionship received outside OWCs. These findings emphasize the effectiveness of social support varies depending on its content and sources.

Second, this paper highlights users’ social needs in OWC engagement. Our research supports that users are social actors rather than individualistic entities (Lamb and Kling, 2003; Zhang et al., 2013). Users can benefit from social interactions by not only

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter estimate (4)</th>
<th>Hazard ratio (4)</th>
<th>Parameter estimate (12)</th>
<th>Hazard ratio (12)</th>
<th>Parameter estimate (20)</th>
<th>Hazard ratio (20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenge Info</td>
<td>−1.690***</td>
<td>0.185</td>
<td>−1.663***</td>
<td>0.190</td>
<td>−1.616***</td>
<td>0.199</td>
</tr>
<tr>
<td>Challenge Emo</td>
<td>−2.508***</td>
<td>0.081</td>
<td>−2.531***</td>
<td>0.080</td>
<td>−2.577***</td>
<td>0.076</td>
</tr>
<tr>
<td>Challenge Comp</td>
<td>−2.185*</td>
<td>0.113</td>
<td>−2.073*</td>
<td>0.126</td>
<td>−2.038*</td>
<td>0.131</td>
</tr>
<tr>
<td>Forum Info</td>
<td>−0.551***</td>
<td>0.576</td>
<td>−0.533***</td>
<td>0.586</td>
<td>−0.538***</td>
<td>0.586</td>
</tr>
<tr>
<td>Forum Emo</td>
<td>−0.521**</td>
<td>0.594</td>
<td>−0.523**</td>
<td>0.593</td>
<td>−0.522**</td>
<td>0.593</td>
</tr>
<tr>
<td>Forum Comp</td>
<td>−0.474</td>
<td>0.623</td>
<td>−0.455</td>
<td>0.635</td>
<td>−0.492</td>
<td>0.612</td>
</tr>
<tr>
<td>Group Info</td>
<td>−0.495***</td>
<td>0.610</td>
<td>−0.536***</td>
<td>0.585</td>
<td>−0.528***</td>
<td>0.590</td>
</tr>
<tr>
<td>Group Emo</td>
<td>−0.385**</td>
<td>0.681</td>
<td>−0.359**</td>
<td>0.698</td>
<td>−0.332**</td>
<td>0.717</td>
</tr>
<tr>
<td>Group Comp</td>
<td>0.243</td>
<td>0.784</td>
<td>0.207</td>
<td>0.812</td>
<td>0.191</td>
<td>0.826</td>
</tr>
<tr>
<td>Weight Change</td>
<td>−23.500***</td>
<td>0.000</td>
<td>−23.540***</td>
<td>0.000</td>
<td>−23.920***</td>
<td>0.000</td>
</tr>
<tr>
<td>Challenge Size</td>
<td>−0.019*</td>
<td>0.981</td>
<td>−0.041***</td>
<td>0.960</td>
<td>−0.048***</td>
<td>0.953</td>
</tr>
<tr>
<td>Challenge Length</td>
<td>−1.694***</td>
<td>0.184</td>
<td>−1.735***</td>
<td>0.173</td>
<td>−1.761***</td>
<td>0.172</td>
</tr>
</tbody>
</table>

Note: Significant at codes: *p < 0.05; **p < 0.01; ***p < 0.001
receiving social support but also enhancing their continued participation behavior in online healthcare activities, which is consistent with the results of previous studies (Fang and Neufeld, 2009; Zhang et al., 2013; Joyce and Kraut, 2006). We confirmed that, in addition to the community responses received in OWCs, the community responses received outside OWCs also influences continued participation. Our results suggest that users experience different levels of arousal when receiving different types of social support from different sources. While users exposed to emotional support and companionship within OWCs are more likely to be at a lower risk of dropout than those exposed to informational support within OWCs, informational support outside OWCs shows more impact than other support outside OWCs.

Finally, the present study provides crucial management implications for online health community service providers. Although gamification design elements are increasingly being used to encourage competition between users and improve their performance, gamification designs that focus solely on competitive elements (e.g. badges and leaderboards) cannot accomplish the expected behavioral changes (Fogel, 2015). A practical challenge facing service providers is deciding which strategies to employ to keep users in OWCs. Our study establishes a relationship between social interaction elements and continued participation behavior in OWCs. The survival analysis suggested that social interaction elements provide an efficient means to retain users in OWCs. More interestingly, our results revealed that the effects of various forms of social interactions are heterogeneous. These findings highlight how designers can make use of social interaction elements to improve users’ engagement in OWCs. For example, because people care more about the presence of other users and emotional support received within OWCs, platform practitioners can consider integration of more social interaction elements, which can support the relational roles of social interactions, such as the use of kudos and virtual gifts, when designing their gamified systems.

Despite these implications, our study had several noteworthy limitations. First, our data consisted of different types of observable online digital footprints. However, the real world contains unobserved individual heterogeneity, which may cause users to stop using an OWC feature. For instance, users may join a local health club or begin following a diet plan from a health consultant, and thus abandon their online engagement. Second, we used the change in the self-reported weight to measure the weight-loss outcome in our empirical test. However, research has suggested that competitive pressure may increase false reporting (Schwieren and Weichselbaumer, 2010). People who feel their performance was poor may report false results that exaggerate their weight-loss achievements as a “face-saving strategy,” causing our data to not reflect actual weight change. Future research should examine whether the effects of healthcare performance evolve over time. Third, we only considered online social interaction within an online weight-loss community. However, users may also have offline interactions, which may be another type of community response received outside OWCs. Future studies should conduct an online survey to obtain offline social interaction data, as suggested by Chung (2013), to examine the effects of offline interaction.

Acknowledgments
The authors would like to thank the Editor-in-Chief and anonymous reviewers of Online Information Review for their help in improving this paper. This research was supported in part by the Research Funds of the National Natural Science Foundation of China (Grant Nos 71601034, 71401047, 71601100), China Postdoctoral Science Foundation (Grant No. 2016M600210), General Program of Humanities and Social Sciences of Liaoning Provincial Education Department (Grant No. LN2016YB007), the MOE Project of Key Research Institute of Humanities and Social Sciences at Universities of China (Grant No. 15JD790002) and DUFF Postdoctoral Science Foundation funded project (Grant No. BSH201715).
References


Further reading


**Corresponding author**
Xiaolong Song can be contacted at: xlsong@dufe.edu.cn
Topical evolution patterns and temporal trends of microblogs on public health emergencies

An exploratory study of Ebola on Twitter and Weibo

Lu An
Center for Studies of Information Resources, Wuhan University, Wuhan, China and School of Information Management, Wuhan University, Wuhan, China

Chuanming Yu
School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan, China

Xia Lin
College of Computing and Informatics, Drexel University, Philadelphia, Pennsylvania, USA, and Tingyao Du, Liqin Zhou and Gang Li
School of Information Management, Wuhan University, Wuhan, China

Abstract
Purpose – The purpose of this paper is to identify salient topic categories and outline their evolution patterns and temporal trends in microblogs on a public health emergency across different stages. Comparisons were also examined to reveal the similarities and differences between those patterns and trends on microblog platforms of different languages and from different nations.

Design/methodology/approach – A total of 459,266 microblog entries about the Ebola outbreak in West Africa in 2014 on Twitter and Weibo were collected for nine months after the inception of the outbreak. Topics were detected by the latent Dirichlet allocation model and classified into several categories. The daily tweets were analyzed with the self-organizing map technique and labeled with the most salient topics. The investigated time span was divided into three stages, and the most salient topic categories were identified for each stage.

Findings – In total, 14 salient topic categories were identified in microblogs about the Ebola outbreak and were summarized as increasing, decreasing, fluctuating or ephemeral types. The topical evolution patterns of microblogs and temporal trends for topic categories vary on different microblog platforms. Twitter users were keen on the dynamics of the Ebola outbreak, such as status description, secondary events and so forth, while Weibo users focused on background knowledge of Ebola and precautions.

Originality/value – This study revealed evolution patterns and temporal trends of microblog topics on a public health emergency. The findings can help administrators of public health emergencies and microblog communities work together to better satisfy information needs and physical demands by the public when public health emergencies are in progress.

Keywords Twitter, Microblog, Weibo, Public health emergency, Temporal trends, Topical evolution pattern

Paper type Research paper

Introduction
Public health emergencies are important types of events that attract high attention. As public health emergencies usually occur unexpectedly and deteriorate rapidly, emergency management departments need to quickly and fully grasp the development status of emergencies, predict potential problems and take effective measures. Currently, microblog
platforms, such as Twitter and Weibo, are popular channels for obtaining and generating a great deal of information and comments on various events, such as public health events. According to Reuter and Spielhofer (2017), 75 percent of citizens used or plan to use social media to share information during an emergency, and 77 percent of citizens believe that social media provides emergency information faster than traditional media channels. Microblogs are an important type of social media and provide a wealth of clues to public emergency management. In September 2013, Twitter Alert, an emergency alert system was released, and the American Red Cross, Federal Emergency Management Agency, and global non-profits such as the World Health Organization have participated in this system (Mashable, 2013). The topical analysis of microblogs on public health emergencies can reveal the development track of emergencies and, thus, facilitate understanding the public’s concerns about the emergencies. As a microblog platform often has a fixed range of users, e.g., from the same or similar countries or of the same language, the analysis of microblogs on one platform may produce limited findings. A comparison between two different microblog platforms may reveal different topical evolution patterns or temporal trends for the same emergency and help emergency management departments obtain insights and identify problems that cannot be found upon the analysis of one microblog platform.

Since February 2014, a severe Ebola outbreak has raged in West Africa and has led to the death of more than 11,000 people as of March 16, 2016 (World Health Organization, 2016). Hundreds of thousands of tweets on the Ebola outbreak have been generated on microblog platforms since then. Tens of thousands of tweets and internet searches have been generated upon each news video of Ebola (Towers et al., 2015). To explore temporal characteristics of topic categories in related microblogs, we took the Ebola outbreak in West Africa since February 2014 as the investigation case.

The purpose of this study was, first, to detect salient topic categories in microblogs on a major public health event at different stages, second, to outline the evolutionary pattern and temporal trends of topic categories in the related microblogs, and third, to reveal and compare usage patterns of Chinese and English microblog platforms when a major public health event unfolds. The findings can be used to understand temporal characteristics of topic categories in microblogs of public health emergencies and usage differences between Chinese and English microblog platforms. Accordingly, public health emergency administrators can provide informational and physical support to mitigate the consequences of infectious diseases in similar cases, such as the current Zika virus.

Related research
Topical analysis of microblogs
Common methods of topical analysis of microblogs include statistical techniques and manual classification. Studies on microblogs involving general topics, e.g., Yun (2012), were usually based on fairly rough topic classifications. More fine-grained topic categories were identified in the topical analysis of microblogs focusing on a specific topic, such as classifying main topics of public concern toward Ebola into several categories (Lazard et al., 2015). However, manual classification of topics limited the number of investigated microblogs.

As major events usually trigger a large number of tweets, researchers generally sample a small percentage of related tweets (Qu et al., 2011). To automatically analyze large-scale microblog entries, topic modeling is an indispensable step. Latent Dirichlet allocation (LDA) and term frequency vectors are common approaches for topic modeling. Topics of microblogs were often studied in a static manner, and the evolution of topics was not sufficiently explored.

Temporal analysis of microblogs
To reveal the temporal characteristics of microblogs, some novel topic models were developed, such as trend sensitive-latent Dirichlet allocation (TS-LDA) (Yang and Rim, 2014)
and the Topics Over Time model (Wang and McCallum, 2006). Various topic models were employed to detect emerging topics in microblogs (Vosecky et al., 2014), to analyze news and reports on public health events, e.g., SARS (Wilson et al., 2008), and to examine daily quantities of tweets on Hurricane Sandy (Wang et al., 2015).

Extant case studies have not sufficiently explored the evolution of topics in microblogs when major public health events unfold; instead, these case studies usually investigate only one microblog platform of a single language or nation. Comparisons between different microblog platforms are rare.

Visual analysis of microblogs
Due to huge volumes of microblogs, researchers have developed visual analytical tools to intuitively analyze topics or events and their evolution in microblogs, such as WeiboEvents (Ren et al., 2014) and SocialHelix (Cao et al., 2015).

Some typical information visualization techniques, such as the self-organizing map (SOM) and Treemap techniques, were also employed to explore microblogs. The SOM technique can visualize high-dimensional input data in a low-dimensional space while preserving the topology of input data (Kohonen, 2001). This technique can apply to general data distributions with little a priori knowledge required (Zhang and Li, 1993) and can help users understand the comprehensive data structure (Raubar et al., 2002). The SOM technique provides a convenient way to visualize topical analysis. The U-matrix is a common SOM display in which the value of each unit equals the sum of the Euclidean distances of an SOM node to all its immediate neighboring nodes normalized by the largest occurring value in the SOM grid. Due to its many advantages, the SOM technique has been widely used in many fields (An et al., 2011, 2014; Zhang et al., 2009). However, few studies utilized the SOM technique to analyze the topical evolution of microblogs, especially microblogs on public emergencies.

Treemap was used to generate group-in-a-box meta-layouts to analyze microblogs (Chaturvedi et al., 2014). However, the display was mainly used to analyze community members and relationships among communities, instead of topics of microblogs.

It seems that most related visual studies aim to develop visual tools to analyze topics or events in microblogs, such as WeiboEvents and SocialHelix. The analytical functions of those tools involve retweets paths, spatiotemporal characteristics, and detection and tracking of events or topics. Little has been done to visually analyze evolution patterns of topics on public emergencies in microblogs, by which much invaluable knowledge may be found useful.

Social media analysis of public emergencies
As social media provide rich information about public emergencies, the former is often associated with the latter. First, social media can be used as the data source of detecting public emergencies, such as emergent disease surveillance. Typical examples include Google Flu Trends, the real-time allergy surveillance system (Lee et al., 2015) and surveillance of the Ebola outbreak by Twitter (Yom-Tov, 2015).

Second, social media can be leveraged to assist public emergency management (Palen and Anderson, 2016). To this end, Wu et al. (2013) developed a collaborative sense-making system in which social media is one of the many information sources used to support team work in emergency management. Such synthetic systems are very helpful, and their usability is based on effective topical analysis of various information sources, including social media data on public emergencies. To do so, Stowe et al. (2016) developed feature-rich classifiers to categorize disaster-related tweets into seven classes and summarized the topical characteristics of each class. However, the evolution patterns and development trends of those classes are unknown as the emergency evolves. Sidana et al. (2016) proposed a new model, named Temporal Ailment Topic Aspect (TM-ATAM), to identify transitions of health-related topics in tweets. The topical transitions they found (e.g. from smoking to...
respiratory diseases) aggregated at the ailment level. Thus, the findings are more useful for public health research and practices than for public emergency management. The latter needs to summarize more generalized categories from the perspective of public emergency management (e.g. emergency status description, or emergency response action) and reveal topical evolution patterns and temporal trends of those categories.

In summary, the topical analysis of microblogs can reveal the development track of public emergencies and the public’s attention toward them. However, this task was often performed in a static manner, and the evolution of topics was not sufficiently explored when major public health events unfold. Previous case studies usually investigated only one microblog platform of a single language or nation. Comparisons between different microblog platforms in different nations or languages are seldom conducted. Information visualization techniques can intuitively facilitate revealing the topical evolution patterns of microblogs on public health emergencies. The topical analysis of microblogs on public health emergencies is more useful for emergency management when aggregated at the class level, such as emergency status description, or emergency response action, than at the symptom or disease level, such as smoking and respiratory diseases.

Data and methods
As Burkholder and Toole (1995) described in their three-phase evolution model of complex disasters, problems and priorities of emergencies shift over time in different phases, i.e., the acute emergency phase, the late emergency phase and the post emergency phase. Inspired by the three-phase evolution model, we wondered if the microblog users have different concerns toward the public health emergencies during different phases. If it is true, which topic categories are salient in each phase and how do they change across different phases? How similar or different are the users of different microblog platforms in terms of the topical evolution patterns and temporal trends of the tweets? These questions can be answered with topical analysis of tweets on a public health emergency on different microblog platforms.

Data set
This study adopted a purposive sampling strategy. Twitter.com and weibo.com were chosen as the data source since the two websites are well-known microblog platforms in the USA and China, respectively. As of March 2016, Twitter and Weibo had more than 310m and 261m monthly active users, respectively (Twitter, 2016; Sina Technology, 2016). The 2014 West African Ebola outbreak was taken as the investigated case in this study since it has been one of the most prevalent and major public health emergencies in recent years. As the searches for “Ebola” or “埃博拉” (the Chinese translation of Ebola) on Twitter and Weibo, respectively, found that related tweets first appeared in February 2014 and the monthly aggregate Ebola cases in the West Africa peaked in October 2014 (West African Ebola virus epidemic, 2016), the time span between February 1, 2014 and October 31, 2014 was chosen in this study. The tweets that contained the term Ebola or “埃博拉” during this period were collected with Metaseeker. The contents, posting dates and retweet counts for each tweet were recorded for further analysis. Since the tweet contents included both original and retweet contents, the retweet counts were not processed.

Words and phrases were extracted from microblogs using word frequency statistical tools. Stop words, which usually referred to the most common words in a language and have no specific meaning, such as the, is, at, which, and so forth, were excluded.

Detection of topics in tweets
The LDA model was employed to detect topics in tweets on Ebola on two platforms. The LDA model is a generative topic model presented by Blei et al. (2003), in which a document is
believed to comprise a certain number of topics and each word in the document belongs to a topic with a certain probability. To infer the topics in the documents and the probabilities with which each topic involves a certain word, Gibbs sampling and expectation propagation are often used. Finally, a number of topics were generated, each of which contained some related words.

Training and visualization of daily tweets
To reveal topical characteristics of daily tweets, the tweets posted on the same day and platform were merged and labeled with the corresponding dates. For example, 2–1 represents February 1. The SOM technique was employed to train and visualize the daily tweets. An SOM input matrix \( M_1 \) was constructed for the tweets from Twitter and Weibo, respectively, as shown in Equation (1). \( M_1 \) is a day-term matrix, of which each row represents a date (an object) and each column represents a term in the tweets (an attribute), e.g. medical:

\[
M_1 = \begin{pmatrix}
d_{11} & d_{12} & \ldots & d_{1n} \\
d_{21} & d_{22} & \ldots & d_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
d_{p1} & d_{p2} & \ldots & d_{pn}
\end{pmatrix}
\]  

(1)

where \( p \) is the number of days in the data set of tweets, \( n \) is the number of terms and \( d_{kj} \) \((k = 1, 2, \ldots, p; j = 1, 2, \ldots, n)\) is defined as the occurrence of term \( j \) in the tweets of day \( k \).

For each data set (the Twitter or Weibo data set), matrix \( M_1 \) was trained by the SOM algorithm. After training, an SOM display was generated and composed of a number of SOM nodes. Each object (a date) was projected onto an SOM node. According to the principles of the SOM technique, objects with similar attributes were projected onto the same or adjacent SOM nodes, while those with quite different attributes were projected onto the SOM nodes in the far distance. Thus, the dates that were projected onto the same or neighboring SOM nodes were believed to share similar terms and those that were projected onto the SOM nodes in far distance were believed to have quite different terms. As the U-matrix can reveal the clustering structure of the objects, which is explained in the section “Visual analysis of microblogs,” the U-matrix was calculated and applied to the background color of the SOM nodes. The dates that were projected onto the SOM nodes with low U-matrix values were believed to share similar terms and vice versa.

Determination of the most salient topics in daily tweets
As the number of terms is usually very large and many terms are actually related, it is difficult and unnecessary for researchers to observe all the terms in the tweets even in one day. Thus, we need to determine the most salient topic in daily tweets and label each SOM node with the most salient topic in the tweets posted on the corresponding dates. Then, researchers can intuitively observe the most salient topic in those days’ tweets. To this end, a novel aggregation and competitive SOM node-labeling strategy is proposed in which each SOM node is labeled with the most salient topic in the tweets posted on the corresponding dates. The procedures are as follows.

Suppose \( k \) topics are detected in the tweets by the LDA model, as described in the section of “Detection of topics in tweets.” Each topic is composed of a certain number of terms, denoted by term\(_{r1}, \text{term}\(_{r2}, \ldots, \text{term}\(_{rn}. It is known that each SOM node is associated with a weight vector. The weight vector elements of the \( i \)th SOM node can be denoted by \( w_{i1}, w_{i2}, \ldots, w_{in} \), where \( n \) is the number of attributes (terms). To find the most salient topic in the tweets posted on the corresponding dates, weight vector elements of each SOM node can be
aggregated according to the topics detected by the LDA technique. Suppose topic \( j \) \((j = 1, 2, ..., k)\) is composed of term \( r_1 \), term \( r_2 \), ..., term \( r_s \). See Equation (2) for the calculation method of aggregated weight vector elements for the \( i \)th SOM node, where \( t = r_1, r_2, ..., r_s \):

\[
    u_{ij} = \sum_{t=r_1}^{r_s} w_{jt}
\]

Find the maximum value among \( u_{i1}, u_{i2}, ..., u_{ik} \). For example, \( u_{ijk} \) is the maximum value. The topic \( j_v \) will be used to label the \( i \)th SOM node. All the SOM nodes can be labeled with the most salient topics in this way.

**Analysis method**

The SOM display labeled with dates and the one labeled with the most salient topics can be compared to identify the most salient topic in tweets posted on each date. The topics and terms that they contain were thoroughly examined and divided into several categories by three researchers in an iterative process, which will be described in detail later.

The dates can be further aggregated at the month level, and the temporal trends of each topic category can be summarized. To recognize the topical evolution pattern of the tweets on the public health emergency, the investigated time span was divided into three stages and the most salient topical categories were determined for each phase. The results of Twitter and Weibo were compared to identify the similarities and differences between Chinese and English microblog platforms. Some suggestions regarding public health emergency management are also provided based on the comparison results.

**Results analysis and discussion**

**Data description and preprocessing**

There are two data sets in this study. The first data set contained 228,992 tweets from Twitter covering 271 days, and the second data set contained 230,274 microblogs from Weibo covering 246 days; there are 25 days when no microblog entry on Ebola was found on Weibo.

To improve processing efficiency, the top 4,000 words and phrases with high frequencies were extracted from the two data sets and used to construct two day-term matrices \( M_1 \).

**Analysis of topics in English Ebola tweets**

In this study, the SOM Toolbox in the MATLAB environment is employed to execute the SOM training process. The input matrix \( M_1 \) for Twitter was constructed, which was composed of 271 rows and 4,000 columns. Notice that different terms may have different occurring frequencies in the investigated tweets. For example, the frequency of outbreak varied from 0 to 734, while the frequency of actor varied from 0 to 15. Thus, different attributes (columns) in the input matrix \( M_1 \) may have different value ranges. To prevent the attributes with large value ranges dominating the SOM display, the input matrix \( M_1 \) was normalized by the “var” method, in which variances of attributes are linearly normalized to 1.

The toroid space was adopted in the SOM display to avoid “border effect” (Kohonen, 2001).

The linear initiation and batch learning algorithms are employed to train the matrix \( M_1 \) since related studies (An et al., 2014) have shown that the smallest final quantization error was achieved by this combination among all of the combinations of random/linear initiation and sequential/batch learning. The background color of the SOM display was determined by the U-matrix. See Figure 1 for the results. The color bar on the right indicates U-matrix values for each color. The SOM display is labeled with dates. For example, 2–1 represents February 1.

According to the SOM principle, the dates that are projected onto the same or adjacent SOM nodes or those with low U-matrix values tend to have similar terms. Note that the SOM
display adopts the toroid space. Thus, the “lower” and “upper” borders and “left” and “right” borders are actually connected. It is seen in Figure 1 that temporally close dates tend to be projected onto SOM nodes in close vicinity. This finding means that tweets within a certain time span, e.g., a month or consecutive days, tend to have focal topics. As time went by, the focal topics changed.

To explore topical characteristics of Ebola tweets, the LDA model was employed and executed by an open source Java application JGibbLDA. A pilot study showed that the perplexity value for 50 topics was lower than that for 10 to 40 topics. For the sake of efficiency, the number of 50 was chosen to execute the LDA algorithm, and the top 20 terms with high probability values were considered for each topic. The Dirichlet prior $\alpha$ was set to 0.5, and $\beta$ was set to 0.1. The learning process was iterated 1,000 times.

As explained previously, weight vector elements of each SOM node were aggregated according to topics detected by the LDA technique, and each SOM node was labeled with the most salient topic. Namely, the $i$th SOM node was labeled with the topic with the maximum value of $u_{ij}$ as shown in Equation (2). See Figure 2 for the results.
Comparing Figures 1 and 2, we can summarize each topic and its dominance periods in Table AI. For example, topic 19 was the most salient topic on March 20 and 21. It was found that 21 topics (of 50) dominated English Ebola tweets for at least one day. Among them, dominance of topic 19 was the shortest, which lasted only two days, while topic 1 dominated Ebola tweets for the longest period, i.e., 50 days.

All of the topics and terms that they contained were thoroughly analyzed with examination of corresponding tweets and divided into several categories by three researchers in an iterative process. First, each researcher was asked to annotate the topics with a multi-label scheme and to classify the topics into several categories by themselves. Second, they discussed the summaries of the topics, categories and classification results together. The two steps were repeated several times until they finally reached agreement.

Among the produced topic categories, "status description" and "statistics and description" need to be clarified. Status description refers to a qualitative description about the Ebola outbreak, such as the Ebola virus reaching a new area, while statistics and description refers to a quantitative description about the Ebola outbreak, such as the death toll or the number of infected people in an area. The categories were also compared with those by Stowe et al. (2016) to find that five of their categories (i.e. sentiment, action, preparation, reporting and information) corresponded to ours (i.e. emotions of the public,
actions of the public/organizations, precautions and response, statistics and description, and background knowledge of Ebola) and some of our categories were more fine-grained. For example, we differentiate “actions of the public” from “actions of organizations.”

A summary of the topics is also listed in Table A1. Figure 3 shows the 11 categories that English Ebola tweets involved and the number of days when each topic category dominated the tweets in each month. The vertical axis indicated the number of days when a topic category was the most salient. If a topic category is invisible in a month, as in Figure 3, it does not mean that the category did not occur in microblogs in the corresponding month. The category may still have existed in the periods but was not the most salient.

Figure 3 clearly reveals temporal development of topics in Ebola tweets. At the beginning of the Ebola outbreak in West Africa in 2014, people were curious about the possible cause of Ebola. Many of them pointed to the immigration problem and predicted consequences of population reduction. A few days later, attention was attracted to some irrational actions of the public, which was also a secondary event. Tweets on various secondary events existed during most of the time span and tended to intensify as time went by. In the second month, some organizations, such as WHO, NIH and governments of some West African nations, realized the necessity and importance of taking appropriate measures to counter Ebola, including long-term measures, such as research on Ebola, and short-term measures, such as shutting borders and closing schools. With the spread and raging of Ebola, tweets of statistics and status description also increased, e.g., death toll and new areas infected with Ebola. Although it only lasted for a short time, other simultaneous important news such as ISIS surges and the Ukraine cease fire were also mentioned together with Ebola. In the middle of the investigated time span, people focused on evaluating the Ebola outbreak in West Africa in 2014 as among the “most challenging” ever. At the final phase, people’s anger and hatred toward Ebola peaked and Topic 41, even comprising profanity, dominated the whole month of October 2014, which reflected that many people lost their temper and felt depressed.

Among all of the topic categories, status description and secondary events were the most salient while evaluation and external environment were the least salient. The findings were

![Figure 3. Dominance durations of the most salient topic categories in English tweets in each month](image-url)
similar to those by Alicino et al. (2015). The latter found that top Ebola-related Google queries included Ebola news, Ebola updates, about Ebola, Liberia Ebola and so on. This finding meant that the public was concerned with the status and secondary events of Ebola.

To reveal the topical evolution pattern of Ebola tweets, the investigated time span was divided into three stages; i.e., from February to April, May to July and August to October. At each stage, the days were added up when each topic category was the most salient one in each month of this stage (see Figure 3). For example, at the first stage, the topic category of prediction was the most salient for 19, 18 and 4 days in February, March and April, respectively. The total sum of the days was 41 days. For each stage, the topic category with the largest sum of days was the most salient one at this stage. It was found that prediction and possible cause of Ebola were the most salient topic categories at the first stage, with statistics and description and secondary events the most salient at the second stage and status description and emotions of the public the most salient at the third stage. See Figure 4 for an illustration of the topical evolution of Ebola tweets.

Please be aware that, at each stage, only the top two categories are shown. Although the secondary events category was more salient at the third stage than at the second stage, it was not as salient as status description and emotions of the public at the third stage. Thus, it is not present at the third stage in Figure 4.

To reveal the temporal trends of individual topic categories, for each category, the days when the category in question was the most salient in each month were counted, which is shown in Figure 3. For example, the topic category of emotions of the public was the most salient for one day in September and 30 days in October. Thus, it was an increasing topic category. It was found that precautions and response, statistics and description and emotions of the public kept increasing. This finding partly coincided with the findings by Alicino et al. (2015), who also found that Ebola cure, Ebola nurse (related to precautions and response), Ebola patient, Ebola cases and Ebola latest news (related to statistics and description) were among the Ebola-related rising Google queries.

Possible cause of Ebola, actions of the public, prediction and evaluation tended to decrease. Secondary events, status description, external environment and actions of organizations fluctuated, and had peaks in August, September, July and August 2014, respectively. See Table I for the temporal trends of topics in Ebola tweets. The findings provided useful clues for decisions made by public health emergency management departments. At the early stage, the departments need to provide sufficient and accurate information about causes and reasonable consequences of infectious diseases. With the

---

**Figure 4.**
Topical evolution of English Ebola tweets

**Table I.**
Temporal trends of topic categories in English Ebola tweets

<table>
<thead>
<tr>
<th>Temporal trends</th>
<th>Topic categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing</td>
<td>Emotions of the public, statistics and description, and precautions and response</td>
</tr>
<tr>
<td>Decreasing</td>
<td>Prediction, actions of the public, possible cause of Ebola and evaluation</td>
</tr>
<tr>
<td>Fluctuating</td>
<td>Secondary events, status description, external environment and actions of organizations</td>
</tr>
</tbody>
</table>
spread of the diseases, the departments need to better meet people’s increasing information needs about precautions and responses, statistics, descriptions, etc. At the later stage, the departments need to calm the high emotions of the public. Temporal trends of topic categories can also provide guidance for public health emergency management departments in allocating attention and efforts to different topic categories. For increasing topics, such as emotions of the public and statistics and description, public health emergency management departments need to be highly alert and make long-term plans to alleviate the consequences of public health emergencies. For decreasing topics, such as prediction and possible cause of Ebola, the management departments can adopt appropriate guidance and corrective measures to eliminate misleading information. However, excessive worries or attention are not necessary since those topics tend to decrease. For fluctuating topics, such as secondary events and status description, more analyses and explorations are to be conducted to reveal the reasons and conditions of topical fluctuation.

Analysis of topics in Chinese Ebola microblogs

The same process was executed on Chinese Ebola microblogs. The SOM display for the day-term matrix of Weibo labeled with dates and that labeled with topics were similar to Figures 1 and 2 and, thus, are omitted here. Details for each topic and its dominance periods are shown in Table AII. A scrutiny of topics and summaries revealed that topics in Chinese microblogs can be classified into 12 categories, as shown in Figure 5.

Figure 5 reveals temporal trends of topics in Chinese Ebola microblog entries. Since the beginning of the Ebola outbreak in West Africa in 2014, Chinese users have kept disseminating background knowledge of Ebola until the mid-late phase. In the second month, some purchasing agents took advantage of Ebola hotspots and posted advertisements that were actually semantically irrelevant to Ebola. For example, in the tweet of “抵抗 […] 埃博拉病毒.行动起来！日本代购三次元舒 […]” (To fight against

![Graph showing dominance durations of the most salient topic categories in Chinese microblogs in each month.](image.png)

**Note:** The vertical axis indicates the number of days when the topic category was the most salient.
the Ebola virus! Purchasing Japanese three-D Shu [...]], the product for sale was irrelevant to Ebola. Chinese microblogs of precautions and response and actions of organizations became salient nearly two months later than the English tweets. As early as late March, many Twitter users noticed NIH’s measures of awarding a five-year grant of up to $28m to fight the deadly Ebola virus. However, not until mid-May did Chinese users focus their attention on the precautions and response of dispatching expert panels to Zaire by the World Health Organization and other agencies.

A comparison between Figures 3 and 5 revealed that Chinese and English Ebola microblogs shared nine topic categories in common, such as precautions and response, actions of organizations, and prediction. Three different topic categories, i.e., background knowledge of Ebola, advertisements and rumors, uniquely existed in Chinese Ebola microblogs. Two other categories, i.e., possible cause of Ebola and evaluation, uniquely existed in English Ebola tweets.

In Chinese Ebola microblogs, Background knowledge of Ebola and actions of organizations were the most salient and external environment, prediction and rumors were the least salient. Background knowledge of Ebola was the most salient topic category both in the first and second stages. This finding was supported by Alicino et al. (2015), who also found that Ebola symptoms, what is Ebola, about Ebola and what Ebola were among the top Ebola-related Google queries. The finding revealed that the public is eager to know the background knowledge of Ebola. Advertisements, precautions and response and actions of organizations were also salient compared with other categories, but not as salient as background knowledge of Ebola. Thus, these categories are shown in dashed lines in Figure 6. The most salient topic categories at the third stage were actions of organizations and precautions and response. See Figure 6 for an illustration of the topical evolution of Chinese Ebola microblogs.

It is seen in Figure 6 that Chinese users tended to take microblog platforms as the channel of spreading medical knowledge of infectious diseases and informing the public of organizations’ actions, many of which are related to measures to fight diseases. Chinese users tended to accept the existing status and adopt pragmatic attitudes, such as knowing about Ebola and how to fight it, while Twitter users were enthusiastic to explore the cause and consequences of Ebola.

In contrast to Chinese microblogs, the absence of background knowledge of Ebola and the downplaying position of precautions and response among salient categories in English tweets are worthy of much attention. It was found that the public is actually highly concerned with symptoms of Ebola virus (relevant to background knowledge of Ebola), safety travel and guidelines for protection from Ebola (relevant to precautions and response) (Lazard et al., 2015); thus, it is necessary for public emergency administrators to inform the public of the Ebola knowledge, infection risk during travel and precautions in Ebola prevention.

![Figure 6. Topical evolution of Chinese Ebola microblogs](image_url)
As for the developing trends of individual topic categories, actions of organizations, secondary events and advertisements kept increasing. Background knowledge of Ebola and actions of the public kept decreasing. Precautions and response, status description and statistics and description fluctuated. Prediction, emotions of the public, rumors and external environment were ephemeral. See Table II for the temporal trends of topic categories in Chinese Ebola microblogs.

Compared with Chinese microblogs, English tweets showed significant strengths in the categories of prediction, status description, statistics and description, secondary events, actions of the public and emotions of the public. Among them, Emotions of the public in Chinese microblogs were mostly gratitude and respect, while this category in English tweets was mostly anger and hatred toward Ebola.

Unfortunately, the category of rumors was found to be salient uniquely in Chinese microblogs for two days. The contents were about whether salmons were infected with Ebola and able to spread this disease. Many Chinese users really believed this rumor until it was refuted two days later. A thorough search for this content in English tweets in the same period retrieved only one tweet, which was actually about the phenomenon in China; i.e., “You can get Ebola from Atlantic salmon? That’s what many in China believe. Market op for NZ king salmon? www.stuff.co.nz/business/10490846/Ebola-a-boost-for-NZ-salmon […]” Then, why did this happen only in Chinese microblogs? The abnormal phenomenon is worthy of attention, and it is necessary to remind Chinese users to stay calm and not to forward suspicious and unverified information.

In the category of external environments, Twitter users tweeted about ISIS surges and the Ukraine cease fire, usually co-occurring in news tweets. Weibo users tweeted about biomedical stock prices and seemed to believe that those companies were related to Ebola, although there was no such clue.

The results of this study revealed several important implications. First, the topical evolution patterns and temporal trends of microblogs on a public health emergency vary depending on the microblog platforms of different regions and languages. The invisibility of the possible cause of Ebola and prediction categories in Chinese microblogs indicated that Chinese users tended to accept existing conditions, seldom exploring causes or consequences of the public health emergency. Instead, they were quite pragmatic and utilized the microblog platform as a channel of spreading background knowledge on the infectious disease. They were concerned about effective measures to fight the deadly disease. Emotions of the public were mostly gratitude and respect toward medical workers who assisted patients in infected areas. In contrast, Twitter users seemed to pay much more attention to the dynamics of the Ebola outbreak than Chinese users and kept long term strong foci on many aspects of Ebola, such as secondary events and status description. They were keen to discuss the causes and consequences of Ebola and regularly evaluate the severity of the Ebola outbreak. Background knowledge of Ebola here was also less salient than that in Chinese microblogs. Emotions of the public, which are mainly anger and hatred toward Ebola, peaked at the end of the investigated time span. Thus, the public health emergency management departments need to inform the public of sufficient background knowledge of infectious diseases to better meet the information needs of the public.

<table>
<thead>
<tr>
<th>Temporal trends</th>
<th>Topic categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing</td>
<td>Advertisements, secondary events and actions of organizations</td>
</tr>
<tr>
<td>Decreasing</td>
<td>Actions of the public and background knowledge of Ebola</td>
</tr>
<tr>
<td>Fluctuating</td>
<td>Precautions and response, status description and statistics and description</td>
</tr>
<tr>
<td>Ephemeral</td>
<td>Prediction, emotions of the public, rumors and external environment</td>
</tr>
</tbody>
</table>

Table II. Temporal trends of topic categories in Chinese Ebola microblogs
regarding the cause and consequences of the disease and to comfort high emotions of the public when necessary.

Second, as a public health emergency develops, the relevant tweets exhibit stage-specific topical characteristics. At the first stage of the public health emergency, people usually wonder about the cause and consequences of the infectious disease. In the second stage, people are concerned with the statistics and description of the emergency. Some secondary events may also draw their attention. Background knowledge of the disease is often what people want to know at the first two stages. At the third stage, people were concerned with effective precautions/responses and actions of organizations, while keeping an eye on the current status. If there is no sign of improvement, the public may lose control of their emotions. The findings can help emergency administrators to better understand the concerns of the public and take effective measures, including but not limited to providing the public with corresponding information.

Third, microblog users in some regions are susceptible to rumors or misleading information and can be harassed by irrelevant advertisements during a public health emergency. For example, some Chinese users took advantage of Ebola hotspots and posted advertisements that were actually irrelevant to Ebola. Many users believed rumors, such as salmon infected with Ebola were able to spread this disease. Thus, Chinese public health management departments need to identify and refute rumors and misleading information in time and effectively. Advertisements that linked themselves to irrelevant hotspots should also be controlled on microblog platforms.

Conclusion
The visual temporal analysis of topics in microblogs can reveal topical evolution patterns when public health emergencies unfold. The research findings of this study can help emergency management departments better understand the development track of a major public health event and concerns of the public at different stages and take effective measures to fight infectious diseases in similar cases.

In this study, we illustrate how an effective information visualization technique SOM in combination with the LDA technique can be used to analyze temporal distributions of topics in microblogs of infectious diseases and summarize topical evolution patterns of microblogs as the disease outbreak progressed. More than 450,000 Chinese and English Ebola microblog entries on the Ebola outbreak in West Africa were explored in two well-known microblog platforms between February 1, 2014 and October 31, 2014.

A novel aggregation and competitive SOM node-labeling approach was proposed, in which each SOM node is labeled with the most salient topic. Topics in Chinese and English Ebola microblogs during different periods were explored and summarized into 14 categories, such as actions of organizations, actions of the public and background knowledge of Ebola.

Different topical evolution patterns and temporal trends were found in microblogs on two microblog platforms. Four types of temporal trends were found for the involved topic categories, i.e., increasing, decreasing, fluctuating and ephemeral ones. The findings of our study can facilitate understanding of the development track of a major public health event and concerns of the public toward infectious diseases and topical differences between Chinese and English microblog platforms. Practical suggestions are provided to public health emergency management departments based on these findings. The constructed methods can be applied to other social media analyses of emergencies or major events.

Limitations and future work
A limitation of this study lies in its data source. Only two microblog platforms were investigated, while others exist, such as Tencent Weibo (Chinese) and Tumblr. Only one
public health emergence case was examined, and it is unknown whether microblogs on Zika and other emergencies have similar topical evolution patterns and temporal trends.

In the future, we will take other public health emergencies and other microblog platforms as cases and data sources and test whether the 14 topic categories apply to the related microblogs and similar temporal development patterns can be found. We will also study the evolution patterns and temporal trends of microblogs on public health emergencies by users of different types to find their specific characteristics.

References


Further reading

<table>
<thead>
<tr>
<th>Dates</th>
<th>Days</th>
<th>Topics</th>
<th>Terms</th>
<th>Summaries</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1<del>2-4, 2-6, 2-8</del>2-9, 2-11, 2-13, 2-15<del>2-22, 2-26</del>2-27, 3-1, 3-3, 3-8<del>3-19, 4-26, 4-29</del>5-1, 5-13<del>5-18, 5-21</del>5-25, 6-1, 6-16, 2-28, 3-2, 4-4<del>3-7, 4-5-3, 5-7</del>5-12, 5-19~5-20, 6-2, 6-13, 6-15, 7-22</td>
<td>50</td>
<td>Topic 1</td>
<td>Medical, immigration, pop, problem, quarantine, population, chief, officer, solve, service, suggest, Jean Marie, explore, day, team, global, France, medicine, staff, French</td>
<td>Immigration and population problem</td>
<td>Possible cause of Ebola/prediction</td>
</tr>
<tr>
<td>3-20~3-21</td>
<td>2</td>
<td>Topic 19</td>
<td>World, organization, health, say, security, drive, council, tell, global, said, people, recent, you've got, sneeze, threat, dinner, worldwide, declare, Spain, party</td>
<td>Activity of health institution and a metaphor about Ebola</td>
<td>Precautions and response/actions of organizations</td>
</tr>
<tr>
<td>3-22<del>3-23, 3-29, 3-31, 5-2</del>5-5, 5-20<del>5-31, 6-4</del>6-10, 6-14, 6-17~6-18</td>
<td>24</td>
<td>Topic 47</td>
<td>Death toll, Ebola death, rise, Guinea, hit, double, reach, Ebola toll, jump, African Ebola death, number, lenone Ebola death, pass, reduce, climb, Africa rise, guinea Ebola death, almost, deaths rise, toll jump</td>
<td>Death toll and status report in Africa</td>
<td></td>
</tr>
<tr>
<td>3-24, 3-26~3-28, 4-8, 6-19</td>
<td>6</td>
<td>Topic 18</td>
<td>Outbreak, spread, African Ebola outbreak, serious, still, panic, spread to Europe, spreading situation, Europe, fever death, outbreak spreads panic, Spanish Arabia, out of Africa, Aljazeera, situation serious, spreading situation, tropical, remain, inquisitr, Geneva</td>
<td>Possible spread of Ebola out of Africa</td>
<td>Prediction</td>
</tr>
<tr>
<td>3-25, 4-20</td>
<td>2</td>
<td>Topic 26</td>
<td>Virus in Africa, new strain, Guinea, cause, study, identify, report, scientist, surprise, important, origin, Zaire, kill, circulate, Ebola strain, strain of disease, identify, report Ebola, Ebola virus circulate, variant</td>
<td>Discussion about new strain description</td>
<td>Status description</td>
</tr>
<tr>
<td>3-30, 4-1, 4-6<del>4-7, 4-13</del>4-16, 4-21, 4-24, 7-27~7-28, 7-30, 8-29</td>
<td>14</td>
<td>Topic 30</td>
<td>Guinea, killer, deadliest, worst, Conakry, Dr, capital, history, report, place, track, Sanjay Gupta, stop, CNN, Conakry guinea, explain, reach, record, show, sure</td>
<td>Ebola reaches Conakry, Guinean capital</td>
<td>Status description/Secondary event description</td>
</tr>
<tr>
<td>4-2, 4-4, 6-20~6-29</td>
<td>12</td>
<td>Topic 46</td>
<td>West Africa, combat, Morocco, wave, outbreak plague, spread in west, telegraph, crime, vox, mass, virus outbreak, early, online, mother, hold, mobile, whatshot, piece, west African nation, obamacare</td>
<td>Ebola outbreak spreads panic in West Africa</td>
<td>Status description</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Dates</th>
<th>Days</th>
<th>Topics</th>
<th>Terms</th>
<th>Summaries</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-3, 4-17<del>4-18, 4-22</del>4-23, 6-3, 6-30~7-6</td>
<td>13</td>
<td>Topic 9</td>
<td>Worse, ISIS, insider, suspect, enter, surge, fire, beltway, Ukraine, cease fire, please, worse than, general, Iraq, Ferguson, surgeon, get worse, outbreak worse, ISIS and Ebola, enterovirus</td>
<td>Other simultaneous important news like ISIS surges and Ukraine cease fire</td>
<td>External environment</td>
</tr>
<tr>
<td>4-9~4-11, 7-7, 7-21</td>
<td>5</td>
<td>Topic 14</td>
<td>Outbreak, challenge, respond to Ebola, outbreak in Africa, most, map, news Ebola outbreak, a lot of, most challenging, speed, cutting down forest, Ebola outbreaks worse, Ebola outbreak ever, mother, who says Ebola, making Ebola worse, began, geographic, recent Ebola outbreak, unusual</td>
<td>West Africa Ebola outbreak is among “most challenging” ever and respond to Ebola</td>
<td>Evaluation</td>
</tr>
<tr>
<td>4-12, 8-28</td>
<td>2</td>
<td>Topic 15</td>
<td>Virus, disease, know, Nigeria, situation, prevent, music, viral, disease outbreak, April, should, June, product, truth, shock, feat, Ebola situation, disease in west, program, site</td>
<td>Nigeria at risk of Ebola Virus and prevention</td>
<td>Status description/Precautions and response</td>
</tr>
<tr>
<td>4-19, 4-25, 6-12</td>
<td>3</td>
<td>Topic 17</td>
<td>Guinea, border, close, story, Senegal, spread, shut, school, list, fight, dying, south, open, close, top, alarm, fever, change, stories, listen</td>
<td>Shut borders and close schools to fight Ebola</td>
<td>Precautions and response/actions of organizations</td>
</tr>
<tr>
<td>4-27~4-28, 5-6, 6-11, 8-25, 9-11</td>
<td>6</td>
<td>Topic 31</td>
<td>Survivor, face, second, stigma, tell, blood, return, video, UK, heat, caught, Dr, shun, infect, luck, market, friend, Kent Brantly, end, explain</td>
<td>Ebola survivors face stigma</td>
<td>Status description</td>
</tr>
<tr>
<td>7-8~7-20, 7-23</td>
<td>14</td>
<td>Topic 37</td>
<td>Dead, deadly virus, flu, clinic, person, infect, river, rare, drop, avoid, bodily fluid, potential, deadly disease, glimpse inside, dan rive, brown, republic, fall, push, candidate</td>
<td>Death in west Africa and rare glimpse inside deadly Ebola clinic</td>
<td>Statistics and description</td>
</tr>
<tr>
<td>7-24, 8-11, 8-14, 8-16~8-20, 9-10, 9-21, 9-23</td>
<td>11</td>
<td>Topic 10</td>
<td>United States, Ebola update, Lagos, governor, nation, red, news, special, today, alert, Nigeria, prevent, important, host, sale, bat, virus case, united nation, Ebola virus case, issue</td>
<td>Ebola case in Lagos</td>
<td>Status description/secondary event</td>
</tr>
<tr>
<td>7-25, 7-29, 7-31<del>8-5, 8-7, 8-10, 8-21, 8-24, 9-4, 9-6, 9-8, 9-27</del>9-29</td>
<td>18</td>
<td>Topic 29</td>
<td>American doctor, treat, top, Nebraska, admit, NIH, expose, chief Ebola doctor, hero, doctor with Ebola, doctor infected, arrive, key, American Ebola patient, play, two American, progress, role, doctor contracts Ebola, fight</td>
<td>American doctor contracted Ebola</td>
<td>Secondary event</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Dates</th>
<th>Days</th>
<th>Topics</th>
<th>Terms</th>
<th>Summaries</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-26, 8-23, 8-30</td>
<td>3</td>
<td>Topic 38</td>
<td>Nation, net, Ghana, BBC news, link, citizen, man, live, recent, positive, daily, British, quarantine, west African nation, tests negative, matter, Ebola test, danger, negative for Ebola</td>
<td>A Briton in Sierra Leone tested positive for the Ebola virus</td>
<td>Secondary event</td>
</tr>
<tr>
<td>8-6, 8-8<del>8-9, 8-15, 9-1, 9-7, 9-9, 9-14</del>9-20, 9-22, 9-24~9-26</td>
<td>18</td>
<td>Topic 5</td>
<td>Help, send, more, million, volunteer, troop, donate, battle, Obama, Ebola fight, grow, money, journal, found, provide, bill, street, tech, military, peace</td>
<td>Nigeria seeking volunteers and donation to fight Ebola</td>
<td>Status description/ Precautions and response</td>
</tr>
<tr>
<td>8-12, 8-22, 8-27, 9-2, 9-13</td>
<td>5</td>
<td>Topic 40</td>
<td>West African country, Ebola hit, flight, affect, rest, Ghana, strike, support, follow, pledge, meet, stop, finance, axe, restrict, Africa travel ban, bank, lift, Kenya, advice</td>
<td>Cancel or ban flights from west Africa and support Ebola effort</td>
<td>Precautions and response/actions of organizations</td>
</tr>
<tr>
<td>8-13, 8-26, 8-31, 9-3, 9-12</td>
<td>5</td>
<td>Topic 27</td>
<td>Experimental Ebola drug, use, monkey, fund, funny, develop, anti, Nigeria, hold, treat, small, mass, infect, company, serum, stay, away from, test, save, effect</td>
<td>Use of experimental Ebola drug and it cured monkeys</td>
<td>Precautions and response/actions of organizations</td>
</tr>
<tr>
<td>9-30<del>10-1, 10-3</del>10-31</td>
<td>31</td>
<td>Topic 41</td>
<td>Get Ebola, fuck, hope, shit, because, joke, catch, think, everyone, want, sick, real, nigga, like, going to, someone, kid, told, Ebola shit, because of</td>
<td>Anger and hatred toward Ebola</td>
<td>Emotions of the public</td>
</tr>
</tbody>
</table>

Table AI.
<table>
<thead>
<tr>
<th>Dates</th>
<th>Days</th>
<th>Topics</th>
<th>Terms</th>
<th>Summaries</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-3, 2-11<del>2-13, 2-24</del>2-26, 3-1<del>3-3, 3-5</del>3-11, 3-13<del>3-14, 3-17, 3-19</del>3-28, 3-30<del>5-6, 5-8</del>5-14, 5-18<del>6-21, 6-23</del>7-6, 7-11~7-26</td>
<td>140</td>
<td>Topic\textsubscript{cn} 28</td>
<td>等级,蠢,视为,白人,陈坤,李奇,下层,恶棍,源头,母亲,陆,续,婴儿,呕,冲击,粪便,援引,排出,祖母,幸,好,荣耀</td>
<td>Symptom and explanation of Ebola</td>
<td>Background knowledge of Ebola</td>
</tr>
<tr>
<td>3-29, 6-22, 7-7~7-10, 10-23</td>
<td>7</td>
<td>Topic\textsubscript{cn} 38</td>
<td>购物,蜗牛,牛奶,修复,货源,胶原,香皂,小黄,蛇毒,洗面,耳朵,月饼,金牌,击剑,狂笑,美好,排球,不公平</td>
<td>Purchasing agents and commodities</td>
<td>advertisements</td>
</tr>
<tr>
<td>5-15<del>5-17, 7-27</del>7-28, 8-22, 8-24, 10-30~10-31</td>
<td>9</td>
<td>Topic\textsubscript{cn} 15</td>
<td>级,危,视为,白人,陈坤,李奇,下层,恶棍,源头,母亲,陆,续,婴儿,呕,冲击,粪便,援引,排出,祖母,幸,好,荣耀</td>
<td>Actions by health organizations</td>
<td>Precautions and response/actions by organizations</td>
</tr>
<tr>
<td>7-29, 8-2</td>
<td>2</td>
<td>Topic\textsubscript{cn} 24</td>
<td>吐,病毒,病毒性,疼,肌肉,潜伏期,疼痛,虚弱,极度,外出,内出血,急性病,感染者,外出血,皮疹,后会,休克,特征,肝脏,肾脏,呕吐,毒性,病毒,毒性,呕吐,中毒,急性,出血,感染,疾病,慢性,出血,内出血,变化,历史上,疗法,出血,中风,恶心,疫苗,恐怖主义,内出血</td>
<td>Description of medical features and symptoms of Ebola</td>
<td>Background knowledge of Ebola</td>
</tr>
<tr>
<td>7-30, 8-3</td>
<td>2</td>
<td>Topic\textsubscript{cn} 33</td>
<td>死亡率,传染率,死亡率,烈性,原因,特效药,吐,内容,体内,外出,改变,历史上,疗法,出血,中风,恶心,疫苗,恐怖主义,内出血</td>
<td>Discussion about the (possible) spread of Ebola</td>
<td>Prediction/status description</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Dates</th>
<th>Days</th>
<th>Topics</th>
<th>Terms</th>
<th>Summaries</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-31<del>8-1, 8-4, 8-7</del>8-8, 8-11~8-12, 10-24</td>
<td>8</td>
<td>Topic cn 12</td>
<td>虐,蔓延,博,文,分享,丧失,失控,面临,可怕,客户,多国,指出,背后,致死率,中新社,暴发,博弈,转机,地境,没救,孔洞</td>
<td>Death toll and number of infected people</td>
<td>Statistics and description</td>
</tr>
<tr>
<td>8-5~8-6</td>
<td>2</td>
<td>Topic cn 14</td>
<td>感染者,人体,预计,设备,美国医生,布兰特利,出院,回到,埃里大学医院,加强,级别,助理,配备,专机,戒备,怀特博尔,极度,有病,有望,升级</td>
<td>Condition of American Ebola patient</td>
<td>Status description/secondary event</td>
</tr>
<tr>
<td>8-9~8-10</td>
<td>2</td>
<td>Topic cn 6</td>
<td>利比里,塞拉,前往,公共卫生,我国,新华网,兔,事儿,蓝,不在,通知,警惕,威武,频道,就医,新华社,塞拉利昂埃博拉,蓝,雄壮,启程</td>
<td>Chinese medical aid and description of Ebola status in West Africa</td>
<td>Actions by organization/status and description</td>
</tr>
<tr>
<td>8-13~8-14, 10-20</td>
<td>3</td>
<td>Topic cn 47</td>
<td>报道,蔓延,结束,宣布,非洲人,好战,路透社,尼日利亚埃博拉,大面积,不可能,召开,视野,染上,揭密,开会,成熟,否认,面临的,本人,身着,report, spread, end, announce, African, become better, Reuters, Nigeria Ebola, large area, impossible, convone, vision, infected with, dense, meet, mature, deny, face, I, wear</td>
<td>American Ebola patient got better and death of Nigeria Ebola patients</td>
<td>Status description</td>
</tr>
<tr>
<td>8-15, 8-26, 8-29</td>
<td>3</td>
<td>Topic cn 21</td>
<td>人数,凶,工作者,前线,凶猛,至关重要,路径,惨烈,总和,珍贵,致敬,松,三,邱医生,实验室,拥有,医学界,宿主,来源,级别Number of people, fierce, worker, frontline, ferocious, crucial, path, tragic, sum,</td>
<td>Death toll and death of medical workers</td>
<td>Statistics and description</td>
</tr>
<tr>
<td>Dates</td>
<td>Days</td>
<td>Topics</td>
<td>Terms</td>
<td>Summaries</td>
<td>Categories</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>8-16, 8-25, 8-27~8-28</td>
<td>4</td>
<td>Topic.cn 3</td>
<td>precious, salute, squirrel, SARS, Dr. Hao, lab, has, medical community, host, source, level</td>
<td>Thanks and salute to the dead Chinese medical workers</td>
<td>Emotions of the public</td>
</tr>
<tr>
<td>8-17, 8-30~9-1, 9-5</td>
<td>5</td>
<td>Topic.cn 23</td>
<td>世界卫生组织, 人数, 警告, 蔓延, 案例, 失控, 须要, 通报, 日内瓦, 花费, 声明, 证兆, 西非地区, 数据, 速度, 评估, 国外, 数量, 低估</td>
<td>Announcement by world health organization</td>
<td>Statistics and description/actions by organization</td>
</tr>
<tr>
<td>8-18~8-19</td>
<td>2</td>
<td>Topic.cn 30</td>
<td>我国, 面, 在线, 超人, 于莺, 撤回, 再说, 友谊医院, 外籍, 先前, 大意, 擅长, 医疗界, 立交桥, 无稽之谈, 不知道, 回应, 人家</td>
<td>Discussion about dead and infected Chinese doctors</td>
<td>Status description/secondary event</td>
</tr>
<tr>
<td>8-20~8-21, 8-23</td>
<td>3</td>
<td>Topic.cn 27</td>
<td>研制, 实验室, 核酸, 流行病, 我国, 研究所, 达安基因, 证明, 高, 院, 药品, 卫生部, 军队, 缺乏, 指出, 从事, 流行病学, 候, 由埃博拉病毒</td>
<td>Developing Ebora virus nucleic acid detection reagent</td>
<td>Precautions and response</td>
</tr>
<tr>
<td>Dates</td>
<td>Days</td>
<td>Topics</td>
<td>Terms</td>
<td>Summaries</td>
<td>Categories</td>
</tr>
<tr>
<td>-------------</td>
<td>------</td>
<td>--------</td>
<td>--------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>9-2, 9-7~9-8</td>
<td>3</td>
<td>Topic25</td>
<td>bodily fluid, infected person, secretion, pathway, believe, ape, strange, mortality, bat, local people, excrement, grinder, host, Westerners, nervous, saliva, trick, needle, chimpanzee, means</td>
<td>Knowledge of protecting from Ebola response</td>
<td>Precautions and response</td>
</tr>
<tr>
<td>9-3~9-4, 9-6, 9-9, 9-16, 9-18, 9-28</td>
<td>7</td>
<td>Topic26</td>
<td>fight, hotspot, Youth Olympic, future, strictly prevent, baur, upgrade, highly, Liu Weizhong, regret, discriminate, treatment, influence, very much, fat, volunteer, rumor, fever, nationwide, cure the sickness and save the patient</td>
<td>About fighting Ebola, Precautions and including quarantine, tracking health of immigrants, etc.</td>
<td>Precautions and response</td>
</tr>
<tr>
<td>9-12</td>
<td>1</td>
<td>Topic27</td>
<td>customer, share, deceased, fox, crowd, poisoned, street, bury, address, mobile phone, walk, wake up, Liberia Ebola, front page, photos, on the way, die of, morgue, go into the street, flee</td>
<td>Discussion about salmon infected with Ebola</td>
<td>Rumors</td>
</tr>
<tr>
<td>9-14</td>
<td>1</td>
<td>Topic21</td>
<td>statement, human, illness, salmon, infected, rumor, app, advertisement, regulate, stop, open, experimental, impact, infectious, seafood, smile, thinking, middle, channel</td>
<td>Refute rumors about salmon infected with Ebola</td>
<td>Rumors</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Dates</th>
<th>Days</th>
<th>Topics</th>
<th>Terms</th>
<th>Summaries</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>9-20<del>9-22, 9-25, 10-5</del>10-7</td>
<td>7</td>
<td>Topic_{cn} 20</td>
<td>statement, human body, fish, infected person, in vivo, refute rumors, assured, app, China Import and Export Fair, provision, stop, start, experimental, contracted, infectious, seafood, smile, idea, middle, channel</td>
<td>Attack on Ebola education workers</td>
<td>Secondary event/actions of the public</td>
</tr>
<tr>
<td>9-30, 10-8~10-12</td>
<td>6</td>
<td>Topic_{cn} 45</td>
<td>Attack on Ebola education workers</td>
<td>First Ebola case in the United States</td>
<td>Secondary event</td>
</tr>
<tr>
<td>Dates</td>
<td>Days</td>
<td>Topics</td>
<td>Terms</td>
<td>Summaries</td>
<td>Categories</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>10-14<del>10-17, 10-21</del>10-22</td>
<td>6</td>
<td>Topic(<em>c</em>) 41</td>
<td>market, military project, market, expectation, holding, listed, fall, new share</td>
<td>Clothing</td>
<td>Advertisements</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>精彩,外套,真皮,男装,长袖,女装,春秋,瘦,时尚,英伦,正品,皮鞋,欧,欧美,宽松,长裤,连衣裙,中年,上衣,鞋子, wonderful, coat, leather, men's wear, long sleeve, women's wear, spring and autumn, thin, fashion, British, genuine, leather shoes, Dad, Europe and America, loose, trousers, dress, middle age, jacket, shoes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-19</td>
<td>1</td>
<td>Topic(<em>c</em>) 9</td>
<td>中医药,板蓝根,远离,非洲人,管理局,指南,妈,爸,回复,机票,照片,虎,右边,头号,好使,刘臻,心中,共和国,浑身解数,Traditional Chinese medicine, Radix Isatidis, away from, African, administration, guide, mother, father, reply, air ticket, photo, little tiger, right, top, scold, work well, Liu Zhen, in the heart, republic, use all one's skill</td>
<td>Study of using traditional Chinese medicine to cure Ebola</td>
<td>Precautions and response</td>
</tr>
</tbody>
</table>
About the authors

Lu An is a Professor with the School of Information Management of Wuhan University. She received her PhD and Master Degrees of Information Science from Wuhan University. She has published two monographs and more than 40 journal articles on topics relating to visual knowledge discovery and health informatics.

Chuanming Yu is a Professor with the School of Information and Safety Engineering of Zhongnan University of Economics and Law. He received his PhD and Master Degrees of Information Science from Wuhan University. He has published more than 40 journal articles and one monograph on topics relating to knowledge engineering and text mining. Chuanming Yu is the corresponding author and can be contacted at: yuchuanming2003@126.com.

Xia Lin is a Professor with the College of Computing and Informatics of Drexel University. He received his PhD Degree from the University of Maryland at College Park. He has published more than 60 journal articles on topics relating to information visualization and information retrieval.

Tingyao Du is a Master with the School of Information Management of Wuhan University. His research interest is visual knowledge discovery and social network analysis.

Liqin Zhou is a PhD Student with the School of Information Management of Wuhan University. She has published two journal articles on topics relating to crisis informatics and text mining.

Gang Li is a Professor with the School of Information Management of Wuhan University and the director of Center for Studies of Information Resources, Wuhan University. He received his PhD and Master Degrees of Information Science from Wuhan University. He has published more than 140 journal articles on topics relating to information resource management and text mining.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Context congruity effects of online product recommendations: an eye-tracking study

Jing Luan, Zhong Yao and Yongchao Shen
Department of Economics and Management, Beihang University, Beijing, China, and
Jie Xiao
Department of Business Administration, Hunan University, Changsha, China

Abstract
Purpose – The purpose of this paper is to understand how the context congruity effects of online product recommendations (PRs) by recommendation agents (RAs) influence consumers’ attention to and memory of recommended products in an online shopping environment.
Design/methodology/approach – The study focuses on the context congruity effects of online PRs by examining consumers’ browsing patterns and attention characteristics (fixation counts and fixation duration) using an eye-tracking device and by measuring memory performance with an aided memory test. Three types of PRs (highly congruent, lowly congruent and incongruent PRs) and two degrees of involvement (high and low involvement) are considered.
Findings – The results of the gaze data show that context congruity effects can influence consumers’ PR attention, but this effect is not moderated by involvements. The results of the memory data show that PR recognition is influenced not only by context congruity effects but also by involvement. Another significant finding is that attention to a PR does not necessarily guarantee better memory performance.
Practical implications – The study significantly contributes to deepening the understanding of how context congruity can influence consumers’ attention to and memory of PRs. The findings also have important managerial and practical implications, such that selecting and presenting PRs should be based on context congruity effects.
Originality/value – First, introducing context congruity effects to investigate the effectiveness of online PRs by RAs not only provides an important theoretical contribution to research on recommendation effectiveness but also enriches its application. Second, the findings suggest that the relationship between visual attention and memory is not definitely positive. Third, to interpret the complex translation process from attention to memory, the authors propose a methodology that considers stimulus attributes, issue involvement, cognitive capacity and cognitive interference.

Keywords Memory, Eye tracking, Attention, Context congruity, Online product recommendation

Introduction
In the era of the Web 2.0, e-commerce offers consumers great convenience, tremendous product choices, and a significant amount of product-related information. Abundant product information affords consumers not only opportunities to evaluate products more comprehensively but also difficulties in processing mass information due to humans’ cognitive limitations. Identifying products that satisfy consumers is not an easy task. The availability of recommendation agents (RAs) aids consumers in searching, selecting and customizing products (Chung et al., 2016). RAs are web-based software applications that conduct a series of operations to provide users product advices based on their needs, preferences, profile, and shopping history. Thus, RAs have the potential to mitigate consumers’ information overload and search difficulty, further improving their decision quality.
Previous research on RAs primarily concerns its recommendation-generating process (Khusro et al., 2016). However, from consumers’ perspective, the effectiveness of RAs is not only determined by algorithms. Many other factors, such as perceived usefulness, ease of use, credibility, output strategy, and product and user-related factors also influence consumers’ evaluation of RAs. This study investigates RA-related issues beyond algorithms in order to enrich RA research. We mainly focus on how online product recommendations (PRs) by RAs work to attract consumers’ attention and whether consumers’ attention and memory processing varies based on context congruity effects of PRs.

Context or thematic congruity has previously been addressed by online and physical advertising-related research in order to understand how to achieve greater advertising effectiveness and impact (Moorman et al., 2002a, b; Porta et al., 2013). Referring to this line of research, context congruity in this paper is defined as the extent to which a PR and its context are congruent in terms of theme, content or context. More specifically, it refers to the relationship between the content of a PR and its surrounding context, and to the association that consumers can generate between them. By providing PRs based on techniques such as the content-based approach, collaborative filtering and hybrid approaches, the degree of congruity in PRs by RAs can be generally categorized as high congruity (HC), low congruity (LC) and incongruity. The context has HC when the PRs are similar to the browsing product. For example, if one is browsing rice cookers and the RA’s recommended products are other rice cookers, the context is highly congruent. With LC, the PRs are less related to the browsing product, such as complementary products. For example, for the rice cookers, if the products recommended by the RA are other items that people who bought rice cookers also purchased, such as a rice storage bin or a rice paddle, the context is lowly congruent. The third category is incongruity (IN), in which the PRs are unrelated to the browsing product and may be based on consumer’s previous browsing history, for example. To evaluate whether these recommendations are effective for consumers with regard to the context congruity effects, we will examine consumers’ attention and memory performance toward PRs as outcome measures. In an online shopping environment with overwhelming information, it is increasingly difficult for marketers to compete for limited consumer attention (Webb and Ray, 1979) and for effective product memory. Thus, in the context of congruity, understanding how consumers comprehend, process, memorize and use PRs by RAs contributes to the research on context congruity effects, and can serve as a reference for RA designers to optimize their creative efforts in assisting marketers to rise to this challenge.

Related research has examined the influence of congruity between the media context and advertising on consumers’ attitudes, reactions, purchase intentions and recall in the context of online newspapers or online portal websites (Porta et al., 2013; Rieger et al., 2015; Kim and Choi, 2012) as well as the congruity effect on consumers’ information procurement and processing (Luan et al., 2016; Rieger et al., 2015). However, online shopping websites or e-commerce portals that are popular among consumers for offering opportunities to address the everyday tasks of problem solving and information searching (Heinz et al., 2013) have not been explored enough. Heinz and Mekler (2012) examined the impact of advertising placement and navigation style (aimless browsing or goal-oriented information searching) on consumers’ advertisement recognition by displaying food-related ads in an online shop selling men’s and women’s apparels. They found that navigation style, not banner placement, can influence recognition. The authors suggested that when conducting research in the context of online shopping, researchers should vary the level of task involvement, such as providing opportunities for participants to choose what they want to search for and using an eye tracker to trace participants’ eye movements, in order to obtain additional insights into how consumers orientate themselves during online shopping. Although they presented very interesting and promising further research, the authors ignored the influence...
of context congruity. This paper narrows this gap by investigating the context congruity effects of PRs by RAs in online shopping portals such as Taobao in China and Amazon in the US and by considering task involvement and using an eye-tracking method.

**Literature review**

Although product RAs are popular in e-commerce, limited research has examined their recommendation effectiveness. PRs based on RAs, a type of online advertising, can share amounts of findings regarding the effectiveness of the online advertisements. Previous studies on traditional advertisements have revealed the obvious finding that context plays a significant role in information processing and advertisement effectiveness (Moorman et al., 2002a, b). In the internet context, literature reviews show that the degree of congruity between the advertising theme and web content has a critical influence on how ads are watched and remembered. Due to banner blindness, banner ads are increasingly losing their effectiveness. At the same time, more and more researchers have considered mitigating this issue from the perspective of context congruency or theme congruity. Therefore, using context congruity to explore the effectiveness of PRs by RAs has both practical and academic significance.

In general, congruity refers to the relationship between advertising content and the surrounding context of the media vehicle. Media context is an important situational factor that can influence consumers’ attention and information processing. Using empirical methodologies, many studies have found that congruent stimuli can help generate a comfortable situation where an individual can find expected and repeated items that conform to his/her already developed mental schema, resulting in a positive and focused attitude (Mandler, 1982; Jeong and King, 2010) and even enhanced memory (Heckler and Childers, 1992). Thus, consumers’ reaction to an advertisement will depend on the extent of the match between the media context and their schematic expectations.

Porta et al. (2013) studied the influence of congruity between ad subjects and news content on users’ ad watching and memory behavior in online newspaper. The findings revealed that thematic congruity did increase fixations on the ad, but it did not have a substantial effect on memory. Similarly, studying embedded ads in online newspaper, Rieger et al. (2015) tested the effect of context congruency, including pictures or text elements or both, on consumers’ awareness, memory and attitude toward an ad. The results showed that complete context congruency (both visual and textual elements) generates higher visual awareness, better memory performance and better attitudes toward the ad. In order to understand consumers’ underlying psychological mechanisms when processing online video advertising, Kim (2015) examined the impact of key online video ads on consumers’ ad avoidance and subsequent ad attitude and brand memory. The author noted that online video ads that were similar to the online video were perceived as more relevant and generated more positive attitudes and lower ad avoidance. On the other hand, online video ads that were dissimilar to online video caused greater ad avoidance and lower brand recognition. Segev et al. (2014) explored the influence of ad-context congruity on consumers’ responses to banner ads on blogs. The results demonstrated that an ad that is thematically congruent with the blog generates more favorable and positive responses. The authors also suggested that individuals’ issue involvement can moderate this congruity effect. Namely, highly involved individuals reacted more positively to the ad in a congruent context, while less involved individuals reacted more favorably to the ad in an incongruent context. All of these studies have investigated the context congruity effects on consumers’ responses, such as attention, attitude, and memory, based on the context of online newspapers, online videos, and blogs, but the effect in the online shopping context is still unclear. In addition, all of these studies verified the positive relationship between context congruity and attention, but the findings regarding its influence on memory are inconsistent. Therefore, in this
paper, we will explore context congruity effects in online shopping context and consider individuals’ issue involvement to investigate its relationship with memory.

In addition to human observation methods, some studies have employed machine observation methods such as eye-tracking device to capture consumers’ fixation behavior on ads (Porta et al., 2013; Kim, 2015; Bang and Wojdynski, 2016; Hsu, 2015; Rieger et al., 2015). Regarding the “eye-mind hypothesis,” Just and Carpenter (1976, 1980) claimed the existence of a direct correspondence between an individual’s gaze and his/her attention. Thus, eye-tracking methods are feasible for studying attention to PRs because they can reveal shifts and the dispersion of visual attention (Wedel and Pieters, 2008).

To date, a number of studies have used eye-tracking methodologies to explore congruity effects in an online environment. Porta et al. (2013) used fixation counts and total fixation duration to examine the influence of the thematic congruity of banner ads in an online newspaper. Their findings showed that thematic congruity led to more fixation counts, while the results for fixation duration depended on the testers’ reading condition. Meanwhile, the authors found that the relationship between fixation and memory is contingent. On the other hand, Wojdynski and Bang (2016) examined the influence of contextual congruency on users’ ability to process online news content. By analyzing fixation duration, the authors found that greater attention was paid to relevant ads than to irrelevant ads. In addition, while in a strong argument condition, users paid more attention to the story content than to the relevant ads. By measuring eye fixation, Hsu (2015) investigated the influence of ad-context congruity on advertising attention as well as the influence of advertising attention on advertising recall. His findings revealed that ad-context congruity had no significant impact on advertising attention; however, attention does have a positive relationship with aided recall, namely higher attention leading to higher recall. Based on these studies, the conclusions regarding the relationship between context congruity and fixation and between attention and memory are not consistent, so further research is necessary.

Development of hypotheses

Context congruity and PR attention

Based on the literature review of context congruity effects, with regard to both traditional media and web media, a positive role of context congruity on advertising information processing and advertising effectiveness has been documented. In particular, thematic congruity between an ad and its website context makes consumers generate positive reactions, such as more attention paid to the ad (Jeong and King, 2010; Segev et al., 2014; Kim, 2015; Porta et al., 2013). One possible reason that context congruity helps attract more attention is the cognitive priming effect. Priming increases the availability of the constructs or concepts used to assess stimuli in their surroundings (Higgins et al., 1985). Re-exposure to a construct increases its accessibility and potential, thus contributing to its usage in subsequent stimulus processing, judgment and evaluation (Herr, 1989). Therefore, context congruity effects prompt consumers’ susceptibility to PRs that are congruent with their background, making them process information more intensively and understand and evaluate both PRs and information better. Thus, the congruity between a PR and its context can lead consumers to pay more attention to content that fits with their processing schema or category.

In addition, involvement is considered to be one of the most significant consumer-related factors that influence advertising information processing (Segev et al., 2014) and advertising evaluation (Jeong and King, 2010). Involvement here means that the action is personally relevant, important and strongly associated with the individual’s self-concept and values (Petty and Cacioppo, 1981). Involvement can affect individuals’ level of attentiveness to informational content (Lee et al., 1999) and can moderate the amount and type of information
that they choose to seek and process (Cho, 2003). Regarding eye-tracking techniques, eye movement to an object means that the individual pays attention to it (Kim, 2015). When facing PRs, consumers extract information and elements from them during eye fixations, which are a reflection of visual attention. Fixation length and fixation count are commonly used eye-tracking metrics related to individuals' cognitive processing and visual attention. Longer fixation duration on an area of interest (AOI) indicates that the object of that AOI is more attractive, or it requires more investigation and engagement (Just and Carpenter, 1976). Similar to fixation duration, higher fixation counts on an AOI indicate that it is more noticeable or more important than other areas. Therefore, we hypothesize as follows:

**H1a.** Compared to lowly congruent and incongruent conditions, participants will, in an imposed condition, have a longer fixation duration in an highly congruent condition.

**H2a.** Compared to lowly congruent and incongruent conditions, participants will, in a free condition, have a longer fixation duration in an highly congruent condition.

**H1b.** Compared to lowly congruent and incongruent conditions, participants will, in an imposed condition, have higher fixation counts in an highly congruent condition.

**H2b.** Compared to lowly congruent and incongruent conditions, participants will, in a free condition, have higher fixation counts in an highly congruent condition.

**H3a.** Compared to the incongruent condition, participants will, in an imposed condition, have a longer fixation duration in an lowly congruent condition.

**H4a.** Compared to the incongruent condition, participants will, in a free condition, have a longer fixation duration in an lowly congruent condition.

**H3b.** Compared to the incongruent condition, participants will, in an imposed condition, have higher fixation counts in an lowly congruent condition.

**H4b.** Compared to the incongruent condition, participants will, in a free condition, have higher fixation counts in an lowly congruent condition.

**H5a.** With regard to fixation length, a high involvement condition will produce different results compared to a low involvement condition.

**H5b.** With regard to fixation counts, a high involvement condition will produce different results compared to a low involvement condition.

**Context congruity and PR memory**

Academic research focused on advertising effectiveness has revealed a positive association between context congruity and memory effects (Moorman et al., 2002a, b; Rieger et al., 2015; Kim, 2015). With regard to marketing practices, the most interesting strategic goal for marketers is to achieve a memory trace based on contextual adjustment. Schema-congruity theory (SCT) (Mandler, 1982) has been used to understand how people perceive and process information. It can help explain what may happen when people receive new information that it is either congruent or incongruent with their expectations. Consistent with the cognitive priming effect, SCT suggests that humans can develop a brain-stored schema for a stimulus based on their basic understanding, impression, or even prior related experience. Based on this brain-stored schema, when individuals evaluate a new input, if it is closely related to those processed by the current schema, schema congruity occurs, prompting their stimulus judgment and evaluation. Thus, recommended products that are congruent with their contexts can contribute to more intensive information processing, finer understanding, and more positive evaluation of both information and PRs, resulting in a better advertising memory.
Therefore, we hypothesize the impact of context congruity effects on advertising memory as follows:

\( H6a. \) Compared to lowly congruent and incongruent conditions, in an imposed condition, information in an highly congruent condition will be better recognized.

\( H7a. \) Compared to lowly congruent and incongruent conditions, in a free condition, information in an highly congruent condition will be better recognized.

\( H6b. \) Compared to lowly congruent and incongruent conditions, in an imposed condition, product recommendations under an highly congruent condition will be better recognized.

\( H7b. \) Compared to lowly congruent and incongruent conditions, in a free condition, product recommendations under an highly congruent condition will be better recognized.

\( H8a. \) Compared to the incongruent condition, in an imposed condition, information in an lowly congruent condition will be better recognized.

\( H9a. \) Compared to the incongruent condition, in a free condition, information in an lowly congruent condition will be better recognized.

\( H8b. \) Compared to the incongruent condition, in an imposed condition, product recommendations under an lowly congruent condition will be better recognized.

\( H9b. \) Compared to the incongruent condition, in a free condition, product recommendations under an lowly congruent condition will be better recognized.

\( H10a. \) With regard to information recognition, a high involvement condition will produce different results compared to a low involvement condition.

\( H10b. \) With regard to PR recognition, a high involvement condition will produce different results compared to a low involvement condition.

Research methods
The major objective of this study is to examine the influence of context congruity effects on consumers’ attention to and memory of PRs by RAs under two types of involvement. To examine these relationships, a 3 (PR type: highly congruent PR, lowly congruent PR and incongruent PR) × 2 (involvement: low and high involvement) factorial within-subject experiment was utilized. When we designed and conducted the eye-tracking experiment, we referred to Holmqvist et al.’s eye-tracking guidebook (Holmqvist et al., 2011). The experiment was conducted in the Human Factors and Ergonomics Laboratory at Beihang University. During the experiment, experimental stimuli were displayed on a Dell 20-inch monitor with a resolution of 1,024×768. Participants’ eye movements were recorded using Tobii Eye-Tracker T120 with a sampling rate of 120 Hz. The eye tracking data were processed using the Tobii Studio Analysis Tool 2.3.2.

Participants
In total, 35 healthy college students (18 female) from Beihang University, age from 20 to 30 (average age is 26.8), were recruited to participate in our eye-tracking experiment. The students were enrolled from a college-wide paid subject pool at Beihang University in exchange for a money reward. The participants were screened to ensure that they had normal or correct-to-normal vision. All of the participants submitted written consent forms and were paid $5 as a reward. All participants contributed valid and integrated data.
Materials and design
An eye-tracking laboratory experiment was designed to test our hypotheses. To create a real online shopping environment, we referred to a product display webpage on Taobao as our stimulus because Chinese consumers are very familiar with it. Our experimental webpage mainly includes five parts: product images, product descriptions, product reviews, recommended products and retailer information, as shown in Figure 1. All of this information was obtained from Taobao, Tianmao, Jingdong and ZOL (well-known Chinese online shopping portals). We chose computers, clothes, shoes, cell phones and movies as browsing products based on college students’ online shopping preferences and the most popular commodity list from the “Chinese online shopping market research report in 2015.”

Referring to real recommendation condition in an online shopping context, for each product in the experiment, there are three types of PRs to match it. Highly congruent products include two similar products (SPs), lowly congruent products include two related products (RPs), and incongruent products include two unrelated products (UPs), as shown in Table I. Typically, banner ads include three elements: the brand, the text and a picture (Pieters et al., 2002). Thus, the PRs by RAs displayed in our experiment contain the brand,
Referring to Porta et al. (2013), we introduce two shopping conditions: an imposed or a free condition, which represents consumers’ online shopping involvement. Here, involvement means that the action is personally relevant, important and strongly associated with the individual’s self-concept and values (Petty and Cacioppo, 1981). For the “imposed shopping condition,” participants are assigned to view a specific product without considering their interest and preferences, which is the low involvement condition. For the “free shopping condition,” subjects can choose to view a product from several product classes according to their demands or interests, which is the high involvement condition. Computers were used in the imposed online shopping condition (low involvement) due to their universality and utility. Clothes, shoes, cell phones, and movies were utilized in the free online shopping condition (high involvement) due to their better attractiveness and hedonistic attributes.

**Procedure**

Each participant signed up for a 20 min lab session. When the participants arrived at the lab, the researchers explained the experimental procedure to them and informed them that an eye-tracking device would unobtrusively track their eye movements. After they signed informed consents, the calibration process for the eye-tracking equipment was conducted.

After the calibration process was completed, the experiment began. The first scenario is the imposed online shopping condition. Computers are the browsing products in this condition, and the experiment scenario provided is “Now you will see three computers from the best-seller list for this week. Please choose the one computer that you favor the most.” The participants are given 50 sec to view each product page. The screen jumps to the next page automatically. After viewing three computers, on the last screen, the participants are asked to report their decisions. Next, the free online shopping condition begins. Clothes, shoes, cell phones and movies are the viewing products in this condition, and the experiment scenario is, “There are four categories of products (clothes, shoes, cell phones, movies), and for each one, three products will be presented to you. Please select one product category to browse and choose one product from this category for yourself.” Similarly, on the last screen, the participants are asked to report their decisions. The procedure in this part of the experiment is shown in Figure 2.

After the eye tracking experiment, an aided recall (recognition) test was administered. In total, 30 PRs that were created for the experiment were used for this test. The first scenario is the PR information recognition test. Textual information for each PR, not including the brand, was arranged randomly and presented on a page. Participants were asked to select the information they noticed during their online shopping process. The second scenario is the PR recognition test. A total of 30 PRs were displayed on another page, and the participants were asked to choose the PR they noticed while browsing the products.
Analysis and results

All 35 participants provided integrated and valid gaze data and memory data. Among them, 12 participants chose to view clothes, 7 shoes, 10 cell phones and 6 movies. As shown in Figure 1, the PRs (No. 4) are our AOI. To test our hypotheses, the eye tracking metrics fixation counts and fixation duration on an AOI were used for the analysis.

Testing the hypotheses related to context congruity and PR attention

In the imposed condition, the results of an ANOVA analysis of the fixation duration values by PR type show that there is a significant effect of PR type ($F(2, 68) = 4.953$, $P = 0.01$, $\eta^2 = 0.127$). The results of a post hoc analysis are shown in Table II, and hypothesis $H1a$ is supported and $H2a$ is not supported. For the results of the fixation counts, an ANOVA analysis of the PR types was conducted. The results show that effect of PR type is substantial ($F(2, 68) = 4.663$, $P = 0.013$, $\eta^2 = 0.121$). The results of a post hoc analysis are displayed in Table II, and hypothesis $H1b$ is supported and $H2b$ is not supported.

In the free condition, because different subjects chose different product categories to view, a preliminary ANOVA on fixation length with four product categories and three PR types was first conducted to check whether the differences among the product categories influence consumers’ ad attention behavior. The results indicate that the interactive effect of product category and PR type ($F(6, 62) = 1.457$, $P = 0.208$, $\eta^2 = 0.124$) and the effect of product category ($F(3, 31) = 0.306$, $P = 0.821$, $\eta^2 = 0.029$) are not significant. Thus, there are no substantial differences among the four product categories.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Matrix</th>
<th>$H$</th>
<th>Sig. Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imposed condition</td>
<td>$H1a$ Fixation duration</td>
<td>HC &gt; LC, HC &gt; IN</td>
<td>Yes [Results]</td>
</tr>
<tr>
<td>$H1b$ Fixation counts</td>
<td>HC &gt; LC, HC &gt; IN</td>
<td>Yes [Results]</td>
<td></td>
</tr>
<tr>
<td>$H2a$ Fixation duration</td>
<td>LC &gt; IN</td>
<td>No [Results]</td>
<td></td>
</tr>
<tr>
<td>$H2b$ Fixation counts</td>
<td>LC &gt; IN</td>
<td>No [Results]</td>
<td></td>
</tr>
<tr>
<td>Free condition</td>
<td>$H3a$ Fixation duration</td>
<td>HC &gt; LC, HC &gt; IN</td>
<td>Yes [Results]</td>
</tr>
<tr>
<td>$H3b$ Fixation counts</td>
<td>HC &gt; LC, HC &gt; IN</td>
<td>Yes [Results]</td>
<td></td>
</tr>
<tr>
<td>$H4a$ Fixation duration</td>
<td>LC &gt; IN</td>
<td>No [Results]</td>
<td></td>
</tr>
<tr>
<td>$H4b$ Fixation counts</td>
<td>LC &gt; IN</td>
<td>No [Results]</td>
<td></td>
</tr>
<tr>
<td>Imposed vs Free</td>
<td>$H5a$ Fixation duration</td>
<td>Difference</td>
<td>No [Interactive effect and main effect are not significant]</td>
</tr>
<tr>
<td>$H5b$ Fixation counts</td>
<td>Difference</td>
<td>No [Interactive effect and main effect are not significant]</td>
<td></td>
</tr>
</tbody>
</table>

Notes: HC, high congruence; LC, low congruence; IN, incongruence

Table II. Results of the testing hypotheses related to attention
Thereafter, without considering the product categories, an ANOVA analysis of the three PR types on fixation length was performed. The results show that the effect of PR type on ad attention is significant (F(2, 68) = 5.759, P = 0.005, η² = 0.145). The results of a post hoc analysis are shown in Table II, and thus, hypothesis H3a is supported and H4a is not supported.

For the results of the fixation counts in the free online shopping condition, a preliminary ANOVA analysis of fixation counts was first conducted. The results demonstrate that the interactive effect of PR type by product category (F(6, 62) = 1.356, P = 0.247, η² = 0.116) and the effect of product category (F(3, 31) = 0.449, P = 0.72, η² = 0.042) are not significant. Thus, product categories did not generate an appreciable influence. Then, an ANOVA analysis of the fixation counts on PR type indicates that the impact of PR type is significant (F(2, 68) = 4.27, P = 0.018, η² = 0.111). The results of a post hoc analysis are shown in Table II, and thus, hypothesis H3b is supported and H4b is not supported.

To compare the advertising attention divergence between the free and imposed online shopping conditions, a repeat measures ANOVA analysis of fixation length was conducted. The results show that the main effect of PR type is significant (F(2, 68) = 10.45, P < 0.001, η² = 0.235). However, the main effect of the online shopping conditions (F(1, 34) = 0.32, P = 0.573, η² = 0.009) and the interactive effect of PR type by shopping conditions (F(2, 68) = 0.025, P = 0.975, η² = 0.001) are not noticeable. Therefore, hypothesis H5a is not supported. For the fixation counts, the results of a repeat measures ANOVA analysis show that the main effect of PR type is appreciable (F(2, 68) = 8.374, P = 0.001, η² = 0.198), while the main effect of shopping conditions (F(1, 34) = 1.01, P = 0.322, η² = 0.029) and their interactive effect (F(2, 68) = 0.067, η² = 0.002) are not substantial. Thus, hypothesis H5b is not supported either.

**Testing the hypotheses related to context congruity and PR memory**

For the imposed condition, the influence of context congruity on consumers’ PR information memory was tested using an ANOVA analysis. The results indicate that different context congruity levels have a significant impact on consumers’ PR information memory (F(2, 68) = 20.629, P < 0.001, η² = 0.378). The results of a post hoc analysis are displayed in Table III, and hypothesis H6a and H7a are supported. Regarding the recognition of PRs, the ANOVA analysis results demonstrate that the influence of context congruity on consumers’ PR recognition is statistically significant (F(2, 68) = 7.466, P = 0.001, η² = 0.18). The results of a post hoc analysis as shown in Table III, and H6b and H7b are not supported.

For the free shopping (FS) condition, the results of a preliminary ANOVA analysis of information recognition accuracy show that the interactive effect of product category and PR type (F(6, 62) = 1.42, P = 0.221, η² = 0.121) and the effect of product category (F(3, 31) = 0.86, P = 0.472, η² = 0.077) are not significant. Next, to test the effect of context congruity on consumers’ PR information recognition without considering the product categories, the results of an ANOVA analysis show that there is no appreciable difference among the different context congruity levels (F(2, 68) = 1.312, P = 0.276, η² = 0.037). Therefore, hypotheses H8a and H9a are not supported. For the PR recognition, the results of a preliminary ANOVA analysis show that the interactive effect of product category and PR type (F(6, 62) = 0.826, P = 0.554, η² = 0.074) and the effect of product category (F(3, 31) = 1.561, P = 0.219, η² = 0.131) are not significant. The results of an ANOVA analysis of PR types without considering product categories demonstrate that there is a noticeable context congruity effect (F(2, 68) = 5.036, P = 0.009, η² = 0.129). The results of a post hoc analysis are shown in Table III, and H8b is partially supported and H9b is supported.
To compare the differences in PR memory between the imposed shopping (IS) condition and the FS condition, a repeated measures ANOVA analysis of PR information recognition accuracy was performed. The results show that the main effect of PR type (F(2, 68) = 9.95, P < 0.001, \eta^2 = 0.226) and its interactive effect by shopping condition (F(2, 68) = 5.745, P = 0.005, \eta^2 = 0.145) are significant. However, the main effect of shopping condition is marginally significant (F(2, 68) = 3.678, P = 0.063, \eta^2 = 0.098). The results of a simple effect analysis show that for the two shopping conditions, FS and IS have a significant difference only in the incongruent condition (FS > IS, P < 0.001). For the PR recognition, the results of a repeated measures ANOVA of PR recognition accuracy indicate that main effect of PR type (F(2, 68) = 6.024, P = 0.004, \eta^2 = 0.151) and its interactive effect by shopping condition (F(2, 68) = 6.387, P = 0.003, \eta^2 = 0.158) are appreciable. However, the main effect of shopping condition is not significant (F(1, 34) = 0.011, P = 0.918, \eta^2 = 0.0). The results of a simple effect analysis demonstrate that for shopping conditions, there are substantial differences between PR recognition accuracy in the FS condition and the IS condition only in the highly congruent context condition (FS > IS, P = 0.008).

**Discussion**
Ad-context congruity effects have been investigated in different contexts such as online newspapers and web portals. The findings in this line of research have been used to direct advertisement placement and to promote advertising effectiveness. Considering the extensive use of online PR agents and the common concern of PR effectiveness, in this paper, we introduce context congruity effects into RA effectiveness research. We investigate the influence of context congruity on consumers’ PR attention and memory using eye-tracking method and a recognition test, and the results are shown in Tables II and III, respectively.
Congruity and attention

With regard to the eye-tracking data, the results show that context congruity does have a noticeable influence on consumers’ attention to PRs in both the imposed and the FS conditions. Specifically, highly congruent PRs attracted more fixations (fixation duration and fixation counts) than lowly congruent and incongruent PRs in both shopping conditions. These findings are consistent with previous research that supports a positive relationship between ad-context congruity and advertising attention (Porta et al., 2013; Rieger et al., 2015). However, for both conditions, the fixations, including fixation duration and fixation counts, on PRs in the lowly congruent condition are not significantly higher than those in the incongruent condition. This result means that when the extent of congruity is low and not very noticeable, the effect of context congruity will not be obvious either. Regarding the different fixation effects under the imposed (low involvement) and the free conditions (high involvement), their impact on consumers’ eye fixations on PRs is not significant. This finding is consistent with the study of Porta et al. (2013) who first considered imposed and free conditions. The result shows that no matter whether consumers are purchasing what they consider to be interesting products or not, they are engaged in their shopping process, and context congruity can impact their attention to PRs.

Congruity, attention and memory

With regard to memory data, the results are contingent. For PR information recognition, in the imposed condition, the context congruity effect is significant. Here, we note that even though fixations on PRs between the lowly congruent condition and the incongruent condition are not significant, the memory performance regarding PR information is significantly different between them. These results are consistent with some previous research that also indicated that there is no definitively positive relationship between fixations and memory (Jeong and King, 2010; Porta et al., 2013). One possible explanation may be that even if consumers do not pay extra attention to a PR under the lowly congruent condition than under the incongruent condition, context congruity does lead to an improved effect on PR information recognition.

PR recognition, however, does not show similar results. In fact, the findings even contradict our suppositions. Specifically, PR recognition under the lowly congruent condition is significantly greater than that under the highly congruent condition, and PR recognition under the incongruent condition is marginally greater than that under the highly congruent condition. Three perspectives can be considered to interpret these findings. The first perspective concerns the visual processing difference between pictorial elements and textual elements of online ads. Online banner ads typically include the brand, some text, and a picture, which are treated differently in the course and extent of cognitive processing. Pictorial elements are regarded as superficial associations; consequently, they easily catch a consumer’s initial attention by relying on stronger and faster activation, while advertising information (e.g. information in text form) requires deeper processing and evaluation. Generally, it is the low-level aspects of stimuli that determine the attentive location (Rayner and Pollatsek, 1992), whereas cognitive processes can intervene in high-level mechanisms and allocate attention to informational aspects of stimuli to enrich and deepen one’s prior understanding (Rieger et al., 2015). The second perspective is the limited capacity theory (Lang, 2000), which indicates that humans can pay only limited attention to information, and their cognitive load is finite. Individuals will allocate their limited cognitive resources to each task based on the available resources and their intention to process the information. The last perspective is cognitive interference (Furnham et al., 2002), which implies that differences in the meaning of stimuli can trigger less involved/attentive consumers’ interests. For such a group of individuals, a contradiction between an ad and its context makes them feel more novel and compelling (Mandler, 1982).
Based on visual attention processing, for a PR, its pictorial elements are addressed at the early stage, which is a low-level, superficial and automatic process, while textual elements are processed and evaluated later, which is a high-level, deeper and cognition-guiding process. In the low involvement condition, the participants were required to browse products that were not very attractive to them compared with under the high involvement condition, where they could choose to view products based on their preferences. With low involvement, the participants did not exert too many cognitive resources on their shopping task, so they had extra cognitive capacity to notice and process the PRs and related information. Because processing textual elements requires more cognitive resources and is a deeper process, it can generate more memory traces, which contribute to a better memory. Thus, context congruity effects have a greater impact on PR information recognition than on PR recognition. This result explains the different recognition results for PRs and their information. Because processing pictorial elements is superficial and automatic without strong storage and processing, at a sufficient condition of cognitive capacity, the participants are easily attracted by PRs under lowly congruent or incongruent conditions because the differences between a PR and its context trigger less-involved individuals’ interest. Novel and compelling features of lowly congruent and incongruent stimuli induce cognitive interference, which overwhelms the context congruity effects. This is why the results for PR recognition do not complement but contradict to the context congruity effect.

Involvement and memory
In the FS condition, for PR information, the context congruity effect has no significant influence on recognition performance. The result occurs because under the high involvement condition, consumers devote most of their cognitive resources to their shopping tasks, and their limited cognitive capacity makes the deep processing of advertising information less possible, even though the context congruity effect exists. With regard to PR, memory performance under the highly congruent and lowly congruent conditions is significantly greater than that under the incongruent condition, while the differences between them are not statistically significant. The reason for this result is that the pictorial elements of PR are prior and automatically processed because they require fewer resources, and the context congruity effect can exert its influence to make congruent PR easier to recognize. Perhaps because PRs under the lowly congruent condition are still related to the viewing product, high involvement, automatic processing and the context congruity effect make the recognition of differences in PRs under the highly congruent condition and the lowly congruent condition not noticeable.

In terms of the distinction between the FS (high involvement) condition and the IS (low involvement) condition, the main effect of the shopping conditions was not significant, except that under the incongruent condition, better PR information recognition was found with the imposed online shopping condition, and under the highly congruent condition, greater PR recognition performance was identified in the FS condition. These results occur because in the imposed condition with sufficient cognitive capacity, it is easily to elicit cognitive interference with low involvement, resulting in more resources and processing allocated to the incongruent condition. However, in the FS condition with high involvement, the participants were highly engaged in their shopping task and exerted most of their cognitive resources and processing toward products that were closely related to their task.

Conclusion and implications
This paper innovatively investigates the context congruity effects of PRs by RAs based on real online shopping conditions on consumers’ visual attention and memory performance. Three types of PRs, namely highly congruent, lowly congruent and incongruent, and two degrees of involvement, namely high involvement and low
Involvement conditions were considered. Eye tracking matrices, fixation counts and fixation duration, and memory recognition of PR information and PRs were used to testify our hypotheses. The results of the gaze data show that context congruity effects can influence consumers' PR attention, but this effect is not moderated by issue involvement. The results of memory data show that PR recognition is influenced not only by context congruity effects but also by issue involvement.

Based on our results, context congruity does have a positive influence on visual fixation on PRs. Both high and low involvement conditions support this finding which indicates that no matter the degree of involvement, consumers are engaged in their shopping process. It is important to note that the differences between LC and incongruity are not substantial, meaning that the congruity level or intensity can influence its effect. Another significant finding is the fact that a PR is observed more does not necessarily indicate that it will be better recognized. The process of translation from visual attention to memory involves complicated cognitive resource allocation and cognitive processing, during which three aspects are important. First, different forms of stimulus (pictures or text) will induce different cognitive processing mechanisms. Second, humans have limited cognition capacity, and issue involvement can intervene in the process of allocating cognitive resources and processing stimuli. Third, when cognitive capacity is sufficient, contextual priming and cognitive interference will compete with each other.

Theoretical contributions
Our study significantly contributes to deepening the understanding of how context congruity can influence consumers' PR attention and memory. The main theoretical achievements can be summarized as following. First, introducing context congruity effects to investigate the effectiveness of online PRs by RAs not only provides important theoretical contributions to research on PR effectiveness, but also enriches the application of context congruity effects. Second, our findings suggest the relationship between visual attention and memory is not definitely positive. Third, to interpret the complex translation from attention to memory, we propose an analysis methodology that considers stimulus attributes, issue involvement, the degree of cognitive capacity, contextual priming and cognitive interference.

Approach contributions
In this study, we used eye-tracking method to investigate the effects of online PRs by RAs. Compared to traditional empirical analyses, eye-tracking method can help provide more insights into consumers' behavior. For instance, eye-tracking equipment can track both where an individual is looking at any given time and the sequence in which their eyes shift from one location to another. All these fixation and saccade information are very important to RA designers and marketers. Therefore, this paper gives a good example in how to use machine observation methods such as eye tracking to examine consumer behavior, and further to aid RA design and market operation.

Managerial implications
In online marketing, our findings also have significant managerial and practical implications for RA designers, marketers and advertisers. First, no matter the degree of involvement, when starting their online shopping, consumers' goal engagement could prompt their attention to PRs or banner ads, and their visual fixation is affected by context congruity effects. Although attention does not necessarily guarantee the formation of memory, memory is definitely not generated without fixation on ads. Thus, when RA designers and marketers plan to present a PR or to issue a new advertisement, they should
consider the context of its placement based on context congruity effects in order to grab consumers’ attention. Also, they can achieve precision marketing by knowing consumers’ preference because it will influence consumers’ involvement in their online shopping and further influence the effectiveness of PRs by RAs. Second, due to the different processing mechanisms for different elements, when designing a PR or an advertisement, they should consider consumers’ involvement or goal orientation. For low-involved or goal-unspecific web page browsing, congruity effects can significantly enhance memory of PRs or advertising information. At this time, using text-dominated advertisement to increase consumers’ product impressions is better, while for high-involved or goal-specific online behavior, congruity effects can increase the recognition of a pictorial element, and therefore a picture-dominated form can help consumers construct the product image.

Limitations and further research
Although both theoretical and practical findings have been achieved in our study, some limitations still need to be discussed. First, the manipulation of issue involvement was not very strict. A better way would be to refer to an involvement scale to quantitatively evaluate participants’ issue involvement such as in Segev et al. (2014). Moreover, even though we considered three primary elements of advertisements – brands, text and pictures – unlike previous researchers who focused on brand memory, in our study, we ignored brand memory. However, brands are very important product-related information and have significant influences on consumers’ purchase decisions. On the other hand, regarding the pictorial and textual elements of an ad, we explored the difference in their recognition, but we did not show any distinction in the visual attention paid to each due to our setting of a whole PR as our AOI. Future investigations should take product brands into consideration and should distinguish AOIs into brands, pictures and text to observe consumers’ specific visual fixations on PRs and the influence of congruity effects on them.

References

**Corresponding author**
Zhong Yao can be contacted at: iszhyao@buaa.edu.cn

---

For instructions on how to order reprints of this article, please visit our website:
[www.emergalgrouppublishing.com/licensing/reprints.htm](http://www.emergalgrouppublishing.com/licensing/reprints.htm)
Or contact us for further details: permissions@emeraldinsight.com
Mood metadata on Chinese music websites: an exploratory study with user feedback

Xiao Hu
Shenzhen Institute of Research and Innovation, University of Hong Kong, Hong Kong
Christy W.L. Cheong
Macao Polytechnic Institute, Macao, China
Siwei Zhang
Faculty of Education, University of Hong Kong, Hong Kong, and
J. Stephen Downie
School of Information Sciences, University of Illinois, Champaign, Illinois, USA

Abstract

Purpose – Music mood is an important metadata type on online music repositories and stream music services worldwide. Many existing studies on mood metadata have focused on music websites and services in the Western world to the exclusion of those serving users in other cultures. The purpose of this paper is to bridge this gap by exploring mood labels on influential Chinese music websites.

Design/methodology/approach – Mood labels and the associated song titles were collected from six Chinese music websites, and analyzed in relation to mood models and findings in the literature. An online music listening test was conducted to solicit users' feedback on the mood labels on two popular Chinese music websites. Mood label selections on 30 songs from 64 Chinese listeners were collected and compared to those given by the two websites.

Findings – Mood labels, although extensively employed on Chinese music websites, may be insufficient in meeting listeners' needs. More mood labels of high arousal semantics are needed. Song languages and user familiarity to the songs show influence on users' selection of mood labels given by the websites.

Practical implications – Suggestions are proposed for future development of mood metadata and mood-enabled user interfaces in the context of global online music access.

Keywords Chinese music website, Global music access, Mood metadata, Music mood, Online music access, Online music repositories

Paper type Research paper

Introduction

With the increasing popularity of digital music, online music services and music websites have become one of the dominant resources for people to seek and consume music information (Hu, Lee and Wong, 2014). Among the various metadata types that describe music, mood[1] has been widely used in many music repositories, online music streaming services and music websites including Spotify, Pandora, Last.fm, Allmusic.com (hereafter Allmusic) and APMmusic.com (hereafter APM) (Hu, 2010; Lee and Waterman, 2012; Lee and Price, 2016). The affective aspect of music is arguably the most important reason people engage with music (Juslin and Laukka, 2004). There are two mood-related concepts in the literature: one is the mood expressed by music pieces which is often used to describe the music; the other is the listeners' internal mood aroused by music which can be a goal of music.

This study is partially supported by the National Natural Science Foundation of China: 61703357 and a Seed Research Fund from the University of Hong Kong. The authors thanks Xueming Ann Wang for her assistance in collecting part of the data.
This study focuses on the former, with the purpose being to investigate the efficacy of mood labels on Chinese music websites in describing music information and in helping users access and select music of certain mood.

Music mood is an obvious criterion by which listeners can organize their personal music collections (Hagen, 2015), but it can be highly subjective, rendering it an intriguing and challenging concept in music information studies (e.g. Lamere, 2008; Yang and Chen, 2012). Although it is increasingly acknowledged that music mood is culturally, linguistically and socially sensitive (Singhi and Brown, 2014), most studies to date have focused on Western music, Western music websites and Western listeners (Hu and Downie, 2007; Serra, 2011; Lee et al., 2012). In the literature, there is a lack of understanding on how music mood is used on non-Western music websites. This Western-centric research trajectory has been criticized in recent years, as music consumption has become increasingly globalized and geographic and cultural boundaries have blurred (Serra, 2011; Xu et al., 2015; Hu and Lee, 2016).

This study contributes to the literature on culturally diversified online music access from a Chinese perspective, which is of theoretical and practical importance not only because of the large listener population China possesses (IFPI, 2015) but also because of the many unique features of Chinese music (Hu and Lee, 2016). While having a long history and a tradition remarkably different from Western music (Fung, 2013), Chinese music is heavily influenced by Western pop culture, which was introduced to the Chinese population following the country’s economic reform and the Chinese Government’s “Open Door Policy” nearly four decades ago.

As one of the first attempts to examine mood metadata on Chinese music websites, this study is exploratory in nature, aiming to further our understanding of three research questions:

**RQ1.** Which mood labels are used on major Chinese music websites?

**RQ2.** To what extent do Chinese listeners agree with the mood labels provided by the websites?

**RQ3.** Are mood labeling patterns on major Chinese websites varied by song language (i.e. Chinese songs vs English songs)?

While answers to the first question help obtain a general picture of mood metadata of music in a Chinese context, the second aims to evaluate whether the mood labels currently used are sufficient in supporting users to access online music by mood. Studies on mood metadata on Western online music services have identified a vocabulary gap between mood labels or taxonomies given by music information services and those of the users (Hu, 2010; Lee et al., 2012). The vocabulary gap has been well recognized in information science (Mai, 2011; Wichowski, 2009; Noruzi, 2006) and can be even wider on music mood, due to the subjectivity and cultural dependence of music mood perception (Lee and Hu, 2014; Singhi and Brown, 2014; Egermann et al., 2015). As Chinese listeners also consume a significant amount of Western music (IFPI, 2015), the third question attempts to investigate whether songs in a foreign language would have different patterns of user feedback with regards to mood labeling and thus should be treated differently on the music websites. Findings of this study can shed light on the current use of mood metadata on Chinese music websites and provide insights to improve listener access. They can also benefit future research on music mood of other non-Western music and cross-cultural music access in the online environment.

**Related work**

Studies on music metadata in online music repositories and services

Traditionally, music digital libraries and repositories use bibliographic metadata (e.g. title, composer, singer, language) to organize music materials. In the last decade, due to increasing user demands, music websites and service providers have started to develop
non-traditional metadata such as genre, mood, and usage to describe music pieces and to facilitate online music access. Due to their popularity, online music services provide a rich resource for studying newly emerging music metadata. One of the pioneering works on mood metadata of music was by Hu and Downie (2007). They collected 179 mood labels and their associated songs from Allmusic. Through clustering analysis, they derived a five-cluster mood taxonomy from these labels. The taxonomy has been widely used in music mood-related studies (for instance, Yang and Chen, 2012; Patra et al., 2013; Singhi and Brown, 2014), but it was developed from labels on Western music and to what extent it can be applied to non-Western music is still largely unknown. While the labels on Allmusic were given by professional editors (and thus can be referred to as a controlled vocabulary), social tags contributed by users (also called folksonomy) can also be exploited to study music metadata. Although not focusing specifically on music mood, Goh et al.’s (2009) study on the effectiveness of social tagging on del.icio.us found that “music” tags, if well-defined and mutually understood, can potentially enhance resource discovery. Hu et al. (2009) studied mood-related social tags posted on Last.fm, one of the most popular social tagging websites for Western music, and derived 18 mood categories frequently mentioned by users, which included “calm,” “sad,” “happy,” and “romantic.” Again, Last.fm contained mostly Western music and its user base was predominantly from Western countries (particularly in Europe) (Lamere, 2008). For non-Western music, Lee et al. (2013) investigated genre labels on Korean music websites for the purpose of studying genres of Korean pop music. They collected genre labels on eight popular Korean commercial websites and identified those unique to Korean music (e.g. Trot). However, to the best of our knowledge, there are no existing studies of controlled or folksonomic mood labels on non-Western music websites. Existing studies on the mood of non-Western music often adopt mood taxonomies created for Western music or borrowed mood categories from models in psychology. For instance, Patra et al. (2013) used the aforementioned five-cluster mood taxonomy and Santos and Silla (2015) adapted the Affective Circumference Model when constructing a Latin music database.

Music mood representations
In the literature of Music Information Retrieval (MIR) and Music Psychology, music mood is often represented by two kinds of models: categorical and dimensional. Categorical models use a set of discrete labels to represent music moods. The most classical categorical models include Hevner’s Adjective Circle which consists of eight mood categories placed in a circle with a set of labels in each (Hevner, 1936). The labels on Allmusic are also an example of a categorical model. Dimensional models represent music moods with numerical values in a small number of dimensions. Among the existing dimensional models, Russell’s two-dimensional model is widely adopted (Yang and Chen, 2012). It contains the dimensions of arousal (measuring level of energy) and valence (measuring level of pleasure) that split the mood space into four quadrants as shown in Figure 1 (Russell, 1980). Frequently seen mood labels are placed in the space according to their corresponding arousal and valence values. From this, one can see that dimensional and categorical models, although mostly used separately, are connected to each other.

While both kinds of representation models have their own advantages, categorical models are often deemed more user-friendly and intuitive as music moods are represented with natural language labels. Consequently, almost all music websites use categorical models and develop mood taxonomies of discrete labels to organize their music collections. However, categorical models are not immune from vocabulary issues. In most languages there can be hundreds of words with mood-related meanings and the differences between them can be subtle (e.g. “brooding” vs “gloomy”). Words selected by music service providers (i.e. controlled vocabulary) may not be understood or used in the same way by website users, and this vocabulary gap can become more complex when words are translated between
languages (Yang and Hu, 2012). In the current study, mood labels collected from major Chinese music websites fall into categorical models, whereas Russell’s (1980) model, as a representative dimensional model, will be referred to in the analysis.

**Cross-language and cross-cultural music mood labeling**
Music listening crosses the borders of countries (Lee et al., 2013), and a number of studies have investigated how listeners perceive the mood of music in a foreign culture and language. Most of these studies compared two or more groups of listeners with different cultural backgrounds, in terms of their mood judgments on a set of songs with the same origin of one of the listener groups. Studies have explored Greek and non-Greek listeners on the mood of Greek songs (Kosta et al., 2013), Canadian and Chinese listeners on English songs (Singhi and Brown, 2014), Korean and American listeners on K-Pop songs (Hu, Lee, Choi and Downie, 2014), American, Korean and Chinese listeners on English songs (Lee and Hu, 2014), and Chinese and American listeners on Chinese songs (Hu and Lee, 2016). All these studies generally concluded that listeners who had the same cultural origins as the songs reached higher agreements on their mood judgments than listeners from other cultures. Although these studies interpreted results in various ways, several common factors can be discerned, including the language abilities of listeners in understanding lyrics in a foreign language, familiarity to the songs and cultural familiarity.

There has been scant research comparing mood judgments by the same group of Chinese listeners on Chinese and English songs. To address this research disparity, the current study seeks to gain some insight into how Chinese listeners agree with mood labels of Chinese and English songs on Chinese music websites.

**Study design**

*Analysis of mood labels on Chinese music websites*
Domain analysis, a method widely used in information science to understand given domains, was conducted to understand the space of mood labels and how often the labels

---

**Figure 1.**
Russell's two-dimensional model of music mood

<table>
<thead>
<tr>
<th>Arousal</th>
<th>Valence</th>
</tr>
</thead>
<tbody>
<tr>
<td>alarmed</td>
<td>satisfied</td>
</tr>
<tr>
<td>afraid</td>
<td>content</td>
</tr>
<tr>
<td>tense</td>
<td>serene</td>
</tr>
<tr>
<td>distressed</td>
<td>calm</td>
</tr>
<tr>
<td>anxious</td>
<td>at ease</td>
</tr>
<tr>
<td>relaxed</td>
<td></td>
</tr>
<tr>
<td>content</td>
<td></td>
</tr>
<tr>
<td>happy</td>
<td></td>
</tr>
<tr>
<td>pleased</td>
<td></td>
</tr>
</tbody>
</table>

Source: Russell (1980, p. 1168)
were applied to songs. Hjørland (2002) summarized the approaches to domain analysis including those identifying terminology and vocabularies from empirical data. In this study, the domain concerned was music mood and the empirical data were mood labels from six popular Chinese commercial music websites (iResearch, 2015). From each site listed below, we included all mood labels shown on the main catalog page:

- Baidu Music: http://music.baidu.com/
- Xiami Music: www.xiami.com/
- NetEase Cloud Music: http://music.163.com/
- Kuwo Music: www.kuwo.cn/
- Kugou Music: www.kugou.com/
- 9ku Music: www.9ku.com/

We downloaded the pages listing the songs under each label and used scripts to extract the song titles and the artist names by batch. The number of songs associated with each label on each website was reported and compared. The song titles collected from two of the listed websites were then further used to develop the online listening test as detailed below.

User feedback on mood labels
To answer RQ2 and RQ3, an online listening test was designed. For this purpose, we focused on mood labels on two of the listed websites, Kugou and Kuwo. They were selected because of their prominent status in the industry and their large user groups. The top two most popular music websites in China (iResearch, 2015), Kugou and Kuwo had 10.3m and 6.5m monthly unique visitors during the first eight months of 2015. The online listening test aimed to explore how many mood labels given by the websites would be selected by users. It was administered in simplified Chinese in line with the demographics of the targeted participants and the mood label data obtained. In the test, the participants were asked to listen to a number of songs and select the mood label(s) they found appropriate in describing the mood expressed by each song. An open-ended question was included to solicit opinions on the mood expressed by music, rather than listeners’ internal moods. Although the latter may influence their selection of music (i.e. what they wanted to listen to, Çano et al., 2017), no convincing published evidence has shown that listeners’ internal mood states influence their judgments on the mood expressed by music, or their ability to make reasonable judgments as to whether certain mood labels might or might not be useful or appropriate. As such, listeners’ internal mood states were not measured in this study. Thirty songs were included, as music listening takes time and more songs would make the listening test too time-consuming to be practical. To answer RQ3, we included 15 Chinese songs and 15 English songs in the listening test.

Previous studies have found that listeners in East Asian cultures tend to choose from existing music mood labels, rather than suggesting their own (Hu and Lee, 2016). Having more options is thus more advantageous in soliciting participants’ opinions, and using labels that were applied to each song by the two websites reflects actual practice in the real world. To ensure participants would have more options to choose from, we selected test songs that were given the largest number of mood labels on both websites and had minimum label overlaps across websites. In other words, the selected songs were those with the largest number of unique mood labels across the two websites. For each selected song, we included as options all labels provided by both websites, resulting in 3 to 13 unique labels presented across songs. An “other” option was also included allowing participants to suggest their own labels when the given options were deemed insufficient. The songs and the labels were arranged in a random sequence to avoid possible biases. As previous
studies have found that one song can express a mixture of different moods (Lee et al., 2012; Hu and Lee, 2016), each participant could select multiple mood labels for each song. Questions concerning the participants’ familiarity with the given songs were included to measure the possible influence of familiarity on their song mood perception. Figure 2 shows a screenshot of the test.

The perceived mood of a song is often culturally dependent, context specific and lyric related (Santos and Silla, 2015; Singhi and Brown, 2014). To ensure the applicability of the findings to the targeted cultural group, the participants were all native speakers of Mandarin living in mainland China. Invitations were distributed through popular social media platforms in mainland China, such as WeChat and Weibo. Participation was completely voluntary.

Results
Overview of mood labels on Chinese music websites
From the six major Chinese music websites, we collected 41 unique mood labels from the first level of mood categories. Table I lists the mood labels that appeared on at least two websites as well as the number of songs associated with each of them (as of May 2015). The mood labels are sorted by frequency of appearance across the websites. There were a number of labels that appeared only in one of the six websites, including “waiting” (等待, Xiami), “healing” (治愈, NetEase), “mournful” (悲伤, Xiami), “exciting” (兴奋, NetEase), “tuneful” (好听, 9Ku), “soft” (轻柔, 9Ku), “clean” (干净, 9Ku), “fresh” (清新, NetEase), and “sexy” (性感, NetEase).

Among the mood labels, “calm” and “lonely” were present across all websites. Also appearing on five out of the six websites were the labels “melancholic,” “joyful,” “wistful,” “encouraging” and “sweet.” In terms of the number of songs associated with each mood label, the labels “sad” and “calm” were the most popular, followed by “melancholic” and “passionate.” Interestingly, “sad” and “calm” were both associated with the largest

Figure 2.
A screenshot of the online listening test (English translations added for illustration purposes only)
numbers of songs, and are placed in the 3rd or 4th quadrants in Russell’s model (Figure 1), with low arousal. On the other hand, there is a lack of radical moods (i.e. 2nd quadrant moods with low valence but high arousal, such as “aggressive,” “angry” and “tense”), in comparison to the presence of such moods on Western websites (Yang and Hu, 2012). These observations are in accordance with previous studies on mood labels of Chinese music. Yang and Hu (2012) manually annotated 500 Chinese songs with a list of 36 mood labels and found the most used labels were “tender” and “wistful” while very few songs were regarded as “aggressive.” Hu and Lee (2016) also found that listeners from Hong Kong and the USA often opted for the category of “wistful, bittersweet” when they were asked to judge the mood of a set of Chinese songs in six different styles. As these previous studies involved limited numbers of Chinese songs, this study, with a fairly complete domain analysis of major music websites in China, helps verify these findings with empirical data on a larger scale.

**Mood labels on Kugou and Kuwo**

Table II shows the mood labels collected from Kugou and Kuwo (as of May 2015). It is notable that the moods embodied in a majority of these labels are not radical emotions. Similarly, there were more labels carrying low-arousal features (e.g. “calm,” “sad”) on both websites. These findings seem to be in line with those of Wu and Xie (2008) and Yang and Hu (2012) that introverted emotions and emotion expressions with soft terms are more appreciated in Chinese culture than in Western cultures (McCrae *et al.*, 1996).

Table III provides statistics of the data collected from the two websites. Chinese songs dominated both websites. The number of song titles available on both websites
(referred hereafter as “overlapping songs”) was 1,080 in total, among which 973 were in Chinese and 107 in English. These song titles, as stated in the previous section, contributed to the design of the listening test.

**Listeners’ responses**

In total, 64 responses completed all 30 songs. There were 19 male participants and the average age of the male and female participants was 25.6 (SD = 1.116) and 25.0 (SD = 3.467), respectively. This age group is often involved in MIR user studies as they are major music consumers (Lee and Cunningham, 2013; iResearch, 2015). The sample sizes in previous studies involving music listening often range from 20 to 35 participants from each cultural group (Kosta *et al.*, 2013; *Lee et al.*, 2013; Singhi and Brown, 2014). The sample size of this study thus conforms to the norm of the field and can help generate useful findings for this exploratory study of music moods on Chinese websites.

**Selection ratios**

In examining listeners’ feedback on song moods, we calculated the label selection ratio of each song. The metric of song label selection ratio is defined as the ratio of the total number of labels selected by all the participants for a song to the theoretically maximum number of labels the participants could select (i.e. the number of labels listed for the song multiplied by the number of participants). This metric implies the level of agreement in mood labeling between the participants and the websites (since the listed labels were from the websites). A song label selection ratio closer to zero suggests a lower level of agreement, whereas a value closer to one indicates nearly perfect agreement. Figure 3 presents the label selection ratio of each song. The overall selection ratio across all songs was 0.334. The fact that the participants chose the labels given by the websites for only about one third of the time suggests the existence of a vocabulary gap between the websites and users.
Table IV shows the song label selection ratios averaged across all songs by website and by song language. An independent $t$-test was conducted to compare the song label selection ratios between the two websites and the result showed a significant difference ($p = 0.001$). However, the difference was found for English songs (0.30 vs 0.49, $p = 0.003$), but not for Chinese songs (0.30 vs 0.34, $p = 0.165$). In particular, the song label selection ratio on English songs was much higher than that on Chinese songs on Kugou (0.49 vs 0.34), which might be related to the fact that more mood labels were given by the websites to Chinese songs (5.27 labels per song on average) than English songs (2.33).

The most and least selected labels

The selection ratio of a label across songs is calculated as the ratio of the total number of times the label was selected by the participants and the maximum number of times the label can be selected (which is the number of songs with this label multiplied by the number of participants). Table V shows the selection ratios across songs of all labels and their source websites. To help characterize the mood labels, we also present in Table V the signs of valence and arousal values corresponding to each label based on Russell’s model (Figure 1). For example, “sweet” can be regarded as a mood with positive valence (i.e. “pleasant”) but the sign of its arousal value can vary. A “sweet song” can be quiet or aroused, and thus the arousal sign is left undefined for “sweet.” As another example, songs that feel “encouraging” are usually “pleasant” and “arousing,” and thus both valence and arousal are positive for the “encouraging” mood label. From Table V we can see some commonality among mood labels with above average selection ratios (i.e. > 33.4 percent). Five out of the six labels that appeared on both websites had above average selection ratios, as did mood labels with relatively clear indications of the arousal dimension (e.g. “melancholic” and “passionate”). In contrast, mood labels with lower selection ratios usually do not have clear semantics on the arousal dimension and are mostly from one website (e.g. “romantic” and “moving”). These results, that mood labels

<table>
<thead>
<tr>
<th>Website</th>
<th>Mean</th>
<th>SD</th>
<th>Avg. number of given labels per song</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kuwo</td>
<td>0.30</td>
<td>0.09</td>
<td>2.67</td>
</tr>
<tr>
<td>English songs</td>
<td>0.30</td>
<td>0.10</td>
<td>1.60</td>
</tr>
<tr>
<td>Chinese songs</td>
<td>0.30</td>
<td>0.08</td>
<td>3.73</td>
</tr>
<tr>
<td>Kugou</td>
<td>0.41</td>
<td>0.15</td>
<td>3.80</td>
</tr>
<tr>
<td>English songs</td>
<td>0.49</td>
<td>0.16</td>
<td>2.33</td>
</tr>
<tr>
<td>Chinese songs</td>
<td>0.34</td>
<td>0.08</td>
<td>5.27</td>
</tr>
</tbody>
</table>

Figure 3. Song label selection ratio across songs
with clear indication of arousal were selected more frequently, seem to be consistent with findings that the arousal dimension of music mood is more agreeable by listeners (Kim et al., 2010; Yang and Chen, 2012; Hu, Lee and Wong, 2014).

Possible influence of familiarity with and language of the songs

It is recognized in the literature that users tend to agree more on the mood of songs with which they are familiar (Kosta et al., 2013; Lee and Hu, 2014; Hu and Lee, 2016). Therefore, we measured the familiarity level of the participants with the songs in the listening test. It is assumed that a song is “familiar” to a listener if the listener has heard it before and “strongly familiar” to a listener if the listener is able to recall the song’s name and singer(s).

We calculated the correlation between the song label selection ratios and the average familiarity level with the songs averaged across all participants, and the result reveals a significantly and moderately negative correlation ($r = -0.555, p = 0.007$). In other words, the more familiar the participants were with a song, the fewer mood labels they would select for that song. A possible reason might be that users may have a clearer understanding of a song that they are familiar with and thus be more certain and focused about the song’s mood.

A closer look at familiarity levels also reveals a large difference between English and Chinese songs. On average, each Chinese song had been heard by 36.7 (57.4 percent) participants before the online listening test, whereas each English song had only been heard by 19.7 (30.8 percent) participants. The difference was even larger on the average number of participants who could name each song or its singer: 26.0 (40.6 percent) for the Chinese songs and 7.6 (11.9 percent) for the English songs. Differences on music mood perceptions on songs in different languages have been often explained in the literature as related to listeners’ familiarity to the songs and the cultures from where the songs originate (Egermann et al., 2015; Kosta et al., 2013). As the listeners were more familiar with Chinese songs and Chinese culture, they tended to have more consistent mood judgments on them, which at least partially explains the higher label selection ratios among English songs than those among Chinese songs.

### Mood metadata on Chinese music websites

<table>
<thead>
<tr>
<th>Mood label</th>
<th>Selection ratio (%)</th>
<th>Number of songs</th>
<th>Source websites</th>
<th>VA quadrant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweet (甜蜜)</td>
<td>79.69</td>
<td>1</td>
<td>Both</td>
<td>+V</td>
</tr>
<tr>
<td>Encouraging (励志)</td>
<td>55.21</td>
<td>3</td>
<td>Both</td>
<td>+V+A</td>
</tr>
<tr>
<td>Melancholic (忧郁)</td>
<td>50.84</td>
<td>13</td>
<td>Kuwo</td>
<td>−V−A</td>
</tr>
<tr>
<td>Wistful (思念)</td>
<td>50.09</td>
<td>18</td>
<td>Kugo</td>
<td>−A</td>
</tr>
<tr>
<td>Miserable (难过)</td>
<td>43.40</td>
<td>9</td>
<td>Kugo</td>
<td>−V−A</td>
</tr>
<tr>
<td>Passionate (激情)</td>
<td>42.19</td>
<td>2</td>
<td>Both</td>
<td>+V+A</td>
</tr>
<tr>
<td>Lonely (寂寞)</td>
<td>38.23</td>
<td>15</td>
<td>Both</td>
<td>−V</td>
</tr>
<tr>
<td>Calm (安静)</td>
<td>38.19</td>
<td>18</td>
<td>Both</td>
<td>+V−A</td>
</tr>
<tr>
<td>Nostalgic (思乡)</td>
<td>36.72</td>
<td>4</td>
<td>Kuwo</td>
<td>−A</td>
</tr>
<tr>
<td>Warm (温暖)</td>
<td>31.97</td>
<td>13</td>
<td>Kuwo</td>
<td>+V</td>
</tr>
<tr>
<td>Sad (忧伤)</td>
<td>31.93</td>
<td>16</td>
<td>Kugo</td>
<td>−V−A</td>
</tr>
<tr>
<td>Tearful (想哭)</td>
<td>28.91</td>
<td>12</td>
<td>Kugo</td>
<td>−V</td>
</tr>
<tr>
<td>Romantic (浪漫)</td>
<td>25.20</td>
<td>8</td>
<td>Kugo</td>
<td>+V</td>
</tr>
<tr>
<td>Happy (幸福)</td>
<td>22.17</td>
<td>16</td>
<td>Kugo</td>
<td>+V+a</td>
</tr>
<tr>
<td>Moving (感动)</td>
<td>22.10</td>
<td>7</td>
<td>Kugo</td>
<td>NA</td>
</tr>
<tr>
<td>Tired (疲倦)</td>
<td>15.28</td>
<td>9</td>
<td>Kuwo</td>
<td>−V−A</td>
</tr>
<tr>
<td>Depressed (郁闷)</td>
<td>15.00</td>
<td>10</td>
<td>Kuwo</td>
<td>−V</td>
</tr>
<tr>
<td>Joyful (开心)</td>
<td>13.44</td>
<td>15</td>
<td>Both</td>
<td>+V+A</td>
</tr>
</tbody>
</table>

**Notes:** 
- While “happy” is marked as +V+A in Figure 1, in Chinese the word 幸福, although translated as “happy” here, also resonates with the meaning of “content” which can have a negative arousal in Figure 1; 
- In Chinese, the word 郁闷 has meanings of both “depressed” and “frustrated”

**Table V.** Selection ratios of the mood labels
User comments

There were 22 participants (34 percent of participants) who provided open-text comments during the online listening test. The general themes include:

(1) the mood labels presented were not sufficient in covering all possible moods of the music (ten participants);

(2) more mood labels with negative valence and positive arousal (e.g. “aggressive,” “angry”) were needed (five participants);

(3) some of the labels were rather similar, and thus the participants tended to choose all the similar labels (four participants);

(4) the songs in the online listening test tended to have the feature of lyrical songs with a soft melody, and the mood labels presented did not include enough moods in positive valence and high arousal (e.g. “excited”) (four participants); and

(5) the labels for English songs were too limited (four participants).

Discussions

Mood labels on Chinese music websites

As with their Western counterparts, most Chinese music websites use mood as a major criterion to organize music. However, there are fewer mood labels on Chinese websites, compared to those on Western music websites or services. For instance, Allmusic has about 180 mood labels and APM has 158. In contrast, each of the six Chinese websites surveyed in this study contained fewer than 30 mood labels, and in total there were 41 unique labels. As also suggested by some of the participants in this study, the mood labels given by the Chinese websites were insufficient, seemed mostly mild (with low arousal) and short of terms of high arousal (e.g. “aggressive,” “excited”). This skewed coverage might be due to characteristics of Chinese music as well as to possible censorship imposed on the Chinese music industry (e.g. Shadbolt, 2014). Songs with negative and destructive moods are likely to express dissatisfaction toward reality and might have been censored during the publishing stage or even self-censored by artists in the creation stage (Mittler, 1997).

Significantly, none of the Chinese websites provided definition of the mood labels. It is left to users to interpret the meanings of the labels, which may not always be straightforward. As some participants commented, the meanings of some labels are indistinguishable (e.g. “sad” and “melancholic,” “joyful” and “happy”), and they ended up selecting all similar labels. Some labels on Chinese websites are not or only marginally related to mood (e.g. “fashion,” “aesthetic,” “tuneful,” etc.) and to complicate the problem further, different websites use different labels, sometimes with very similar meanings, such as “忧伤” (sad) in Kugou and “伤感” (melancholic) in Kuwo. The commercial companies behind the websites may not necessarily have the incentive to standardize labels with other companies, but it can also be hard even within one company. Kugou and Kuwo were owned by the same company which was recently merged with QQ music (Soo, 2016). However, at the time of writing the three websites are still using three different sets of mood labels.

There are similarities between the mood labeling practices of Chinese and Western websites. As shown in Table I, the mood labels with most songs on the selected Chinese websites are “calm,” “sad” and “melancholic.” Similarly, previous studies on Last.fm found that mood tags in these mood categories were associated with more songs than other mood tags (Hu et al., 2009). There is a lack of context in listing the mood labels on all websites, irrespective of whether the website is Chinese or Western. As mentioned before, besides the mood expressed by the music, listeners may have their own emotional state and may want to select music accordingly (Çano et al., 2017). Only listing mood labels without indicating whether it refers to music mood or
listeners’ mood may cause confusion for users. Therefore, music websites may differentiate these two concepts of mood by, for example, adding a search box or field of “what is your mood now?” in addition to the existing labels specifically for music mood.

User feedback on mood labels
The 33.4 percent average selection ratio in the listening test indicates that only about one third of the mood labels given by the two websites were selected by the users, despite the fact that the users could freely choose as many labels as they wished. There is clearly a gap between users’ mood feedback and mood labels provided by the websites. As labels on websites are generated and controlled by website editors, this gap reflects tension between a controlled vocabulary and user’s vocabulary (i.e. folksonomy), which has been well recognized in other domains of information science (Mai, 2011; Noruzi, 2006). To bridge this gap, researchers have started to develop music mood taxonomies from social tags (Laurier et al., 2009; Yang and Chen, 2012), with popular online music services in Western countries making use of user generated tags to help organize and search for music, as exemplified by Last.fm and Pandora. Chinese music websites may also incorporate user tags into the labeling of music mood and make them a beneficial complementary to the controlled vocabulary already in existence.

The finding that the selection ratios were higher for mood labels with clearer semantics on arousal dimension (Table V) is consistent with studies on listeners from other countries (Kim et al., 2010; Yang and Chen, 2012; Hu, Lee and Wong, 2014). When labeling songs with mood, either manually or automatically, Chinese music websites could assign high confidence to mood labels with clear semantics in arousal and/or to mood annotations on the arousal dimension.

Implications to MIR in the global context
Differences between mood labels on Chinese and Western websites raise significant challenges in facilitating music access to global music digital libraries by listeners from differing cultures. Interoperability is an important requirement for digital libraries so that more resources can be connected and provided to users with less cost. To reconcile and bridge the two mood labeling systems in Chinese and Western websites, a crosswalk of vocabulary is necessary and can be constructed by integrating categorical and dimensional models of mood representation. By providing two dimensions (i.e. valence and arousal), Russell’s model has minimum dependence on an accurate interpretation of the meanings of mood labels. This characteristic can be particularly useful in bridging users from one mood space of 41 labels (Chinese websites) to another with 180 labels (allmusic.com), and vice versa. Incorporating dimensional models can thus alleviate the dependence on language, facilitating cross-language and cross-cultural music access in the global context.

Similarly, dimensional models can also serve as a middle ground to connect music mood vocabularies between users and service providers, as well as those between different companies. Although Russell’s model has been extensively applied in music retrieval and music psychology studies, it has yet to be adopted in real-world end-user facing services. It is thus suggested that future music websites could consider mapping their mood taxonomies into a well-accepted dimensional model such as Russell’s model, which can improve the interoperability of their websites and also help make their websites more user friendly.

Another enduring issue in music mood is the definition of mood labels. Few, if any, websites give definitions of mood labels they use, possibly due to the difficulty in defining music mood. In categorical models such as Hevner’s Adjective Circle (1936) and the five mood clusters (Hu and Downie, 2007), a category is often defined with a set of mood-related words, rather than a single word, to mitigate the ambiguity and subjectivity in understanding word meaning. In dimensional models, it is also challenging to define the intensity of music mood represented by each dimension. One way to resolve this is to let music rather than words be

Mood metadata on Chinese music websites
the medium of translation. That is, to provide representative music examples to demonstrate
the meaning of the mood. Examples have been used in training annotators or participants in
existing studies on music mood recognition (e.g. Hu and Downie, 2007; Yang and Chen, 2012)
where multiple music clips with clear indications of certain mood were presented, to show
users in a concrete way, what music in certain moods would sound like.

The different number of labels given to Chinese and English songs as well as the
different selection ratios from users suggest that the language of the songs matters.
More mood labels are needed for English songs, as suggested by participants in this study.
As lyrics can play an important role in music mood (Lee et al., 2012), not being able to
understand the lyrics in English songs might have hindered website editors from applying
more mood labels. One way to improve this situation is to borrow mood labels given by
Western music websites, especially those with clearer semantics on the arousal dimension
as they are more likely to be selected by users.

The negative correlations between song label selection ratios and users’ familiarity to the
songs suggest that the Chinese music websites could consider making use of users’ listening
records and trim down mood labels of music that has been frequently listened to (and thus
more familiar to the users). Perhaps when a song is newly released, more mood labels can
help promote the songs to more listeners, especially for English songs that are in general not
as familiar to the listeners as Chinese songs. When users become familiar with a song, they
may develop more focused opinions on the moods of the song, and thus less relevant labels
could then be removed.

Conclusions and future work
This study surveyed mood labels on six popular music websites in mainland China,
revealing a much smaller mood space compared to that in Western websites. In addition, the
coverage of mood labels in Chinese music websites is skewed toward low arousal moods and
away from high arousal moods. An online user listening test was conducted on the mood
labels of two of the surveyed websites. The results show only one third of the given
mood labels were selected by the users even when more labels were allowed.
The comparisons on Chinese and English songs demonstrate that the song language
mattered for editors who generated the mood labels as well as listeners who made use of the
labels. Implications of the findings in music information research were discussed, with focus
areas on interoperability across websites in different countries, integration of categorical
and dimensional models of mood representations and the definition of music mood.

Future work can utilize other methods such as listening behavior analysis which can
exploit users’ music listening records on a larger scale. To reach the goal of global online
music access, more studies are needed for examining music services in other non-Western
countries with significant online music consumers such as India, Mexico and Korea.

Note
1. In the domain of music information retrieval, the term “mood” is often used interchangeably with
the term “emotion” (Yang and Chen, 2012).

References
recommendations”, in Maglaras, L.A., Janicke, H. and Jones, K. (Eds), Industrial Network and

related psycho physiological responses: comparing Canadian listeners to Congolese Pygmies”,


Further reading

Corresponding author
Xiao Hu can be contacted at: xiaoxhu@hku.hk

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Sexual health information-seeking behavior on a social media site: predictors of best answer selection

Yong Jeong Yi
Department of Library and Information Science, Sungkyunkwan University, Seoul, Republic of Korea

Abstract
Purpose – The purpose of this paper is to identify sexual health information needs and the cognitive and affective factors correlated with the best answer chosen by social Q&A users.
Design/methodology/approach – The study collected questions and answers regarding sexual health information on a social Q&A site, and analyzed the questions and a paired sample composed of best and non-best answers (n = 480).
Findings – The main information needs of consumers are human development, sexual behavior, and sexual health. Best answers are more likely to include both cognitive (higher level of readability, risky information, social norms) and affective factors (empathy, positive/negative feelings, and optimistic information) than non-best answers.
Research limitations/implications – The study illuminates the roles of social Q&A as a unique platform to discuss sensitive health topics due to the fact that consumers use such social media sites as critical complementary health information sources.
Practical implications – If health information providers develop information with the factors that the study suggests, not only will it be more adopted by consumers, but it will also ameliorate the quality concerns about online health information.
Originality/value – Previous studies only investigated the most prevalent factors, rather than the most effective ones, which have a greater influence on best answer selection. This study compares the best answers and the non-best answers to overcome the limitations of the previous studies. Above all, the study applied the persuasion concepts to address the cognitive and affective perspectives to the answer evaluations of social Q&A.

Keywords Consumer health information, Social Q&A, Answer quality evaluation, Health information-seeking, Sexual health information

Paper type Research paper

Introduction
The advancement of Web 2.0 technologies has spurred laypeople, i.e., consumers, to seek and share health information through online communities (Bowler et al., 2013; Jin et al., 2016). It is widely known that health information is more sensitive than general information due to its critical effect on human life and health (Yi et al., 2012). In the healthcare field, information is updated and changes quickly, even with the growth of controversial issues. Whenever such issues revolve around unfamiliar healthcare problems or contradictory health information, consumers tend to tap into the diverse health information sources available to them to make sense of uncertain health issues (Genuis, 2012). Such activities are salient in online communities.

Online community users share and construct collaborative knowledge by asking and answering questions. These sites are referred to as community-driven knowledge sites, or social Q&A sites. There are two fundamental reasons why social Q&A sites are thriving. One reason is that they meet the information needs of consumers who prefer interpersonal information sources, that is, consumers tend to acquire health information through
interaction between people, either online or offline (Johnson and Case, 2012). The other is that social networking sites incorporating social Q&A provide consumers with a range of benefits and opportunities to empower themselves through sharing both emotions and knowledge (Lessick, 2015). Questioners can choose one of the answers provided for their questions as a best answer. The rest of the answers, which are not chosen, are tagged as non-best answers (Arai and Handayani, 2013). Consumers’ answer evaluations are influenced by their emotions, as well as their knowledge and thoughts (Yi et al., 2012). Since a best answer is selected by a consumer who is not a healthcare expert, it does not ensure its quality.

In South Korea, Naver Knowledge iN is one of the successful social Q&A sites frequently compared with Yahoo! Answers (Lee and Jang, 2010). Since its launch on 2002, Naver Knowledge iN has accumulated over 44m users, 150m questions, and 210m answers (BizSpring, 2002/2016; Newsis.com, 2015). On this site, consumers actively search and share private, yet sensitive health information allowing health professionals to provide answers for free with the benefit of advertising their workplaces; hence, Naver Knowledge iN is an excellent site for examining sexual health information behavior among the general Korean population and the influence of expertise on selecting the best answers (Bae and Yi, 2017).

Culturally, Koreans do not feel comfortable opening up about their sexual health or sexual life. Social Q&A sites help consumers obtain collective knowledge and insight from a large number of people. Additionally, by assuring anonymity, privacy, and easy access, these types of internet sites provide a unique place to discuss sexual topics, which range from general issues such as reproduction and sexual identity to highly sensitive issues such as sexually transmitted diseases (STD) and sexual abuse (Bowler et al., 2013; Buhi et al., 2010; Jin et al., 2016).

Although sexual health is perceived as a trendy topic, there is inadequate research on what kind of sexual health information needs and information-seeking behaviors consumers have on social Q&A sites (Magee et al., 2012). Previous social Q&A studies focused on investigating the most prevalent factors to see which factors influenced the best answer selection. This was done by analyzing only the best answers; the methodology was not appropriate for identifying the most “effective” factors which have a greater influence on best answer selection (Bowler et al., 2013; Jin et al., 2016). To overcome the limitations of the previous studies, the present study compares best answers and non-best answers. Above all, the study applies the persuasion concepts to address the cognitive and affective perspectives to the answer evaluations of social Q&A, which has not been attempted by previous studies. Thus, the purpose of the present study is to investigate sexual health information needs and the cognitive and affective factors correlated with the best answer. Meticulous attention to sexual health information-seeking behavior on social Q&A can inform health information providers on how to guide consumers to quality health information or inform on how to design health literacy education. To bridge the research gap mentioned above, the present study proposes the research questions as follows (specific H1 to H10, will be presented in the following conceptual framework and hypotheses):

**RQ1.** What are the sexual health informational needs of social Q&A users?

**RQ2.** Do cognitive factors influence the askers’ selection of best answers (H1–H5)?

**RQ3.** Do affective factors influence the askers’ selection of best answers (H6–H10)?

**Conceptual framework and hypotheses**

Using the internet or social media as a source of health information is prevalent (Fox and Duggan, 2013). Social Q&A sites have changed the rules for seeking health information by facilitating community-based social searches. Compared to web search results, social
search provides a networking opportunity to find potential answers that consumers are seeking and tailored to them, whereas traditional searches do not offer these options. Consumers distinguish these sites as opportunistic places to obtain divergent thoughts and perspectives from others, which they take into account when making a health-related decision.

Previous studies discussed diverse topics related to sexual health information needs (Evers et al., 2013; Jones and Biddlecom, 2011; Magee et al., 2012; Priest et al., 2016). Jones and Biddlecom (2011) analyzed 177 sexual health websites and identified the following seven topics: sexual anatomy, sexual terminology, pregnancy, abortion, reproductive cancers, pornography, and sexual pleasure. In addition to these topics, many studies examined people’s information-seeking behaviors about sexual health, such as contraception, sexual health testing, risk prevention (Magee et al., 2012), sexuality, medication and sex, and sexual abuse (Buhi et al., 2010; Spink et al., 2004). Priest et al. (2016) found gender differences for sexual health information needs on the ChaCha, a social Q&A; while female questioners tended to ask pregnancy and menstruation, male questioners tended to ask body image and sexuality. Regarding sexuality, people relied on different types of non-professional sexual information sources including friends, social media, and the internet (Doornwaard et al., 2015; Evers et al., 2013). Kubicek et al. (2010) illuminated the absence of sexual health information for LGBT people and strongly suggested appropriate sexual education to protect them from risky sexual behavior.

Along with limited information for particular sexual health topics, there have been concerns about the quality of online health information (Jin et al., 2016). The explosion of online health information has led to confusion for information seeking and use, although online information provided advantages such as usefulness and convenience. This was due to the fact that non-health professionals could provide this online health information (Buhi et al., 2010). The effect of inaccurate health information can be quite harmful for people; for example, people perceive unprotected sex has low risk (Kubicek et al., 2010). Similarly, many studies have discussed the quality of answers in social Q&A (Harper et al., 2008), and endeavored to identify the criteria/indicators for judging the credibility and quality of health information to reduce the uncertainty surrounding the question of validity (Genuis, 2012).

Regarding the quality evaluation of online health information, a majority of studies have discussed the cognitive quality criteria/indicators such as authority, accuracy, expertise, ease of understanding, and experience (Hong, 2006; Yoo et al., 2014). Along with the proliferation of Web 2.0, however, which features the interaction between health information consumers and providers, consumers identified affective criteria as one of the important virtues to discern quality answers because they do not tend to trust the information coming from sources who do not seem to empathize about how they feel about (Yi et al., 2012).

One of the key characteristics of social Q&A sites is that consumers can rate responses; a questioner can select the best answer among the answers posted. Selecting best answers depends on how effectively the responder persuades a questioner compared to the other answers provided. Persuasion is closely associated with emotions; information that evokes emotion is more effective than information that is non-emotional when influencing people’s decision-making (Yoo et al., 2014). According to Nahl (2007), “all information needs, seeking, reception, and use are processed through emotion” (p. 4), while emotion can be encoded and delivered by information; thus, the impact of emotion in information seeking cannot be ignored. On social Q&A sites, quality evaluations by consumers reveal the importance of social supports. Social supports pertain to the affective and cognitive aspects by integrating psychological and informative, or practical, information. Affective aspects refer to emotional dimensions (e.g. feeling, attitude, and temperament), whereas cognitive
aspects or factors can be defined as the intellectual or perceptual dimensions (e.g., knowledge and thought) involved in understanding or interpreting something (Savolainen, 2015). It is obvious that for people who seek help or answers for serious health issues, socio-emotional support is critical (Jin et al., 2016). Thus, people highly assess the answers that have empathy and encouragement as well as helpful information for their health problems.

The present study classifies the factors that influence answer quality evaluation into two categories: cognitive factors and affective factors as presented in Table I. The present study examines only the factors that are relevant and clearly identified with the content analysis method. As for cognitive factors, expertise and experience have been the most frequently used factors in previous studies. Regarding new concepts derived from the persuasion framework, readability is related to health literacy, and risk and social norms are frequently employed particularly in the context of health. As for affective factors, empathy is a salient factor used in previous studies. Other affective factors, positive/negative feelings and optimistic/pessimistic information are derived from the persuasion framework since they are appropriate for delivering health messages.

### Cognitive factors

Social Q&A sites consist of consumers who normally do not have any healthcare background. From the perspective of consumers, the present study investigates five

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition and example</th>
<th>Related studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>Providing healthcare background, e.g., “Hello, I am a doctor consultant for obstetrics and gynecology”</td>
<td>Gazan (2006), Hong (2006), Wathen and Burkell (2002), Yi et al. (2012)</td>
</tr>
<tr>
<td>Readability</td>
<td>Levels of readability according to Flesch–Kincaid, e.g., Flesch–Kincaid Grade Level: 8.9</td>
<td>Blooma et al. (2012), Chua and Banerjee (2015), Fong et al. (2014), Frew et al. (2009)</td>
</tr>
<tr>
<td>Risk</td>
<td>Providing risk information, the possibility of getting certain diseases, e.g., “Herpes is likely to recur again”</td>
<td>Freimuth and Hovick (2012), Frew et al. (2009), Ledford (2009), Parera and Suris (2004), Wright et al. (2013)</td>
</tr>
<tr>
<td>Experience</td>
<td>Providing one’s experience, e.g., “In my case, after a while, the symptoms disappeared”</td>
<td>Banerjee and Greene (2012), Blooma et al. (2012), Braverman (2008), Jeon and Rieh (2013), Silience et al. (2006), Yi et al. (2012)</td>
</tr>
<tr>
<td>Social norms</td>
<td>Discussing what the majority of people think or do, e.g., “Many people go through the rehabilitation surgery in order to prevent sexual trouble with their spouse”</td>
<td>Bae and Yi (2017), Fong et al. (2014), Lally et al. (2011), Mollen et al. (2013), Robinson et al. (2014), Stok et al. (2012)</td>
</tr>
<tr>
<td>Empathy</td>
<td>Understanding/sharing an asker’s feelings, e.g., “I think you’re having a hard time with homosexual issues. I have had a lot of trouble with them, too”</td>
<td>Bowler et al. (2013), Hong (2006), Nambisan (2011), Osborne (2012), Yi et al. (2012)</td>
</tr>
<tr>
<td>Positive feeling</td>
<td>Expressing positive feelings, e.g., comfortable, content, glad, happy, e.g., “It’s just a happy thing that I have someone I love”</td>
<td>Bowler et al. (2013), Nahl (2007)</td>
</tr>
<tr>
<td>Negative feeling</td>
<td>Expressing negative feelings, e.g., anxious, horrible, frustrated, nervous, e.g., “I was terrified by the same symptoms that you have mentioned”</td>
<td>Lang and Yegiyan (2008), Nabi et al. (2008), Rains and Tuckachinsky (2015), Yoo et al. (2014)</td>
</tr>
<tr>
<td>Optimistic information</td>
<td>Judging current situation as optimistic, e.g., “It is nothing to worry about”</td>
<td>Bae and Yi (2017), Bowler et al. (2013), Genuis (2012)</td>
</tr>
<tr>
<td>Pessimistic information</td>
<td>Judging current situation as pessimistic, e.g., “It’s very likely to be a yeast infection”</td>
<td>Bae and Yi (2017), Bowler et al. (2013)</td>
</tr>
</tbody>
</table>

Table I. Definitions and related studies of variables
cognitive factors that are associated with best answer selection: expertise, readability, risk, experience, and social norms.

**Expertise.** When pursuing health information, consumers normally critique credibility by checking where the information is coming from, or whether the information is provided by healthcare experts (Wathen and Burkell, 2002). The level of expertise is indicated by knowledge, competence for, or credentials in a certain area. According to previous findings, the effect of expertise on selecting best answers has not been consistent (Gazan, 2006; Sillence et al., 2006). Blooma et al. (2012) suggested that expertise is one of the determining factors for selecting best answers for certain topics. The consumers who question topics such as computers, mathematics, or business prefer information to be provided by experts, whereas those who ask health-related questions tend to seek practical advice from non-experts due to skepticism toward typical answers given by health experts. It is notable that people tended to put more value on personal testimony provided by people who empathized with the askers’ health issues rather than the knowledge of health experts (Sillence et al., 2006). Askers wanted to apply such firsthand experience to their own health issues and overcome the limitation of conventional medical treatments (Yi et al., 2012). In sum, while people did rely on health experts’ authority, they also wanted to learn from someone else’s experience, which then could be applied to themselves and could not be provided by health professionals. Considering the particular context of sexual health, it can be expected that expertise is not influential for choosing the best answers; therefore, the following can be hypothesized:

**H1.** There will be a difference of expertise between best answers and non-best answers.

**Readability.** Health literacy is defined as the essential ability to identify health information needs and to search, understand, evaluate, and use health information (Osborne, 2012). Seeking and evaluating health information requires a certain level of literacy because a large body of health information is written at a ninth grade level or above (Nielsen-Bohlman et al., 2004). Frew et al. (2009) found that higher levels of readability were unlikely to be a persuasive feature for recruitment campaigns for a HIV vaccine study because the advanced reading levels hindered people’s ability to process such information. That is, concise and simple language would help people understand the information, which would improve the clarity of it (Blooma et al., 2012; Fong et al., 2014). Thus, readability was a significant factor affecting social Q&A users’ quality evaluation for answers (Chua and Banerjee, 2015). However, there have been different findings. When appraising health information, people are likely to trust a text that has a professional writing style or that sounds scientific (Yi et al., 2012). Health information written with high-level vocabulary possibly leads consumers to conclude that these answers are written by health experts, and the information given is of high quality. Thus, the study hypothesizes:

**H2.** There will be a difference of readability levels between best answers and non-best answers.

**Risk.** Health information highlighting risks can be helpful for healthcare decisions (Straub, 2006). Communication of risk between adolescents and trusted mentors such as parents, teachers, or religious leaders is expected to affect getting accurate sexual knowledge or perceiving sexual risk, such as teen pregnancy and STDs (Wright et al., 2013). Communication of sexual risks, thus, reduces the likelihood of having sexual intercourse (Wright et al., 2013) or, multiple partners (Parera and Suris, 2004), and increases self-efficacy for contraceptive use (Mueller and Powers, 1990). Many studies have investigated risk perceptions in terms of cognitive aspects and the association
between risk perception and health behaviors (Leppin and Aro, 2009). In comparison, some studies discussed that changes in health behaviors are made because of emotional responses, such as anxiety and fear, caused by cognition of risk (Freimuth and Hovick, 2012; Ledford, 2009). According to Frew et al. (2009), when a health behavior was perceived as risky, people were more responsive to the information that emphasized negative outcomes than positive ones. Based on previous studies on risk, the present study hypothesizes that risk is associated with best answer selection:

**H3.** Risk information will be presented more in best answers than non-best answers.

*Experience.* People who frequently visited social Q&A sites usually asked uncommon health questions and expected non-popular/tailored answers to their concerns (Jeon and Rieh, 2013). They favorably assessed answers based on others’ firsthand experience. Learning from another’s experiences may be one of the most beneficial aspects to social Q&A sites, seeing that it promotes practical interaction with people online. Those who frequently used social Q&A sites were engrossed in “collective personal experience” (Gazan, 2006). Particularly in the case where a health issue was peculiar, people were likely to seek health information from others’ experiences because it can complement professional healthcare services (Yi et al., 2012). That is, when physicians cannot determine one’s health issue or how to treat it, consumers who have gone through the same problem can answer the health question. The health information conveying personal experiences are more effective in health promotion that focuses on changing unhealthy behaviors, such as antidrug or antismoking, than the health information carrying normative or statistical messages (Banerjee and Greene, 2012; Braverman, 2008). Thus, we can hypothesize the following:

**H4.** Personal experience will be presented more in best answers than non-best answers.

*Social norms.* People value what others think and do in a given context, as long as it is a socially acceptable behavior. The standards that judges acceptance are social norms (Lapinski and Rimal, 2005). People shape their thoughts and behaviors based on social norms. Consumers visit social Q&A sites with a curiosity of others’ thoughts as well as motivation to solve their health problems by obtaining others’ advice or verifying their own knowledge (Fong et al., 2014). Such collective personal perspectives influence consumers’ decision-making. Responses on social Q&A sites provide diverse views, values, and judgment or coping styles that a majority of people take into consideration, which contributes to consumers’ judgment on appropriateness (Mollen et al., 2013). According to previous studies, adolescents’ sexual behavior and sources of sexual information were associated with social norms. Perceived peer approval of having sex was positively correlated with their exposure to sex or sexually explicit online resources (Doornwaard et al., 2015). Health information based on social norms had significantly encouraged healthy behavior, rather than health information based solely on health without considering the social norms (Lally et al., 2011; Robinson et al., 2014; Stok et al., 2012). Creating new social norms has significantly changed injection behavior by drug users (e.g. syringe or needle-sharing) (Latkin et al., 2013). Thus, social norms should be related to choosing best answers:

**H5.** Social norms will be presented more in best answers than non-best answers.

*Affective factors.* This study examines five affective factors that influence selecting best answers: empathy, positive feelings, negative feelings, optimistic information, and pessimistic information.
Empathy. Empathy is likely to be perceived in information provided by laypeople, whereas expertise is likely to be perceived in information delivered from professional sources (Gray et al., 2005). Empathy is defined as “sensitivity to another person’s feelings, thoughts, and experiences” (p. 231) (Osborne, 2012). Previous studies have examined the notable role of empathy in online support groups where people share similar health issues and coping strategies. People with different lifestyles from all over the world have something in common often seek and offer emotional support in online forums. People voluntarily respond to one another with helpful information that, in return, gives people a sense that they are not alone and that they have people that will listen and understand them (Caplan, 1974). Emotional support should not replace health experts’ care; however, consumers are able to form a sense of collaboration and empowerment through these emotional interactions (Nambisan, 2011). When people believe that they are being cared for, they feel relieved and less anxious about the problematic health issues; empathetic understanding ameliorates their embarrassment and feelings of stigmatization (Cobb, 1976). Furthermore, they are more likely to be “receptive to learn new health information” (p. 231) (Osborne, 2012). In Yahoo! Answers, people who have suffered from eating disorders preferred answers that convey emotional support, rather than those that come from evidence-based medical sources (Bowler et al., 2013). Thus, the present study expects that empathy influences choosing a best answer in a social Q&A site, and hypothesizes:

H6. Empathy will be presented more in best answers than non-best answers.

Positive/negative feelings. Affective responses express positive (e.g. happiness, comfort, contentment) or negative feelings (e.g. anxiety, frustration, panic, and sadness) (Banerjee and Greene, 2012). Previous studies have addressed the impact of emotion on delivering health information; many of them indicated that emotional information is more effective than non-emotional information (Lang and Yegiyan, 2008). People’s decision-making tends to be more influenced by emotion in unfamiliar, risky, or stressful situations than in general situations (Grimm et al., 2014). Consumers who are suffering from health problems use positive emotions to fortify their coping mechanisms to enhance self-efficacy and optimism; positive feelings are beneficial for them in reducing anxiety and improving psychological and health outcomes (Bowler et al., 2013; Nahl, 2007). Due to the fact that affection influences cognitive behavior and vice versa, the role of positive feelings is very important. Discussing sensitive health information is usually distressful, causing people to more likely be in an emotion-laden state (Yi et al., 2012). On social Q&A, people present their informational needs using emotional language (Savolainen, 2015). People with eating disorders typically ask socio-emotional questions using negative language, such as depressed, hurt, hate, bad, etc. The action of asking and answering questions is in itself another concept. Although askers reveal negative feelings, it is difficult to predict whether they prefer answers that express negative feelings in response to such questions. Previous results showed that negative affectivity can aggravate people’s information seeking and hinder their coping skills (Nahl, 2007). Therefore, this did not satisfy their need for information to reduce negative emotions, which were evoked by threatening or dangerous health issues (Rains and Tukachinsky, 2015). Lang and Yegiyan (2008), however, examined that the effect of negative emotional information was greater than positive emotional messages. Some studies investigated the influence of negative emotions on health behaviors (Nabi et al., 2008; Yoo et al., 2014). Yoo et al. (2014) explored the effects of three different types of negative emotions (sadness, fear, and anger) as motivating factors to take preventive actions, such as getting a cervical cancer vaccine or regular mammograms. The study implied that sadness had a stronger impact
than fear or anger on delivering cancer risk messages. Guided by previous studies, the present study hypothesizes:

H7. Positive feelings will be presented more in best answers than non-best answers.

H8. Negative feelings will be presented more in best answers than non-best answers.

Optimistic/pessimistic information. Positive or negative feelings focus on emotional attributes, which differs from optimistic/pessimistic information. In comparison, psychological comfort is correlated with whether the health information has an optimistic or pessimistic message. As described in Table I, optimistic or pessimistic information is likely to encourage or discourage people, respectively, when they are emotionally laden. Since social Q&A consumers value others’ knowledge and judgment in their virtual community, responders’ comments produce more of an impact for the askers when they are making a decision regarding their health issues. In uncertain or even in panicked situations, askers look forward to responses that suggest that they are not in trouble or that their situations are normal (Genuis, 2012). Positive messages given through optimistic information have a great impact on lowering distress. In Yahoo! Answers, thus, answers that are hopeful are often selected as the best answers, when questioners address their problems with a negative conception (e.g. sounds depressed or dark) (Bowler et al., 2013). In contrast, askers tend to poorly evaluate answers that include critical facts or pessimistic information, which aggravates the askers’ emotions and leaves a negative impact on them (Bowler et al., 2013). If people do not perceive emotional or social support to relieve their anxiety in the responses, even though they provide useful information, people do not recognize the replies as helpful because the askers believe that the repliers do not understand what they are going through. Thus, selecting the best answer depends on whether the response is optimistic or pessimistic:

H9. Optimistic information will be presented more in best answers than non-best answers.

H10. Pessimistic information will be presented less in best answers than non-best answers.

Findings from the previous studies have led to the conceptual framework that guides both cognitive and affective factors considered when selecting best answers on a social Q&A site as presented in Figure 1.

Methods
Analyzing only best answers is inadequate for identifying the predictors of best answers, since a factor that is frequently observed in best answers is not necessarily critically important when that same factor is also as frequently presented in non-best answers. On the other hand, although a factor is not frequently present in best answers, it can be a salient indicator of best answers when it is present much less in non-best answers. Thus, the present study directly compared best answers with non-best answers to effectively find out
the characteristics of best answers. The study collected questions and corresponding answers regarding sexual health information on Naver Knowledge iN, which was harnessed for a year from September 2012 to August 2013.

**Sampling**

To compare the characteristics between the best answers selected by the original questioners and non-best answers that were not selected by the questioners, the study was restricted to questions that have at least two answers: one chosen as a best answer and other answer(s) that were not chosen. The study excluded questions or responses that were not adequate (e.g. slander, joke) or did not have best answers since the original questioners had not selected any yet.

After sifting through unqualified questions, the study found that the maximum number of questions for a couple of months were 40. To collect an equal number of questions for each month, the study selected the first 40 questions asked every month and the corresponding best and non-best answers to be used as a paired sample. To select one non-best answer to compare with the best answer, the study selected a random number (= 3), which was generated by MS Excel function (randbetween). Thus, the third non-best answer from the top was selected as a counterpart non-best answer. When there were only one or two non-best answers, the first one was chosen because it revolves to the first answer after the cycle is completed. As a result, the study collected a sample that was composed of 480 questions and answers. Figure 2 presents the sampling process, which is modified from the previous study (Bae and Yi, 2017).

**Measures and inter-coder reliability**

The study employed content analysis to analyze the data; two research assistants coded the presence or absence of the variables to be analyzed after multiple discussions about their definitions. Readability levels were measured by the Flesch–Kincaid formula offered by Microsoft Word 2007 (Spink *et al.*, 2004). The study used Scott’s $\pi$ to test the inter-coder reliability of two category variables (present or absent) since there were two coders (Krippendorff, 2004), and applied Krippendorff’s $\alpha$ only for the ratio variable, readability levels. Table II presents the test results and definitions of the variables.

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer 1: non-best answer (Not selected as sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Answer 2: best answer (Chosen by a questioner, and selected as sample)</td>
</tr>
<tr>
<td></td>
<td>Answer 3: non-best answer (Not analyzed as sample)</td>
</tr>
<tr>
<td></td>
<td>Answer 4: non-best answer (Third non-best answer, and selected as sample)</td>
</tr>
<tr>
<td></td>
<td>Answer 5: non-best answer (Not analyzed as sample)</td>
</tr>
</tbody>
</table>

**Figure 2.** Sampling process of questions, best answers, and non-best answers

40 Questions from each month ($n=480$)

Random number generated to select a non-best answer: 3
- The third non-best answer from the top is selected

Analyzed as a paired sample: Wilcoxon signed-rank test
The paired sample $t$-test
To directly compare two matched samples, best answers and non-best answers, the study conducted the Wilcoxon signed-rank test and the paired sample $t$-test which is only for readability levels.

**Results**

Regarding $RQ1$, the study found that Naver Knowledge iN users have diverse sexual health information needs. The study categorized sexual health information needs based on the Guidelines for Comprehensive Sexuality Education (Sexuality Information and Education Council of the United States, 2004) in the USA. The Guidelines consist of six key concepts: human development, relationships, personal skills, sexual behavior, sexual health, and society and culture. Under these six key concepts are several topics as presented in Table III.

The present study classified the sexual health information needs according to the main topic of the question and the corresponding answers. For instance, if a question talked about two different topics (e.g., reproductive anatomy and physiology and puberty) but the selected answers were mainly about puberty, the question was classified as puberty. For instance, a 14-year-old student asked that he was concerned about the shape and the size of his reproductive organ, which are the characteristics of puberty, and the selected answer provided that the questioner’s concerns were common phenomenon of puberty. In this case, the question was classified as puberty. Out of 480 questions, 41.6 percent (200 questions) were about human development, 35 percent (168 questions) were related to sexual behavior, 13.9 percent (67 questions) to sexual health, 4.3 percent (21 questions) to relationships, 3.7 percent (18 questions) to personal skills, and 1.2 percent (6 questions) to society and culture. In the category of human development, there were high rates of questions about sexual identity/orientation (83 questions) and reproductive anatomy/physiology, specifically, premature growth (62 questions). The category of sexual behavior had high rates of questions about sexuality throughout life (90 questions) and masturbation (42 questions).

With respect to $RQ2$, Table IV presents the frequency of cognitive and affective factors for best answers and non-best answers. Best answers included high rates of optimistic

<table>
<thead>
<tr>
<th>Variable</th>
<th>Scott’s $\pi$</th>
<th>Krippendorff’s $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>Best</td>
<td>0.917</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.914</td>
</tr>
<tr>
<td>Readability</td>
<td>Best</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>1</td>
</tr>
<tr>
<td>Risk</td>
<td>Best</td>
<td>0.842</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.827</td>
</tr>
<tr>
<td>Experience</td>
<td>Best</td>
<td>0.803</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.837</td>
</tr>
<tr>
<td>Social norms</td>
<td>Best</td>
<td>0.928</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.872</td>
</tr>
<tr>
<td>Empathy</td>
<td>Best</td>
<td>0.770</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.964</td>
</tr>
<tr>
<td>Positive feeling</td>
<td>Best</td>
<td>0.780</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.796</td>
</tr>
<tr>
<td>Negative feeling</td>
<td>Best</td>
<td>0.770</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.798</td>
</tr>
<tr>
<td>Optimistic</td>
<td>Best</td>
<td>0.783</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.974</td>
</tr>
<tr>
<td>Pessimistic</td>
<td>Best</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>0.962</td>
</tr>
</tbody>
</table>

Table II. Inter-coder reliability of variables

Sexual health information-seeking behavior

889
<table>
<thead>
<tr>
<th>Variables</th>
<th>Answers</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>Best</td>
<td>55</td>
<td>11.5</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>68</td>
<td>14.2</td>
</tr>
<tr>
<td>Risk</td>
<td>Best</td>
<td>95</td>
<td>19.8</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>55</td>
<td>11.5</td>
</tr>
<tr>
<td>Experience</td>
<td>Best</td>
<td>56</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>60</td>
<td>12.5</td>
</tr>
<tr>
<td>Social norms</td>
<td>Best</td>
<td>175</td>
<td>36.5</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>73</td>
<td>15.2</td>
</tr>
<tr>
<td>Empathy</td>
<td>Best</td>
<td>206</td>
<td>42.9</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>144</td>
<td>30</td>
</tr>
<tr>
<td>Positive feeling</td>
<td>Best</td>
<td>182</td>
<td>37.9</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>103</td>
<td>21.5</td>
</tr>
<tr>
<td>Negative feeling</td>
<td>Best</td>
<td>195</td>
<td>40.6</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>122</td>
<td>25.4</td>
</tr>
<tr>
<td>Optimistic</td>
<td>Best</td>
<td>319</td>
<td>66.5</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>229</td>
<td>47.7</td>
</tr>
<tr>
<td>Pessimistic</td>
<td>Best</td>
<td>58</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>104</td>
<td>21.7</td>
</tr>
</tbody>
</table>

**Table III.**

Sexual health information needs

<table>
<thead>
<tr>
<th>Variables</th>
<th>Answers</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>Best</td>
<td>55</td>
<td>11.5</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>68</td>
<td>14.2</td>
</tr>
<tr>
<td>Risk</td>
<td>Best</td>
<td>95</td>
<td>19.8</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>55</td>
<td>11.5</td>
</tr>
<tr>
<td>Experience</td>
<td>Best</td>
<td>56</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>60</td>
<td>12.5</td>
</tr>
<tr>
<td>Social norms</td>
<td>Best</td>
<td>175</td>
<td>36.5</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>73</td>
<td>15.2</td>
</tr>
<tr>
<td>Empathy</td>
<td>Best</td>
<td>206</td>
<td>42.9</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>144</td>
<td>30</td>
</tr>
<tr>
<td>Positive feeling</td>
<td>Best</td>
<td>182</td>
<td>37.9</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>103</td>
<td>21.5</td>
</tr>
<tr>
<td>Negative feeling</td>
<td>Best</td>
<td>195</td>
<td>40.6</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>122</td>
<td>25.4</td>
</tr>
<tr>
<td>Optimistic</td>
<td>Best</td>
<td>319</td>
<td>66.5</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>229</td>
<td>47.7</td>
</tr>
<tr>
<td>Pessimistic</td>
<td>Best</td>
<td>58</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>Non-best</td>
<td>104</td>
<td>21.7</td>
</tr>
</tbody>
</table>

**Table IV.**

Frequency of categorical variables
information (66.5 percent), and empathy (42.9 percent); next, in order of frequency, best answers included negative feelings (40.6 percent), positive feelings (37.9 percent), risk (19.8 percent), pessimistic information (12.1 percent), experience (11.7 percent), and expertise (11.5 percent). Levels of readability for both best and non-best answers have 0 as minimum and 12 as maximum. The mean of readability score for best answers is 4.88 (SD = 3.17), and that of non-best answers was 4.24 (SD = 3.60).

Best and non-best answers have significant differences in using cognitive and affective factors. For cognitive factors, the paired sample t-test for readability indicates that best answers are more likely to have higher readability scores ($t (479) = 3.07$, $p = 0.002$) than non-best answers. Best answers are more likely to include risk information and social norms as presented in Table V, thus, $H3$ and $H5$ were supported. However, there is no significant difference in using two factors: expertise and experience: $H1$ and $H4$ were rejected. Four affective factors (empathy, positive feeling, negative feeling, and optimistic information) are present more frequently in best answers. In comparison, best answers are less likely to include pessimistic information, which is present more frequently in non-best answers. Therefore, all five hypotheses of affective factors were confirmed ($H6$-$H10$).

According to the findings of the study, Figure 3 presents a modified conceptual framework for predicting best answer selection.

**Discussion**

Guided by health information-seeking behavior on social media sites and the conceptual framework aforementioned, the present study analyzes sexual health information needs based on consumers’ questions and a set of cognitive and affective factors that influence the askers’ selection of best answers on a social Q&A site. A majority of sexual health

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum of ranks (Non-best – Best)</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>1,880.00</td>
<td>-1.34 ($p = 0.178$)</td>
</tr>
<tr>
<td>Risk</td>
<td>3,415.50</td>
<td>-4.04***</td>
</tr>
<tr>
<td>Experience</td>
<td>2,137.50</td>
<td>-0.41 ($p = 0.680$)</td>
</tr>
<tr>
<td>Social norms</td>
<td>11,628.00</td>
<td>-7.82***</td>
</tr>
<tr>
<td>Empathy</td>
<td>10,120.50</td>
<td>-4.72***</td>
</tr>
<tr>
<td>Positive feeling</td>
<td>10,332.00</td>
<td>-6.11***</td>
</tr>
<tr>
<td>Negative feeling</td>
<td>9,877.00</td>
<td>-5.68***</td>
</tr>
<tr>
<td>Optimistic</td>
<td>11,770.50</td>
<td>-6.78***</td>
</tr>
<tr>
<td>Pessimistic</td>
<td>1,689.50</td>
<td>-4.42***</td>
</tr>
</tbody>
</table>

**Note:** $***p < 0.001$

Table V. Wilcoxon signed-rank test

![Figure 3.](image)
information needs are about human development, sexual behavior, and sexual health. The specific topics with these kinds of highly sensitive issues tend to form higher rates in these categories. Such examples are sexual identity/orientation, premature growth, sexuality, and masturbation. The findings suggest that anonymity is one of the biggest advantages of social Q&A. In other words, consumers are in need of sensitive health information, but are reluctant to ask these private questions face to face.

The three cognitive factors (readability, risk, and social norms) were identified as effective for best answer choices, whereas as the two factors (expertise and experience) were not. It is notable that health information seeking on social Q&A sites contrasts with traditional websites. People consider expertise as one of the most important criteria or virtue for judging health information quality in searching traditional websites, whereas they do not on social Q&A sites. The present study confirms a significant feature of social Q&A sites, where askers of health-related topics gravitate toward firsthand practical advice from non-experts due to skepticism toward typical answers given by healthcare experts (Blooma et al., 2012). Such stereotypical answers cannot satisfy the needs of consumers who pursue not only healthcare knowledge but also socio-emotional support, which suggests that healthcare practitioners pay attention to patients’ emotions to provide effective healthcare services.

Unexpectedly, experience is not a significant factor in predicting best answer selection. Previous studies have discussed that people turn from traditional health websites to social Q&A sites to search for information obtained from personal experience (Jeon and Rieh, 2013). The findings of the present study, however, indicate that although experience is one of the most highly considered characteristics that differentiate information seeking on social Q&A sites from that on traditional websites, it is not significant for distinguishing the best answers from non-best answers. The finding suggests that consumers who ask highly sensitive information such as sexual health information are more likely to seek others’ collective thoughts, rather than real experience, before making a decision for their serious issues for which they are not sure of what to do.

According to the findings of sexual health information needs, there are many consumers who ask others’ opinions about their sexual identity or relationship. These issues are usually very personal, and the individual situations are quite different, which means that a questioner is not likely to apply what an answerer has gone through into his or her issues. The present study indicates that social norms are one of the key cognitive factors for answer evaluation. Significance of social norms illustrates that consumers are likely to determine what they should do based on what others approve or believe acceptable. Questioners who post questions relating to sexual health problems are more likely to rate the answers that are consistent with social norms as best answers than are those that do not reference social norms. This verifies the findings of previous studies, which suggest that consumers prefer social Q&A sites to traditional websites because on social Q&A sites, they can find how others think or act regarding their own issues (Jin et al., 2016).

Five affective factors (empathy, positive feeling, negative feeling, optimistic, and pessimistic information) were found to be significant in deciding between best and non-best answers. The findings of those affective factors confirm the critical roles of emotion in delivering information. While the expression of positive feelings is employed to comfort questioners, the expression of negative feelings is mainly used to empathize with questioners or to alleviate their anxiety. In the same vein, humor helps people effectively communicate about serious sexual health issues by reducing the concerns about the stigma attached to such issues (Evers et al., 2013).

There are only a few studies that have explored the effect of optimistic or pessimistic information (Bowler et al., 2013; Yi et al., 2012). These studies indicated that people preferred hopeful answers, whereas they perceived critical or pessimistic answers as unhelpful.
Interestingly, pessimistic information was identified as a predictor for non-best answer selection, that is, the presence of pessimistic information was significantly higher in other answers than the best answers. Such findings illustrate the health information needs of social Q&A users. They do not want to obtain responses that convey negative messages about their issues; rather, they eagerly look for encouragement or social support that ameliorates their anxiety or fear (Genuis, 2012). The social Q&A users' preference for optimistic information over critical information provides an insight on how to deliver sensitive health information. While previous studies have had many discussions about empathy (Yi et al., 2012) — health information consumers highly evaluated information which demonstrated shared understanding for their health problem — only a few studies have researched the use of positive/negative feelings and optimistic/pessimistic information in providing effective information services. Hence, these findings of the present study are fruitful and can lead to improvement in the health information area.

One of the limitations of previous social Q&A studies is that they focus on analyzing only best answers. The previous methodology is inadequate to identify the effective characteristics that distinguish best answers from non-best answers. Thus, this study compared the best answers and the non-best answers to overcome the limitations of the previous studies. Above all, the present study applies the persuasion concepts to address the cognitive and affective perspectives to the answer evaluations of social Q&A, which is not often attempted by previous studies. The findings of the present study can be generalized in cultures that regard sexual health topics as private or sensitive and feel uncomfortable when sharing such issues because people in the same culture tend to share health information-seeking behaviors developed in a certain context (Yi et al., 2012).

Conclusion
The present study examined sexual health information needs and information-seeking behavior in Naver Knowledge iN, which is mainly used by Koreans who are likely to have conservative perspectives on sexual health issues and infrequently talk about these issues with people. The main sexual health information needs of social Q&A users were human development, sexual behavior, and sexual health. The study developed the conceptual framework to illuminate cognitive and affective factors that were associated with best answer selection. With respect to cognitive factors, the research indicated that best answers had a higher level of readability and presented more risk information and social norms than did non-best answers. Four affective factors (empathy, positive/negative feeling, and optimistic information) were present more frequently in best answers than non-best answers. Interestingly, pessimistic information was revealed as a predictor for non-best answers.

The study illuminated the roles of social Q&A as a unique platform to discuss sensitive health topics due to the fact that consumers use such social media sites as critical complementary health information sources. The primary contribution to this field of research is that the present study applied the persuasion concepts to identify predictors of best answer selection. This extends the horizon to discuss quality evaluation by employing and validating new concepts, such as readability, risk, social norms, positive/negative feelings, and optimistic/pessimistic information.

As information and reference services utilize more social media, social interaction behavior occurring in social Q&A will provide practical insights into librarianship on delivering sensitive and private health information to users. When information providers, including librarians, guide quality health information to users, they should consider that users will be satisfied with the health information with the cognitive and affective factors that are identified in the study. Furthermore, if health information providers, including healthcare practitioners, develop health information that the present study suggests,
not only will it be more adopted by consumers, but it will also ameliorate the quality concerns about online health information. This will furthermore lead to effective and thorough health communication.

Future research needs to corroborate a health answer evaluation model that is composed of the cognitive and affective factors that the present study identifies by replicating it in different domains, e.g., Quora, which is one of the most frequently used social Q&A sites (Ovadia, 2011). The answer evaluation model needs to be applied to diverse healthcare topics, such as mental health or chronic diseases, to better understand health information-seeking behavior. Considering the increasing frequency with which consumers seek health information through mobile devices, future research should explore how current mobile healthcare applications support consumers' cognitive and affective needs, and suggest some insights to develop more sophisticated applications embedded with functions to support consumers' cognitive and affective needs.

References


Further reading


About the author

Yong Jeong Yi, MLIS, PhD, is Assistant Professor at the Department of Library and Information Science, Sungkyunkwan University, Seoul, South Korea. She has completed PhD Degree in the School of Information, College of Communication and Information at Florida State University. Yong Jeong Yi can be contacted at: redpapa01@skku.edu

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com
Predicting the success of Twitter in healthcare

A synthesis of perceived quality, usefulness and flow experience by healthcare professionals

Mohamed Gamal Aboelmaged
Department of Business Administration, University of Sharjah, Sharjah, UAE

Abstract
Purpose – The purpose of this paper is to predict Twitter satisfaction by healthcare professionals through integrating constructs of Csikszentmihalyi’s flow theory, quality dimensions and usefulness.
Design/methodology/approach – Survey responses of 108 physicians from a variety of specialisations in the United Arab Emirates have been validated and analysed by means of partial least squares-based structural equation modelling method using smartPLS software.
Findings – Service quality has emerged as the most influential quality dimension that positively impact flow state and perceived usefulness of Twitter, while information quality, surprisingly, does not show any effect. The findings also indicate that flow state plays a significant role in shaping physicians’ satisfaction with Twitter. The study also enhances our understanding concerning the effects of perceived usefulness on flow state and satisfaction.
Research limitations/implications – Understanding factors that influence Twitter satisfaction can help healthcare managers construct appropriate intervention strategies for maximising professional benefits of social media and minimising user resistance. This is important because top managers usually ratify traditional practices that are only of limited effect. Also, the findings help vendors to accentuate user’s concerns in addition to system functionalities in social media applications.
Originality/value – The paper is an early attempt to propose a model for social media success in a professional context in general and healthcare in particular. It also one of first studies that examine social media satisfaction through integrating contemporary information system success and acceptance models with flow theory.

Keywords Information quality, Service quality, Twitter, Satisfaction, System quality, Flow theory

1. Introduction
Social media usage has become an inevitable phenomenon over the past decade. Whereas the term “social media” is elusive and evolving, it represents internet-based networking technologies that allow individuals and communities to create, share, collaborate or exchange information, ideas, interests, images and other content in a social dialogue. Social media are based on interactive Web 2.0 applications that facilitate user-generated content in a virtual community as well as usability and the ability of the applications to operate with other systems using web-based or mobile platforms such as smartphones and tablets (Obar and Wildman, 2015). Social media applications underpin social networking (e.g. Twitter, Facebook, Myspace, Google Plus), professional networks (e.g. LinkedIn), media presentations and networks (e.g. YouTube and Instagram), and platforms that produce or aggregate content (e.g. blogs, Wikipedia). Social media has reshaped the way people share thoughts, choices, behaviours and knowledge by making them visible, interlinked, accessible, reviewable and searchable, whereas they were previously silent or hidden.

In the context of healthcare, the professional use of social media by physicians, nurses, pharmacists, counsellors and caregivers can include the promotion of healthy behaviours, patient care, student education, medical development and training, communication with colleagues, the creation of communities and the endorsement of public wellness. Physicians
also join social online groups where they can share news, cases and expert views, and discuss and disseminate medical research, consult colleagues and communicate with patients to augment clinical care (Ventola, 2014). In a descriptive study on the managerial implications of social media in healthcare, Aboelmaged et al. (2016) confirmed that social media platforms such as Twitter, Facebook and YouTube provide a new forum for marketing, managing human resources, training and development, and customer service within the healthcare context. Ba and Wang (2013) indicated that health-related online networks are starting to provide mechanisms to enhance people’s daily lives by enabling them to monitor their diet and providing motivation for them to change their lifestyles. Similarly, Panahi et al. (2014) identified that social media helps physicians stay connected with their patients, share professional knowledge, consider methods for teaching and engagement in continuing medical education. On the patient side, research shows that social media allows patients to enhance their physical and psychological competencies by using social media as a forum for community dialogue to obtain essential knowledge and support (Antheunis et al., 2013).

However, the confirmation of effective social media interventions in healthcare would require further studies to explore the issue of acceptability and satisfaction (Welch et al., 2016). Prior research has focussed considerable attention on users’ satisfaction as a central concept in the online context due to its role in shaping user’s loyalty and system success (Ding and Lii, 2016). Nevertheless, less attention has been placed on the factors that influence users’ satisfaction with social media in general and within a professional context such as healthcare in particular. This may be due to an excessive focus on the initial adoption of social networks (Gao and Bai, 2014) rather than considering post-adoption issues, such as user satisfaction. An exception is Krishen et al.’s (2016) study on social media satisfaction which highlighted the role of meeting user’s virtual needs such as innovativeness, belonging, affinity and interactivity in online satisfaction. Besides, the prior literature on social media usage in healthcare is primarily technical with a medical orientation. An understanding of factors that may influence physicians’ and nurses’ satisfaction with the professional use of social media applications would be of significant value, not only for managerial benefits but also for technology vendors since the influence of technology investment on hospital efficiency and quality are of great interest to healthcare executives as well as insurers (Gholami et al., 2015). Responding to this gap and being motivated by the need to further understand social media success in healthcare contexts, this paper synthesises the information system (IS) success model (Delone and McLean, 2004), flow theory (Csikszentmihalyi, 1997) and the technology acceptance model (TAM) (Davis, 1993) to hypothesise a new model to explain the factors that influence physicians’ satisfaction with Twitter, as an example of a social media platform, and denoting the importance of Twitter in this context (see section 2). Accordingly, quality factors including system quality, service quality and information quality are examined concerning their impact on the perceived usefulness and flow state towards user’s satisfaction. Examining the level of satisfaction with Twitter is critical for social media success since vendors have invested great effort and resources on releasing and updating the application, and intense competition among applications allows users to switch at a very low cost. From a practical standpoint, understanding factors that may influence users’ satisfaction with Twitter can help administrators construct appropriate intervention strategies for maximising application benefits and minimising user resistance in a professional context such as healthcare. This is important because top management usually ratifies traditional practices, consisting of software demonstrations, user training and help desk staffing, that are only of limited effect. Further, most applications tend to accentuate system functionalities rather than user considerations, such as perceived quality dimensions, the usefulness of the application, and user’s immersion and engagement behaviour.
The next section reviews prior literature on social media practices relevant to the healthcare context, IS success, flow theory, and usefulness constructs. In Section 3, a theoretical model of satisfaction with Twitter is constructed. Section 4 depicts the research methods and instrument construction. Statistical data analysis and results are presented in Section 5. This is followed by a discussion of the findings and managerial implications in Section 6. Finally, research limitations and future directions are delineated.

2. Literature review

2.1 Twitter in the healthcare context

Previous research has confirmed that Twitter is an effective platform that is widely used by healthcare professionals to communicate with patients, colleagues and healthcare providers (e.g. Diddi and Lundy, 2017; Gomes and Coustasse, 2015). It has been used by the clinical community as a source of up-to-date medical information and news (Panahi et al., 2014). Even local governments use Twitter, rather than Facebook, for disseminating public health information (Neiger et al., 2013). In addition, Twitter can clearly expose real-time information that is difficult to access through other platforms concerning personal health (Cole-Lewis et al., 2015), medical education (Choo et al., 2015), health research (Sinnenberg, et al., 2017), and professionals' views concerning health issues (Goff, 2015). Lee et al. (2014) conducted a content analysis of health-related tweets. They showed that most tweets tended to be about testable claims and personal health experiences communicated among patients, physicians and nurses. Fuoco and Leveridge (2015) examined urologists' attitudes towards social media platforms, including Twitter. The analysis depicted that inter-professional communication (67 per cent) and interaction with patients (73 per cent) are the most important applications. Furthermore, an analysis of 2,400 tweets by Canadian healthcare professionals revealed that health services, personal health practices and education are the predominant tweeted themes which reflect various political and social healthcare issues (Donelle and Booth, 2012). Examining a sample of 1,000 tweets, Neiger et al. (2013) found that US health departments use Twitter to disseminate personal and organisational information related to health topics as a tool for facilitating public health transition. Huang and Dunbar (2013) examined 23,300 posts and tweets on US hospitals' web pages and indicated that both the Facebook and Twitter platforms can be used as effective two-way marketing tools with the community. Tyrawski and DeAndrea (2015) analysed the content of Twitter and other social media accounts of the top 15 pharmaceutical companies in the USA. The study revealed that advertising that is consistent with the US Food and Drug Administration regulations and helps consumers seeking behaviour was found in 40.7 per cent of posts. In addition, a systematic review of 876 research papers concerning Twitter use in health education demonstrated its vital role as part of a blended approach to medical education that improves professional practices and accessibility to real-world expertise (Smith and Lambert, 2014).

Twitter has also been considered as a vehicle for participatory medicine where physicians, nurses and other stakeholders build on patient-centred care to digitally connect patients to perform all aspects of care by themselves (e.g. retrieving records, booking appointments, interacting with relevant physicians and support groups, reimbursing bills, etc.; Lim, 2016; Millenson, 2011). Gallant et al. (2011) performed a content analysis of more than 1,000 web pages of top US hospitals. They revealed that Twitter is among top social media platforms that support the involvement of patients in healthcare service, which may transform traditional hospital-patient interactions. In addition, Tsuya et al. (2014) depicted that cancer patients in Japan use Twitter to share information with healthcare providers to enhance knowledge and reduce anxiety levels concerning diagnosis, symptoms and cancer treatment. Tracking activities and progress of patient communities, and their interactions with caregivers can also improve service quality. For example, the exposure of patients' presence and the expressiveness of tweets and retweets empower patients' physical and
virtual involvement in medical panels and meetings to highlight vital service issues in healthcare (Dredze, 2012; Harmel and Young, 2013). Twitter has become a powerful tool in the healthcare context due to many features that motivate healthcare professionals to use it over other social media platforms including Facebook. Among these features, limited characters and content that urge Twitter users in getting to the point, a controlled timeline, searching of conversations, shared keywords and news very quickly connect users with people that they need to know rather than people they already know (Meeney et al., 2015). Examples of the use of Twitter in healthcare include: emergency response team management, supportive care for patients and family members, diabetes management, adverse event reporting in the clinical setting, coordinating preoperative, perioperative and postoperative care, census management/monitoring, arranging outpatient care, crowdsourcing for healthcare resources, shift-bidding for nurses and other healthcare professionals, recruitment of healthcare staff, clinical education coordination, facilitating patient transfer processes, coordinating patient discharges with all services, post-discharge patient consultations and follow-up care, tissue and blood recruitment, and nursing mentoring and collaboration are among them (Baumann, 2009). Moreover, examples of widely used hashtags in healthcare include #Patient, #PatientExperience, #HealthTalk, #medicalbilling, #patientengagement, #nurses, #digitalhealth, #HealthReform, #Healthcosts, #occupyhealthcare, #healthcareforall, #physician, #MedEd (Referralmd blog, 2017). Recognising the importance of the Twitter platform in healthcare, most of the prior studies have investigated the benefits of Twitter using conceptual or descriptive approaches. Rigorous studies that investigate the success of Twitter in the healthcare context using a specific theory or framework barely exist. This is consistent with Panahi et al.’s (2014) claim that a very limited number of studies have explored the experiences of physicians in adopting social media, through the fast growing social media applications in healthcare. Therefore, one of the contributions of this study is to fill this gap through synthesising constructs from well-established theories to examine healthcare professionals’ satisfaction with Twitter. The subsequent sections will shed more light on related IS theories and models that are concerned with users' satisfaction.

2.2 Flow experience
As a result of his research on the stimuli that guide people’s engagement with an activity, Csikszentmihalyi (1997) found that one of the stimuli is a state of mind that he named “flow experience”, which reflects deep immersion in an activity while nothing else (e.g. extrinsic rewards, money or reputation) seems to matter. When people are entirely absorbed by an activity, they lose self-consciousness and feel enjoyment, and their experience flows with little distinction between past, present and future; between self and environment; or between stimulus and response (Csikszentmihalyi, 1997; Jackson and Marsh, 1996). The flow concept can be measured using three aspects: embracing concentration, perceived control and perceived enjoyment (Cheng, 2014; Zhou, 2014). Perceived enjoyment captures users’ intrinsic pleasure, apart from any performance consequences that may be anticipated when using a technology. Concentration reflects an immersion of users’ attention on an activity, while perceived control captures a belief that an individual is capable of influencing and making a difference in his/her behaviour and the surrounding environment. The flow state is not limited to specific types of actions rather it is a general phenomenon associated with carrying out any activity, including climbing, dancing, sailing, playing chess, gaming and communicating. Research also suggests that flow state can be facilitated in the context of online communication (Skadberga and Kimmel, 2004). The concept turned out to be a widely used variable in technology acceptance research. Chang (2013) revealed that flow state plays a mediation role and produces indirect effects in predicting the continuance of using social network sites games. In the context of a mobile social networking service (SNS),
Zhou et al. (2010) showed that flow state has a strong significant impact on SNS users’ loyalty. Similar results concerning the effect of flow state on online behaviour have been demonstrated with various systems involving e-learning (Cheng, 2014; Liu et al., 2009), smartphones (Dery et al., 2014), computer-mediated environment (Hsu et al., 2013), mobile payment (Zhou, 2013, 2014), enterprise resource planning (Choi et al., 2007) and online gaming (Alzahrani et al., 2016). Therefore, flow state has the potential to be integrated into the IS success model to influence user satisfaction.

2.3 Information success

DeLone and McLean’s (2003, 2004) work on theoretical and empirical IS research has resulted in developing a model for IS success that is grounded on three quality factors involving system quality, information quality and service quality, which considerably influence users’ satisfaction and thus maximise IS benefits. The model has been considered one of the most influential and vibrant models in contemporary IS research due to its relevance in examining IS activities at organisational and individual levels. While Delone and Mclean’s success model has received much attention among researchers, very little research has been conducted within the context of social media. Negahban et al. (2016) applied DeLone and McLean’s IS success model to investigate the effects of the mobile customer relationship management dimensions on business performance in South Korea. They indicated that improving quality dimensions will significantly influence business performance. Ramirez-Correa et al. (2017) applied the IS success model to determine the moderating impact of learning styles on the success of learning management systems from a student’s point of view. They indicated that the model can explain the use, user satisfaction and perceived benefits of a learning management system. Moreover, Stefanovic et al. (2016) implied that the IS success model has the capability to examine the satisfaction of employees in government institutions with e-government system implementation. On the contrary, an IS success model was barely used to investigate user satisfaction with social media platforms. Wu and Chen (2015) provided a recent application of IS. They explored the relationship between information quality, system quality, function quality and social influence as key determinants of Facebook usage for educational purposes. The results of this study showed that social influence and information quality were critical and direct determinants that affect users’ continuous intention to use Facebook in learning. In their study to assess blog-based learning systems success, Wang et al. (2014) indicated that the interrelationships between quality dimensions positively affect learning performance. Özata and Er (2015) used the success model to examine determinants of user satisfaction with Facebook as a mobile app in Turkey.

2.3.1 Quality measures: system quality, information quality and service quality

The present study adopts DeLone and McLean’s (2004) quality perspectives, namely, system quality, information quality and service quality. First, system quality focusses on the technical level of success of a system with respect to the design and functionality of a system. Seddon (1997) previously asserted that consistency of the user interface, ease of use, documentation quality and maintenance quality are elements of system quality. However, DeLone and McLean (2004) indicated that crucial aspects of system quality determine the usability, reliability, responsiveness, flexibility, integration, navigation and accessibility of the system. Prior research has suggested that frequent system errors, which enervate navigation or slow processing time, have a sizeable impact on system performance and users’ acceptance (Ahn et al., 2004; Sun, 2010). Although researchers have recognised the role and importance of system quality in contexts such as virtual communities (Zheng et al., 2013), mobile business intelligence (Peters et al., 2016) and three-dimensional television technology (Shin, 2012), few have attempted to investigate its role in social media satisfaction. Second, information quality
has evolved over the last two decades as a prerequisite to users’ understanding of system complexities or implications (Halawi et al., 2007). Information quality captures success at the semantic level through conveying the intended meaning of information. It reflects various standards of content and forms generated by an IS, such as accuracy, precision, currency, reliability, completeness, conciseness, efficiency, meaningfulness, relevance, scope and understandability (DeLone and McLean, 2003; Halawi et al., 2007). Thus, a high level of information quality can reduce system complexity and allow users to process data using appropriate formats and interfaces. In the context of social media, standards of information are user centric and play a crucial role in user acceptance. If the information provided by a social media platform is accurate, comprehensive and frequently updated for current needs, users will feel unperturbed making task-related decisions. Third, service quality refers to the overall supporting services that reflect the values of trustworthiness, meeting users’ needs, responsiveness, personalisation and attentiveness of the system (DeLone and McLean, 2004; Ahn et al., 2007; Wang and Lin, 2012). Prior research has examined the impact of service quality in various settings, including e-service (Udo et al., 2010), virtual community (Elliot et al., 2013), internet systems (Ahn et al., 2007), e-business (Cenfetelli et al., 2005) and mobile services (Wang and Lin (2012). However, research examining the quality of social media service is infrequent.

2.3.2 User satisfaction. Seddon (1997) defined user satisfaction as the net feeling (e.g. pleasure or displeasure) resulting from combining all the benefits or aspirations that a person hopes to receive from interaction with the IS. User satisfaction taps a wider range of system needs and benefits than any other construct (Seddon, 1997). User satisfaction has been extensively underpinned in IS research as a good measure or surrogate of IS success since IS use alone is insufficient to explain the success of IS (DeLone and McLean, 2004; Seddon, 1997). Lee and Chung (2009) have also found that user satisfaction is a common measure of IS success because it can be used for measuring the effectiveness of the system over the course of a long-term relationship with the user. Prior research contended that user satisfaction is a reliable success variable of a wide range of IS applications, such as e-service (Udo et al., 2010), Facebook (Ozata and Er, 2015) and nursing IS (Lin et al., 2016). Thus, user satisfaction is incorporated into the Twitter success model in this study. In line with the preceding review, the IS success model has the plausibility to provide a holistic view on users’ satisfaction with Twitter so that it will be used as a basis for our study and expanded with other constructs that may influence users’ satisfaction.

3. Research model and hypotheses

3.1 Relationship between quality dimensions and perceived usefulness
Perceived usefulness is deduced from the TAM that has received substantial attention since its inception and revision by Davis (1989). Perceived usefulness is regarded in IS research as an influential determinant of accepting various systems and technology applications (Aboelmaged, 2010). Moreover, perceived usefulness has been considered a powerful determinant of social media acceptance including instant messaging and Facebook (Glass and Li, 2013). On the other hand, Service quality is a fundamental factor affecting IS success. It is also found to play a significant role in influencing system’s perceived usefulness through paying more attention to features such as trustworthiness, responsiveness and personalisation (Ahn et al., 2007; Lin, 2007). Research into online shopping adoption has revealed that the dimensions of service quality are far more significant when users perceive the usefulness of an internet system (Ahn et al., 2007). In the e-business environment, Cenfetelli et al. (2005) regarded service quality features as the main determinant of user perceived usefulness of e-business systems. Meanwhile, Wang and Lin (2012) confirmed
that service quality has significant positive effects on perceived usefulness among experienced users of mobile value-added services.

User assessment of perceived information quality is positively related to perceived usefulness in diverse online contexts involving online retailing (Ahn et al., 2007), Internet tax-filing systems (Chang et al., 2005) and virtual communities (Lin, 2007). High-quality information reflects the ability, integrity and benevolence of a service provider, and thus maximises the users’ perceived usefulness. Wang and Lin (2012) confirmed that information quality serves as an important antecedent of perceived usefulness among experienced mobile phone subscribers.

High system quality has been demonstrated to improve system productivity and efficiency, allowing users to perceive a higher level of its usefulness. In line with Lee et al. (2009), system quality involves aspects such as speed of access, interface design and navigational easiness that may influence the users’ evaluation of a mobile applications’ utility. Shin (2012) showed the significant role of system quality in influencing perceived usefulness of three-dimensional television technology. Wang and Lin (2012) established that system quality positively influences users’ perceived usefulness of mobile services. Likewise, Lin and Lu (2000) found that system response time is a predictor of its perceived usefulness. Byrd et al. (2006) argued that perceived usefulness is affected by system quality, as it helps process jobs more efficiently and minimises operation time and cost. For example, if Twitter is operated by healthcare professionals with poor service, insufficient information or slow processing time, they may find it useless. Thus, the hypotheses are developed as follows:

\[ H1. \] Service quality has a significant impact on the perceived usefulness of Twitter.

\[ H2. \] Information quality has a significant impact on the perceived usefulness of Twitter.

\[ H3. \] System quality has a significant impact on the perceived usefulness of Twitter.

3.2 Relationship between quality dimensions and flow state

Providing what is promised to meet users’ specific needs, on time, can augment customers’ knowledge and sharpen their skills as they engage and learn from service providers (Froehle, 2006). Accordingly, high service quality may reduce perceived challenges, improve system engagement, lower online anxiety (Wolfinger and Gilly, 2003) and increase a sense of perceived control (Ding et al., 2007). Chung and Tan (2004) believed that perceived service quality has an important impact on users’ perceived playfulness, which, in turn, signifies flow state. In addition, improved perceptions of information quality attributes delivered through the online application will help users immerse themselves in using that application to achieve flow state. Zhou et al. (2010) and Chau et al. (2000) suggested that information quality presented on the internet and social networks have a significant effect on users’ flow state. Hsu et al. (2012) examined the impact of website quality on customer satisfaction using perceived flow as a mediator. They confirmed that information quality has a positive effect on users’ flow state. Also, Shin (2012) showed that the information content quality of three-dimensional television has a significant effect on users’ flow state. Moreover, poor network reliability, slow response speed, or unreliable platforms may undermine users’ flow state as they wait longer to exchange information with disturbed enjoyment and engagement as well as a lack of control. Peters et al. (2016) demonstrated that engagement is a key to understanding the relationship discrepancy between system quality and the use of mobile technologies. Guo and Poole (2009) and Hsu et al. (2012) also asserted that users’ flow state is influenced by system complexity in an online context. Examining the effect of flow state on users’ loyalty in mobile social networking, Zhou et al. (2010) showed that system quality
significantly affects users’ flow state, which further determines their loyalty. Thus, the following hypotheses are delineated:

- **H4.** Service quality has a significant impact on the flow state of Twitter.
- **H5.** Information quality has a significant impact on the flow state of Twitter.
- **H6.** System quality has a significant impact on the flow state of Twitter.

### 3.3 Relationship between perceived usefulness and user satisfaction

Satisfaction has been considered a key factor in IT/IS success models (DeLone and McLean, 2004; Wu and Chen, 2005; Tong, 2009). Satisfaction represents users’ cumulative feelings while interacting with the service providers as a gap between perceived performance and expectation. As a key factor in the TAM, perceived usefulness reflects the value associated with using an information technology application that users will judge their satisfaction based on. Bhattacharjee’s (2001) expectation–confirmation model has confirmed that perceived usefulness is a significant factor affecting user satisfaction. Chen et al. (2009) developed an integrated model to predict the determinants of user satisfaction of self-service technologies (SSTs). They found that users’ satisfaction with SSTs is significantly influenced by perceived usefulness. Zhou (2011) examined the critical success factors of mobile website adoption by means of structural equation modelling (SEM). The results indicate that perceived usefulness determines user satisfaction with mobile websites. Rai et al. (2002) demonstrated the positive relationship between the perceived usefulness of an enterprise resource planning system and user satisfaction. With regard to e-service, Wen et al. (2011) recognised perceived usefulness as the determining factor of user satisfaction. Similarly, Cenfetelli et al. (2005) stressed that the usefulness of an e-business environment positively influences user satisfaction. In the context of mobile internet, many researchers found that the perceived usefulness of mobile internet sites positively affects user satisfaction (Lee et al., 2007; Zhou, 2014). Twitter is a useful social media application that enables users to communicate instantly at anytime from anywhere using short messages, which may enhance their satisfaction. As a result, **H7** is suggested as follows:

- **H7.** Perceived usefulness has a significant impact on user satisfaction with Twitter.

### 3.4 Linking flow state to satisfaction and perceived usefulness

User engagement is an indication of technology success when flow state positively influences subsequent online attitudes and behaviours, leading directly to satisfaction (O’Cass and Carlson, 2010; Peters et al., 2016). According to Ho and Kuo (2010), satisfaction is attained when system users are able to balance their skills and system challenges while they immerse themselves in an enjoyable flow state. For example, when healthcare professionals plan to use Twitter, they need to have internet skills and knowledge related to Twitter as a social media application. Moreover, they need to deal with challenges concerning privacy, security and operation inconvenience issues. Flow state as a significant predictor of satisfaction has been confirmed in preceding research (Shin, 2006). In his study on nurses’ intentions to continue using a blended e-learning system, Cheng (2014) showed that flow is a significant predictor of nurses’ satisfaction with the system. Hsu et al. (2011) showed that perceived flow positively influences user satisfaction within an online travel context. Previous research has also established that perceived enjoyment, as a part of flow state, can influence user satisfaction with mobile applications (e.g. Lee et al., 2007; Zhou, 2013). In a study on flow in a professional online sporting context, O’Cass and Carlson (2010) established that the flow state is perceived as so enjoyable that users would re-experience it frequently. Similarly, focussed immersion in using instant
Messaging has been recognised to have an impact on user satisfaction (Zaman et al., 2010; Goel et al., 2013; Zhou, 2013). In addition, the state of flow in an online context can be influenced by perceived usefulness (Hsu et al., 2013). Feasibly, flow state can be improved once a user perceives a social media application to be useful. Wolfinbarger and Gilly (2001) argued that an individual will engage, searching and browsing, with a website more often when it is perceived to be useful. Thus, physicians’ perception of usefulness may help filter out irrelevant thoughts and lead to enhanced cognitive absorption and flow state. In line with the preceding argument, we suggest the following hypotheses:

H8. Flow state has a significant impact on user satisfaction with Twitter.

H9. Perceived usefulness has a significant impact on flow state.

Based on the above elaboration, a proposed research model is presented in Figure 1. The model illustrates the research hypotheses through nine paths towards Twitter satisfaction.

4. Methodology
4.1 Development of research instrument

We used well-established scales from prior literature whose validity has been proven, lending high reliability to the results (see Table I). The Likert scale ranging from 1 to 5 is used to operationalise the constructs, where (1) represents “strongly disagree” and (5) represents “strongly agree”. Information quality was assessed by three items that reflect accuracy, comprehensiveness, and how up-to-date the information provided by Twitter is. The items were adapted from Wixom and Todd (2005), Yan et al. (2014) and Zhou (2011). Items of system quality were adapted from Wang and Lin (2012) and Yan et al. (2014). The items represent the reliability of the system, navigation effectiveness and usability of the layout. Items measuring service quality were adapted from Yen and Lu (2008) and Zhou (2011) to manifest the promptness in responding to inquiries, personalisation and the real-time communication of

---

**Figure 1.**
Suggested research model

**Notes:** FL, flow state; IQ, information quality; SAT, satisfaction; SQ, system quality; SVQ, service quality; USE, usefulness
Twitter services. To reflect enhanced work efficiency and effectiveness associated with using Twitter, perceived usefulness was measured using items adapted from Gao and Bai (2014), Hung and Jen (2012) and Cheng et al. (2006). Items representing flow state were adapted from Hsu et al. (2013), Cheng (2014), Zhou (2014) to depict concentration, perceived control and perceived enjoyment associated with using Twitter. Items of satisfaction were adapted from Chen et al. (2009), Gao and Bai (2014) and Lin and Hsieh (2007). The items reflect user’s content, comfortability and positive feeling concerning services provided by the Twitter platform. The second part of the questionnaire was designed to collect basic information about physicians’ demographics, including specialisation, gender, age and usage of Twitter (see the Appendix for a copy of the questionnaire).

### 4.2 Instrument validity and reliability

Cooper and Schindler (2003) suggested that a panel of people can assess how well the instrument meets the standards. Therefore, the scale items were generated based on an extensive review of relevant literature and independently examined by a panel of three professors and five postgraduate students to assess the scale’s representation of relevant constructs. The panel suggested that the measures were appropriate and captured the key factors. Also, the items were translated into Arabic and then a reverse translation into English was independently performed. As a result, minor modifications concerning wording in the English and Arabic versions of the questionnaire were applied to ensure consistency.

### 4.3 Data collection and adequacy

The current study used an online survey administered through Survey Monkey (www.surveymonkey.com). The sampling procedure included sending an e-mail invitation...
message to full-time registered physicians in private hospitals in the United Arab Emirates. The e-mail invitation briefly described the purpose of the study and provided the link to the survey. The recipients were asked to complete the survey questionnaire by a certain time or to forward the e-mail to the most appropriate healthcare professional to respond. The survey was voluntary, and strict confidentiality was assured, as no identifying information was recorded. E-mail reminders were also sent to maximise the response rate. A total of 350 questionnaires were mailed to physicians from a wide spectrum of specialities. Of those, 108 usable questionnaires were completely filled and returned with a response rate of approximately 31 per cent. The response rate was satisfactory given physicians’ busy schedule and limited accessibility due to hospital’s policies. Hair et al. (2010) indicated that a response rate of 30 per cent is sufficient for survey research. According to Podsakoff et al. (2003), Harmon’s one-factor test was provided while collecting the data for the main study to help reduce the potential for common-method bias. It is suggested that the research results were not affected by common-method bias since a single factor did not emerge and one-factor did not account for most of the variance. Multicollinearity was examined using the variable inflation factor (VIF) (Kline, 1998). The maximum VIF for this test was 2.86 which reveals that multicollinearity does not appear to be a significant problem since VIF values for latent constructs do not exceed the threshold level of 5.0 (Hair et al., 2017).

5. Data analysis

5.1 Sample demographics

Respondents represent a wide spectrum of healthcare specialities as shown in Table II. About 50.7 per cent of physicians represent five specialisations including respiratory, surgery, orthopaedic, paediatric and ENT. Also, most of the physicians (82.4 per cent) had more than two years of experience with Twitter, which made a good sample, as respondents are generally known to be familiar with the Twitter platform. The gender distribution of the study subjects was mainly males (71.3 per cent). Respondents above 40 years formed the largest age group (76.9 per cent).

<table>
<thead>
<tr>
<th>Age</th>
<th>n</th>
<th>%</th>
<th>Respondents’ specialisation</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 30</td>
<td>8</td>
<td>7.4</td>
<td>Gastroenterology</td>
<td>9</td>
<td>8.3</td>
</tr>
<tr>
<td>31–40</td>
<td>17</td>
<td>15.7</td>
<td>Endocrinology</td>
<td>5</td>
<td>4.6</td>
</tr>
<tr>
<td>41–50</td>
<td>45</td>
<td>41.7</td>
<td>Respiratory</td>
<td>10</td>
<td>9.2</td>
</tr>
<tr>
<td>&gt; 50</td>
<td>38</td>
<td>35.2</td>
<td>Family medicine</td>
<td>6</td>
<td>5.6</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
<td>100</td>
<td>Geriatric medicine</td>
<td>1</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Surgery</td>
<td>10</td>
<td>9.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>n</th>
<th>%</th>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>77</td>
<td>71.3</td>
<td>Orthopaedic</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Female</td>
<td>31</td>
<td>28.7</td>
<td>Psychiatric</td>
<td>2</td>
<td>1.9</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
<td>100</td>
<td>Obst. and Ggyn.</td>
<td>4</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Paediatric</td>
<td>12</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Urology</td>
<td>3</td>
<td>2.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Twitter experience</th>
<th>n</th>
<th>%</th>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2</td>
<td>19</td>
<td>17.6</td>
<td>Nephrology</td>
<td>6</td>
<td>5.6</td>
</tr>
<tr>
<td>3–5</td>
<td>56</td>
<td>51.9</td>
<td>Ophthalmology</td>
<td>2</td>
<td>1.9</td>
</tr>
<tr>
<td>&gt; 5</td>
<td>33</td>
<td>30.5</td>
<td>Rheumatologist</td>
<td>4</td>
<td>3.7</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
<td>100</td>
<td>ENT</td>
<td>11</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Oncology</td>
<td>5</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Therapist</td>
<td>6</td>
<td>5.6</td>
</tr>
</tbody>
</table>

Table II. Profile of respondents
5.2 Assessment of the measurement model

The SmartPLS 3.0 software was used for analysis through adopting Partial Least Squares (PLS) as a variance-based SEM method. PLS-SEM is preferred over covariance-based SEM methods since it is originally designed for prediction purposes and is less sensitive to sample size (Hair et al., 2017; Lin et al., 2014). Likewise, kurtosis and skewness values of the sample data are within ±1 which indicates no violation of normality assumptions. The quality of measurement model is assessed by multiple measures. Figure 2 visualises the standardised factor loading for each measurement item and the composite reliability (CR) values. All loadings are significant ($p < 0.001$) and exceed the cut-off value of 0.70. The CR values for all variables in the model are higher than 0.84 which demonstrates adequate internal consistency.

Table III shows that Cronbach’s $\alpha$ and rho_A values are greater than 0.72 which indicates an acceptable internal reliability of all variables (Henseler et al., 2016). In addition, the minimum average variance extracted (AVE) value is 0.729 which demonstrates a satisfactory level of convergent validity as it exceeds the required cut-off value of 0.5 (Henseler et al., 2016). Furthermore, the Fornell and Larcker criterion has been applied to assess the discriminant validity of the variables. Table III also demonstrates that the inter-construct correlations are less than the square root of the AVE (as italic diagonal elements) lending a well-maintained quality of the measurement model.

5.3 Hypotheses testing

A bootstrapping technique with 1,000 subsamples in the SmartPLS 3.0 software was used to examine the explanatory power and path significance of the structural model. Model fit
indices involving SRMR = 0.074 (< 0.08), d_{ULS} = 0.93 (< 0.95), d_{C} = 0.702 (< 0.95), \chi^2 = 339.02 and NFI = 0.691 demonstrate a significantly reliable and adequate fit (Henseler et al., 2016). The adjusted $R^2$ indicates that 37 per cent of the usefulness of Twitter is explained by two factors ($t = 4.813, p < 0.001$). As a result, we have found that service quality ($\beta = 0.347, t = 4.307, p < 0.001$) and system quality ($\beta = 0.381, t = 4.269, p < 0.001$) significantly affect the usefulness of Twitter. Therefore, H1 and H3 are supported. In contrast, information quality tends to have no significant impact on perceived usefulness ($\beta = -0.112, t = 1.293, ns$) and flow state ($\beta = 0.018, t = 0.252, ns$). Consequently, H2 and H5 are not supported. Similarly, system quality has no significant effect on the flow state of Twitter ($\beta = 0.079, t = 1.176, ns$). So, H6 is not supported. Moreover, about 0.57 per cent of variance in the flow state ($t = 9.250, p < 0.001$) is explained by perceived usefulness ($\beta = 0.426, t = 4.307, p < 0.001$) and service quality ($\beta = 0.406, t = 5.223, p < 0.001$), providing support for H4 and H9. The impact of perceived usefulness and flow state on physicians’ satisfaction with Twitter explain 0.49 per cent of its variance ($t = 5.725, p < 0.001$). Accordingly, the results indicate that the perceived usefulness of Twitter ($\beta = 0.336, t = 3.355, p < 0.01$) and flow state ($\beta = 0.440, t = 4.789, p < 0.001$) significantly influence physicians’ satisfaction with Twitter, rendering support for H7 and H8. Figure 3 visualises the results of the structural model with path coefficients ($\beta$) and adjusted $R^2$ values. It also highlights (in bold) the supported hypothesised relationships.

6. Discussion and implications
Social media has provided new research opportunities for management and technology. This paper proposes a solid research model that integrates IS success and TAM models with flow theory to investigate how these constructs would influence the satisfaction of health professionals with social media. Healthcare professionals tend to develop different patterns when adopting new technology. Aggelidis and Chatzoglou (2009) argued that healthcare professionals, unlike other users, seem to have different priorities and perceptions concerning implementing new technology. Their adoption process is slower than other users (Lowenhaupt, 2003) and usually ends up with resistance (Paul and McDaniel, 2004). Hence, our study contributes to the body of knowledge related to the professional use of social media in the healthcare context. Clearly, this study enhances our understanding concerning the effects of quality variables (i.e. service quality, information quality, system quality) on perceived usefulness and flow state, which, in turn, influences health professionals’ satisfaction with Twitter as an example of a social media platform. The model is examined by means of PLS-SEM and indicates important findings and implications as follows.

First, contrary to prior research, information quality tends to have no impact on the perceived usefulness and flow state where Twitter is concerned. This is possibly due to the
limited information content in Twitter. Using 140 characters may configure itself more as data than information. Peng et al. (2016) argued that social network messages are characterised by redundancy and noise, which results in the poor quality of information extraction and query-orientation. Information quality could be an influential antecedent in applications where rich content (e.g. images, pictures, and animation), learning needs, styles and knowledge levels are highly considered (e.g. e-learning) (Cheng, 2014). Similarly, system quality tends to have an impact on perceived usefulness only. This contradicts previous research that claims a significant effect of system quality on flow experience (e.g. Guo and Poole, 2009; Hsu et al., 2012; Zhou et al., 2010). This finding implicates that healthcare managers may integrate the Twitter platform into healthcare professional’s day-to-day tasks and work routines as an on-time monitoring platform rather than as a discovery and learning platform which may require high levels of information and system quality to deal with the rich data content. Using Twitter as a monitoring platform by healthcare professionals strengthens professional autonomy as a key professional value that allows a healthcare professional to have control over the medical practice, processes and other non-professional groups (Raelin, 1989; Zuger, 2004). This is in harmony with Walter and Lopez (2008) who depicted that healthcare professionals resist new technology that wears away their professional autonomy and accept technology that reinforces their autonomy.

Second, service quality has emerged as the only quality dimension which demonstrates a significant impact on both perceived flow and usefulness. This reflects the crucial role of on-time services, attention to individual needs, support mechanisms, and security issues for healthcare professionals in shaping the application’s effectiveness concerning promoting health behaviour,
patient care, medical education and development, and sharing professional knowledge. The present evidence is consistent with existing findings implying that new technologies or inventions would be useful to users if they explicitly possessed service quality (Lee and Wu, 2011; Wang and Lin, 2012). In addition, the finding is in harmony with prior studies that relate high service quality to improved flow state through enhancing perceived control, system engagement and anxiety management (Chung and Tan, 2004; Ding et al., 2007; Wolfinger and Gilly, 2003). Similarly, Hsu et al. (2012) noted that service quality is a more influential factor to flow state than system and information quality in the context of online services.

Third, the result of the proposed model also highlights the significant role of perceived usefulness in moulding physicians’ flow state and satisfaction with Twitter. So, no wonder that perceived usefulness is a prominent extrinsic motivational construct that has been integrated into various IS models including Bhattacharjee’s (2001) expectation–confirmation model. It also opens a venue for a relationship between usefulness and flow state that is barely investigated in social media research. This is consistent with previous research which demonstrates that perceived usefulness is the strongest determinant of users’ satisfaction with online systems (Hsu et al., 2013) and microblogging platforms (Barnes and Böhringer, 2011; Shu, 2014). The finding also challenges prior research on microblogs that supports an insignificant relation between flow experience and satisfaction (e.g. Shu, 2014).

Fourth, flow state tends to demonstrate an essential effect on healthcare professional’s satisfaction with Twitter. This opens a venue for a relationship that is barely investigated in IS research and extends Agarwal and Karahanna’s (2000) work from the World Wide Web to social media platforms. These findings challenge the conventional argument that professionals have no time to experience flow while using Twitter at work due to their job nature and busy schedule. The result also contrasts the findings of Özata and Er (2015) who established that flow state has an insignificant effect on user satisfaction with Facebook. However, this finding may lead to a substantial observation concerning the significant role of flow state in social media, similar to its impact in other online contexts (e.g. Cheng, 2014; Goel et al., 2013; Zhou, 2013). An interpretation for this finding may lie in the fact that flow state reflects both goal-oriented and joyful absorption that usually influence satisfaction.

The findings of this study have various theoretical and practical implications for social media researchers, professional users and vendors. For researchers, the study demonstrated that PLS-SEM as a variance-based method can provide a good opportunity for researchers to examine predictive relationships using small sample size with less sensitivity to normality assumptions. The study also offers a good understanding of the factors that influence post-adoption of social media by healthcare professionals. This contributes to current technology adoption frameworks in professional contexts that can guide future research to expand our model to examine the role of professional values (i.e. professional autonomy) in the adoption process. Moreover, the inclusion of perceived usefulness and flow state provide balanced motivational factors that signify both intrinsic and extrinsic experience while interacting with a system. These two factors can be widely used to determine users’ post-adoption satisfaction with various types of technologies. From the practical perspective, healthcare managers first need to spot the importance of Twitter in the healthcare context, then allocate enough resources to acquire or develop social media applications that are compatible with professionals’ expectations, service requirements, support mechanisms and security, rather than being concerned only with the content and functionality of the application. In addition, healthcare professionals should consider skills, task control and autotelic experience that reinforce flow experience while adopting social media platforms in the workplace. Health managers should also pay careful attention to issues concerning ethics, legitimacy and privacy expectations that may influence satisfaction with Twitter in the healthcare context, particularly when using Twitter in health campaigns. For vendors, Twitter service and design should reflect the specific
needs of healthcare professionals to monitor and control healthcare performance. Thus, non-interrupted services, fast response to inquiries, easy navigation and layout design that support interaction with nurses, patients and other partners in both usual situations and outbreaks should be considered.

7. Research limitations and directions for future studies
Even though this research offered meaningful implications, the findings must be interpreted with caution. First, data collection focused on a particular social media, Twitter, while other platforms such as Facebook or medical social networks such as Sermo.com and AllNurses.com may bring different findings. Thus, future research could attempt to apply the model to other social media platforms to verify the findings and broaden its scope.

Second, flow state is an elusive multidimensional concept formed from joy, discovery, control and reduced awareness of time and surroundings (Pace, 2003). The operationalisation of the flow construct has been queried in former research with the lack of a universal viewpoint concerning its components (Zhou et al., 2010). In addition, Hoffman and Novak (2009) noted that what one researcher considers as an antecedent of flow state, another considers a consequence of flow or perhaps a part of flow itself. For instance, our study successfully verified that perceived usefulness is a significant consequence of flow state, while a reverse relationship between the two variables can also be confirmed (e.g. Hsu et al., 2013). Therefore, future research should pay more attention to the multifaceted nature of the concept and possible inclusion of other variables that may associate with flow state, such as attention (Ozata and Er, 2015), interactivity (Wu and Chang, 2005), enjoyment (Agrawal and Karahanna, 2000; Ozata and Er, 2015), knowledge generation and utilisation (Holsapple and Wu, 2008), playfulness (Hsu et al. 2011) and trust (Zhou et al., 2010). Third, this research has been conducted in a developing context where the professional use of social media is still immature. Thus, future research needs to recognise other developed contexts that have reached a maturity stage in using social media. Similarly, users’ perception is a dynamic concept that varies from time to time as experiences accumulate thus, conducting longitudinal studies at several points in time would offer more insights into user’s satisfaction. Imminent research can also thematically analyse social media accounts and differentiate between personal and professional feeds. This study will enhance our understanding of professional intervention in social media and move beyond problems associated with measuring perceptions. Fourth, the boundaries between institutions and experts are blurring within the social media context. Therefore, upcoming studies may focus on mechanisms that help institutionalise social media governance in the workplace beyond the traditional scope of IT solutions. Finally, a substantial volume of retweets is an evident feature of healthcare communication. Thus, imminent research may differentiate between tweets and retweets while examining user’s engagement behaviour or social media acceptance models. In the same vein, exploring the issue of trustworthiness in tweets and retweets would be of great value for social media research, particularly when rumours and true news are mixed in crucial circumstances, and the dynamic relationship between patients and caregivers is influenced by authoritative styles.

References


Predicting the success of Twitter in healthcare


Further reading


Appendix. Questionnaire

PART 1: For each of the item below, please circle the number that best describes the extent to which you agree or disagree with the following statements

(1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree)

(1) Information provided through Twitter is accurate   1  2  3  4  5
(2) Overall, Twitter is useful   1  2  3  4  5
(3) Twitter system is free of operating errors   1  2  3  4  5
(4) Information provided through Twitter is sufficient   1  2  3  4  5
(5) Twitter service meets my specific needs   1  2  3  4  5
(6) Twitter increases my work effectiveness   1  2  3  4  5
(7) Information provided through Twitter is up-to-date   1  2  3  4  5
(8) The navigation of Twitter system is effective   1  2  3  4  5
(9) Twitter help me acquire the information I need   1  2  3  4  5
(10) When using Twitter, I feel in control   1  2  3  4  5
(11) Twitter provides real time communication service   1  2  3  4  5
(12) I find using Twitter to be enjoyable   1  2  3  4  5
(13) The layout of Twitter system is clear   1  2  3  4  5
(14) I feel comfortable dealing with Twitter   1  2  3  4  5
(15) I am satisfied with Twitter   1  2  3  4  5
(16) Twitter service reacts to my inquiries in a timely manner   1  2  3  4  5
(17) My attention is absorbed in what I am doing while using Twitter   1  2  3  4  5
(18) I like the services provided by Twitter   1  2  3  4  5

PART 2: Please indicate the appropriate demographic information

(19) Your age ___________ years
(20) Gender (tick √ where applicable): Female __ Male __
(21) Your medical specialization: _______________
(22) Your experience in using Twitter: _______ years.

About the author
Mohamed Gamal Aboelmaged has a PhD in Management Science from Lancaster University, UK, and MA in Public Policy & Administration from the International Institute of Social Studies, Erasmus University Rotterdam, the Netherlands. Currently, he is Associate Professor of Management at University of Sharjah, UAE. His research interests include social media in workplace, sustainability, lean and enterprise systems. His work has been published in international journals and conference...

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
What drives internet users’ willingness to provide personal information?
Ching-Hsuan Yeh and Yi-Shun Wang
Department of Information Management,
National Changhua University of Education, Changhua, Taiwan
Shin-Jeng Lin
Department of Management, Leadership and Information Systems,
Le Moyne College, Syracuse, New York, USA
Timmy H. Tseng
Department of Information Management,
National Changhua University of Education, Changhua, Taiwan
Hsin-Hui Lin
National Taichung University of Science and Technology, Taichung, Taiwan, and
Ying-Wei Shih and Yi-Hsuan Lai
Department of Information Management,
National Changhua University of Education, Changhua, Taiwan

Abstract
Purpose – Considering that users’ information privacy concerns may affect the development of e-commerce, the purpose of this paper is to explore what drives internet users’ willingness to provide personal information; further, the paper examines how extrinsic rewards moderate the relationship between users’ information privacy concerns and willingness to provide personal information.
Design/methodology/approach – Data collected from 345 valid internet users in the context of electronic commerce were analyzed using the partial least squares approach.
Findings – The result showed that agreeableness, risk-taking propensity and experience of privacy invasion were three main antecedents of information privacy concerns among the seven individual factors. Additionally, information privacy concerns did not significantly affect users’ willingness to provide personal information in the privacy calculation mechanism; however, extrinsic rewards directly affected users’ disclosure intention. The authors found that extrinsic rewards had not moderated the relationship between users’ information privacy concerns and their willingness to provide personal information.
Originality/value – This study is an exploratory effort to develop and validate a model for explaining why internet users were willing to provide personal information. The results of this study are helpful to researchers in developing theories of information privacy concerns and to practitioners in promoting internet users’ willingness to provide personal information in an e-commerce context.
Keywords Big Five personality, Experience of privacy invaded, Extrinsic rewards, Information privacy concern, Risk-taking propensity, Willingness to provide personal information
Paper type Research paper

Introduction
It has been a common practice for internet users to provide their personal information in exchange for website services access. Users become registered members and are allowed to use a website’s services that are either personalized (system-tailored) or customized (user-tailored) by offering limited or detailed demographics or personal preference information, limited (Sundar and Marathe, 2010). While users may continue to enjoy excellent website services, information privacy concerns are increasingly at stake (Liu et al.,
Information privacy concerns arise when a user subjectively perceives a threat resulting from his/her personal information being intruded upon in one or more of the following ways: improper access, unpermitted collection, unauthorized secondary use and incorrect capture (Smith et al., 1996).

After an interdisciplinary review of literature on privacy concerns, Smith et al. (2011, p. 989) proposed the antecedent–privacy concerns–outcome (APCO) model. Application of the APCO model revealed that once users perceive that their privacy is likely invaded, they might be unwilling to share their personal information with websites. However, this negative relationship may not always hold true; users may actually disclose personal information despite their privacy concerns (Lee and Cranage, 2011). Such a phenomenon is known as the privacy paradox (Norberg et al., 2007). The privacy paradox is relevant to the web environment because an increasing number of websites, such as Amazon, invite users to provide personal information and enjoy personalized services (Garfinkel et al., 2008). Researchers explained this paradoxical phenomenon with the privacy calculus theory: extrinsic benefits or rewards such as price discounts should additionally be taken into an account in reaching the decision to provide personal information (Wang and Wu, 2014). Likewise, privacy disclosure behavior is the function of privacy benefits and privacy costs (Dinev et al., 2008). Recent studies have proposed that different levels of information processing (i.e. central or peripheral) exist when users perform privacy calculus (Angst and Agarwal, 2009). When users perform central information processing, they are likely to focus on the interaction of benefits and costs; in contrast, when users perform peripheral information processing, they may only consider the main effects of benefits and costs or one of them. However, few studies have examined privacy calculus mechanisms while considering different levels of information processing. Hence, the first objective of this research is to examine whether privacy benefits and privacy concerns facilitate the intention to disclose personal information independently or interactively.

Moreover, a variety of factors (i.e. individual, institutional and cultural) have been found as the antecedents of privacy concerns (Li, 2014; Smith et al., 2011). Of these factors, individual factors, such as personality, are mostly examined but have inconclusive results (Smith et al., 2011). It has been suggested that more research is needed to clarify and better understand the exiting contradiction (Li, 2014, p. 348). Furthermore, personality traits were indicated by Li (2012) as antecedents to all other personal beliefs. Hence, the second objective of this research is to investigate the relationships between individual factors and privacy concerns. Specifically, general personality (i.e. Big Five typology) and risk-related factors (i.e. risk-taking propensity and experience of privacy invasion) are considered for individual differences.

Li (2012) proposed an integrated theoretical framework of online information privacy concerns. The framework is a useful guide to research because it is developed based on a comprehensive review of theories in the area of online information privacy and in the website context (Li, 2012). Furthermore, a portion of the framework includes theories concerning the individual factors and the privacy calculus of online information privacy concerns. This study chooses a portion of Li’s (2012) framework to develop the research model for the following reasons. First, since the aim of this study is to examine the privacy calculus mechanism with empirical evidence on how privacy concerns are determined by individual factors, the individual factors and privacy calculus parts of Li’s (2012) framework are consistent to the objectives of current research. Second, although both individual factors and privacy calculus have been widely examined in the past studies, the inconsistent results of past studies regarding the effects of individual factors on privacy concerns and the complex psychological process concerning privacy calculus make these two parts of Li’s (2012) framework worth further investigation (Angst and Agarwal, 2009; Li, 2012; Smith et al., 2011).
The rest of the paper is organized as follows. First, relevant literature is reviewed and hypotheses are developed based on privacy calculus theory and personality theories. Next, methodology and research findings are described. Finally, theoretical implications, managerial implications, limitations and suggested future research are discussed.

Literature review and hypotheses development

Theories of online information privacy concerns

Li (2012) systematically reviewed past literature and indicated that theories had examined the four aspects of online information privacy concerns: the origin, influential factors, trade-offs and behavioral consequences of online information privacy concerns. The agency theory and the social contract theory have been applied to address the origin of online information privacy concerns. Influential factors of online information privacy concerns can be divided into institutional factors (i.e. procedural fairness theory, social presence theory and social response theory) and individual factors (i.e. protection motivation theory, information boundary theory, social cognitive theory and personality theories). The privacy calculus theory examines the trade-offs of online information privacy concerns. The theory of reasoned action and the theory of planned behavior have been adapted to investigate the behavioral consequences of online information privacy concerns. Since the first objective of this research is to investigate the mechanism of privacy calculus (i.e. independently or interactively) and the second objective is to examine whether individual factors influence online information privacy concerns, the privacy calculus theory and theories reflecting individual factors (e.g. the Big Five model) are utilized to develop the research model.

Privacy calculus, privacy concerns and privacy benefits

The privacy calculus theory (Culnan and Bies, 2003) is one of the most plausible explanations for the privacy paradox. Privacy calculus is defined as a cognitive process where people evaluate the future consequences of present choices by weighing the potential costs and benefits of sacrificing privacy in order to gain better outcomes (Angst and Agarwal, 2009; Li et al., 2010). Based on the principle of utility maximization, the privacy paradox may occur if the privacy benefits outweigh the costs. Privacy costs in the context of e-commerce relate to the perceived risks of opportunistic behavior as related to the disclosure of internet user personal information submitted in general (Dinev and Hart, 2006); information privacy concerns are merely one of many forms of privacy costs. Privacy benefits, in contrast, denote the overall evaluation of favorable consequences in disclosing personal information; three popular benefits that a user may acquire in return for disclosing his/her personal information are financial rewards, personalization and social adjustment benefits (Smith et al., 2011). The privacy calculus theory offers an excellent theoretical perspective by introducing the idea of trade-off between major drivers and inhibitors for the disclosure of personal information. Some past studies have focused on how to reduce privacy concerns (i.e. privacy costs) through privacy benefits. However, Min and Kim (2015) proposed that both privacy benefits and costs had to be included in the research framework of privacy calculus in order to examine their effects on intention to provide personal information.

The relationship between information privacy concerns and willingness to disclose personal information has been largely examined in the e-commerce context. Past studies have indicated that users with more information privacy concerns would show less intention to share their personal information at a website providing health information services (Zimmer et al., 2010) and an e-store (Faja and Trimi, 2006). Privacy concerns had negative effects on intention to give personal information (Min and Kim, 2015) and
willingness to show the profile information to the public (Feng and Xie, 2014) in the context of social network sites. Therefore, we hypothesize the following:

H1. Information privacy concerns are negatively related to willingness to provide personal information.

Privacy benefits are expected to positively influence willingness to provide personal information. Past studies had indicated that intention to disclose personal information was positively related to the amounts of monetary incentive within the contexts of purchasing mobile computing products (Hui et al., 2007) and financial portals (Hann et al., 2007). Min and Kim (2015) focused on social benefits within the social networking realm; they indicated that the motivation for relationship management through SNS and perceived usefulness of SNS for self-presentation facilitated intention to give personal information. Therefore, we hypothesize the following:

H2. Extrinsic rewards are positively related to willingness to provide personal information.

As per the privacy calculus theory, willingness to disclose personal information is a trade-off function of costs (i.e. information privacy concerns) and benefits (i.e. extrinsic benefits such as money). However, the phenomenon can be more complicated, as extrinsic rewards may facilitate willingness to provide personal information directly and indirectly (i.e. moderating) (Park et al., 2012). Li (2012, p. 477) provided empirical evidence that users’ perception of privacy intrusion may be lessened when provided with benefits; this indicated that extrinsic rewards may mitigate the negative association between information privacy concerns and intention to disclose personal information. Therefore, we hypothesize the following:

H3. When extrinsic rewards are higher, the negative relationship between information privacy concerns and willingness to provide personal information is weaker.

Antecedents of information privacy concerns

General personality: Big Five model. Personality is an individual trait which means an enduring disposition or propensity toward a given event or object (Kim and Jeong, 2015). In privacy studies, personality has been identified as a major determinant of information privacy concerns in that an individual trait could impact information processing style as well as a belief system (Mount et al., 2005; Smith et al., 1996).

Extensive academic efforts have been made in figuring out representative personalities to explore human nature. Five fundamental personalities have been investigated with most interests: agreeableness, conscientiousness, extraversion, neuroticism and openness. The five personalities emerged as a widely accepted personality typology (i.e. Big Five model) in the late 1980s (Costa and McCrae, 1992). Based on numerous empirical evidences, the structures or dimensions of the Big Five personality were consolidated (Junglas et al., 2008).

Multiple studies have attempted to understand how each of the five personalities influence individuals’ information privacy concerns. First, agreeableness refers to a predisposition to maintain harmonious relations with others, including the avoidance of interpersonal conflicts (John and Srivastava, 1999). The agreeable persons would rather sacrifice themselves and show compliance than to be combative. Highly agreeable persons show high trust in others in social interactions for building positive interpersonal relationships. Junglas et al. (2008) showed that people who are more conforming are inclined to think that others’ actions will not cause privacy threats; hence, their information privacy concerns would be low. Therefore, we hypothesize the following:

H4. Agreeableness is negatively related to information privacy concerns.
Conscientiousness is an individual’s tendency to be goal directed and to maintain self-control (John and Srivastava, 1999; McCrae and Costa, 2008). By nature, conscientious people may be more easily aware of the threats underlying the disclosure of personal privacy (Bansal et al., 2010; Junglas et al., 2008), as they tend to deliberate in their encounters. As a result, more conscientious people may be more likely to have greater information privacy concerns. Therefore, we hypothesize the following:

H5. Conscientiousness is positively related to information privacy concerns.

Extraversion is characterized as sociability and engagement of positive emotions in daily life (John and Srivastava, 1999; McCrae and Costa, 2008). Extraverted persons may enjoy social gathering and are full of energy in such interactions (Judge et al., 2002). Considering that extraverted people may favor social interactions and be outgoing, they are less inclined to be concerned about their privacy; there is the increased possibility of personal information exchange in their relations (Bansal et al., 2010; Junglas et al., 2008). It has been found that extraverts may pursue stimulation over having concerns about privacy threats (Badgaiyan and Verma, 2014). Therefore, we hypothesize the following:

H6. Extraversion is negatively related to information privacy concerns.

Neuroticism indicates the extent to which an individual’s emotions are imbalanced across situations (John and Srivastava, 1999; McCrae and Costa, 2008). Individuals with high neuroticism are emotionally unstable and maladjusted. This emotional instability may then lead to vulnerability where neurotic people are prone to experience a series of negative feelings, such as anxiety, anger, depression, embarrassment and stress. The nature of worrying strongly underlies neuroticism, which could fuel information privacy concerns. Given that neurotic people are more likely to concentrate on negative emotions, they may more easily perceive higher information privacy concerns; feelings of possible losses might take center stage as a result (Junglas et al., 2008). Therefore, we hypothesize the following:

H7. Neuroticism is positively related to information privacy concerns.

Openness to experience is defined as an individual’s tendency to have novel experiences in various ways, such as through fantasy, aesthetics, feelings, actions, ideas and values (John and Srivastava, 1999; McCrae and Costa, 2008). More open people have a stronger intellectual curiosity and like to try unfamiliar activities like adventure. Owing to the curiosity characteristic, it is expected that people with high openness would more likely disclose their personal information for novelty experience, i.e. personalization (Bansal et al., 2010). Therefore, we hypothesize the following:

H8. Openness to experience is negatively related to information privacy concerns.

Specific personality: risk-taking propensity. Risk-taking propensity translates to the level of an individual’s tolerance when facing risk or uncertainty (Blais and Weber, 2006; Chen et al., 2011). Individuals with high risk-taking propensity will more likely embrace the risk and are prone to engage in progressive decisions with uncertainty (Fogel and Nehmad, 2009). Risk-taking propensity is recognized as a personality trait specific to risk-related situations.

Given that people with high risk-taking propensity have a more positive attitude toward risk, there is a greater likelihood for them to engage in riskier or more aggressive activities and to have less information privacy concerns. The negative linkage between risk propensity and risk perception was empirically confirmed (Chen et al., 2011). Accordingly, we hypothesized the following:

H9. Risk-taking propensity is negatively related to information privacy concerns.
Experience of privacy invasion. Individual experiences regarding privacy invasion may contribute to one’s overall belief system as well as their willingness to disclose personal information. As many researchers have observed (Awad and Krishnan, 2006; Smith et al., 1996), negative privacy experiences may result in stronger privacy concerns. People who had suffered from privacy infringement such as unauthorized use their collected personal information would be more aware of the potential risks when prompted to reveal their personal profile (Li, 2014). The more experience of privacy invasion people had, the lower the expectation that their personal information could be managed properly by EC stores; thus, the greater privacy concerns would emerge in such instances (Bansal et al., 2010). As demonstrated empirically in prior studies in the e-commerce context, past experience of privacy invasion was positively linked to information privacy concerns (Bansal et al., 2010; Li, 2014). Therefore, we hypothesize the following:

H10. Experience of privacy invasion is positively related to information privacy concerns.

The research model of this study is shown in Figure 1.

Methods
Measurement
A questionnaire-based survey was conducted for data collection. To effectively probe respondents’ opinions, the questionnaire was structured in a concise format which included nine sets of measures for the focal constructs in our research model. All the measures were adopted from prior studies in that their reliability and validity have been confirmed. Personality, which illustrates five distinctive individual characters, was adapted from the Big Five Inventory scale (John et al., 1991); each sub-personality consisted of five items. To reduce the negative effect of response fatigue, some items were stated in the reverse form (John et al., 1991). The eight-item risk-taking propensity scale was adapted from Weber et al.’s (2002) work in the social domain. Information privacy concerns was measured with four items derived from the four possible privacy invasions in Smith et al.’s (1996) study. Measures of extrinsic rewards (two items) were adapted from Bock et al. (2005).
Based on Lee and Cranage (2011), willingness to provide information was assessed with seven items. Finally, respondents were asked to recall how many times they had suffered privacy invasion; these data are used as the measure of experience of privacy invasion (Xu et al., 2011). A panel of academics and professionals on information management were consulted for item and wording refinement. All items were reflectively modeled and scaled in a seven-point Likert form ranging from 1 (strongly disagree) to 7 (strongly agree). Table AI show the measurements used in this study.

Data collection and consumer profile

Given that users’ willingness to provide personal information over the internet is our main concern, it is justified to invite volunteer participants online. The recruitment information was posted on the biggest bulletin board system in Taiwan (telnet://ptt.cc). The surfers who had experience of disclosing personal information online were invited to participate in this study by accessing to the questionnaire on a professional survey website (www.mysurvey.tw/). The sample participants were asked to recall the most recent experience of providing their personal information to a website in exchange for access to its services; they answered survey questions based on that experience. Thus, our respondents can provide reliable and valid survey results because they answer questions based on their actual experience. To encourage participation, a respondent who completes a valid and useable response is eligible to win a lottery ticket valued at NTD 100 as an incentive. The survey was conducted in December 2013 through March 2014. Within four months of data collection, 396 participants submitted their responses, where 345 were found to be valid. As Haenlein and Kaplan (2004) suggested that the sample size should be larger than 150 or be at least ten times the numbers of independent variables while using the partial least squares (PLS) method, our sample satisfied this requirement.

Of the valid sample, 183 (53.04 percent) were male and 162 (46.96 percent) were female. The gender balance was acceptable. Our respondents were young and the mean age was 24.52 (SD = 5.59) years old. As for the job status, 146 (42.32 percent) were employed, 20 (5.80 percent) were unemployed and 179 (52.88 percent) were students. In total, 209 (60.58 percent) respondents reported their monthly income was under NTD 20,000, and 125 (36.23 percent) were between NTD20,001–50,000. The average internet usage hours per valid participant was found to be 33.53 h per week. More importantly, 114 (33.04 percent) respondents indicated that their personal information have never been invaded; the following figures for the remaining participants’ amounts of experiences of privacy invasion were 59 (17.10 percent) one time, 47 (13.62 percent) two times, 40 (11.59 percent) three times, 11 (3.12 percent) four times, 7 (2.03 percent) five times and 67 (19.42 percent) six times or more.

Data analysis

The structural equation modeling (SEM) is a statistical technique for testing proposed hypotheses simultaneously. PLS, which is a variance-based approach of SEM, was utilized for data analysis in this study. Compared with traditional covariance-based SEM, PLS demands lower assumptions on the measurement distributions, residual distributions and sample size (Urbach and Ahlemann, 2010). Within the objective of minimizing the variance of all the dependent variables, the estimation algorithm of PLS intends to acquire the proxy scores of the latent variables from the measurement model (outside approximation) and the structural model (inside approximation), respectively. Once the proxy scores get stable (i.e. not significantly changing in both approximations), the iterative estimation procedures are terminated.

SmartPLS 2.0 M3 was employed to carry out the PLS analysis (Ringle et al., 2005). The estimation of standardized path coefficients was based on the path weighting scheme, while the significance level was the result of 1,000 bootstrapped samples with construct level changes (Tenenhaus et al., 2005).
Results

The measurement model

For the evaluation of the measurement model, the factor loading of each item was examined first. The acceptable value of the factor loading should be greater than 0.6 (Hair et al., 2010), which indicates that the item variance could be highly explained by its corresponding construct. Based on this threshold criterion, eight measured items were removed, resulting in four for agreeableness, four for conscientiousness, four for extraversion and three for risk-taking propensity after item purification. Next, the reliability of the constructs was examined by using the composite reliability (CR) value whose acceptable value is suggested to be higher than 0.7 (Hair et al., 2010). As Table I presented, the CR value of the nine constructs ranged from 0.74 to 0.93, suggesting that a strong internal consistency existed in all the constructs. Convergent validity of the constructs was evaluated in terms of average variance extracted (AVE), which is the average squared factor loadings. AVE should exceed 0.5 for all items measuring a construct to actually tap into the same construct (Hair et al., 2010). Our results showed that the AVE values were greater than 0.5 except for risk-taking propensity (0.48). The convergent validity was primarily verified.

Finally, discriminant validity, which refers to the degree to which measures of different constructs are distinct, was analyzed with the Fornell–Larcker criterion and cross-loadings (Urbach and Ahlemann, 2010). According to the Fornell–Larcker criterion, the square root of AVE of each construct should be higher than its corresponding correlations, which was the case with our results. Our results also showed that each item successfully loaded on its construct with the highest loading and that no cross-loading existed. Thus, the discriminant validity was then satisfied. As the results of the measurement model indicated, validity and reliability were both achieved this confirmed that the sample participants had provided reliable and valid survey results.

The structural model

The coefficients of determinants ($R^2$), significance of path coefficient and effect size were used as indicators of model validity within the sample (Urbach and Ahlemann, 2010). As Figure 2 presented, 14.86 percent of the variance of willingness to provide personal information was accounted for by its determinants in the research model. Extrinsic rewards ($\beta = 0.35, p < 0.001$) were positively related to willingness to provide personal information.

| CR AG CO EX NE OPE RTP EPI IPC ER WP |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AG | 0.90 | 0.84 | | | | | | | | |
| CO | 0.89 | 0.59 | 0.82 | | | | | | | |
| EX | 0.89 | 0.50 | 0.45 | 0.82 | | | | | | |
| NE | 0.88 | −0.56 | −0.55 | −0.70 | 0.77 | | | | | |
| OPE | 0.90 | 0.52 | 0.50 | 0.60 | −0.57 | 0.80 | | | | |
| RTP | 0.74 | 0.08 | 0.06 | 0.23 | −0.15 | 0.18 | 0.70 | | | |
| EPI | 1.00 | −0.04 | −0.08 | −0.01 | 0.02 | −0.01 | 0.03 | 1.00 | | |
| IPC | 0.86 | 0.34 | 0.18 | 0.13 | −0.16 | 0.16 | 0.22 | 0.12 | 0.78 | |
| ER | 0.93 | −0.01 | 0.03 | 0.10 | −0.07 | 0.12 | 0.20 | 0.12 | 0.03 | 0.93 |
| WP | 0.88 | 0.02 | −0.01 | 0.08 | −0.06 | 0.06 | 0.19 | 0.12 | 0.13 | 0.35 | 0.71 |

Table I.
Descriptive statistics, composite reliability, average variance extracted and discriminant validity

Notes: AG, agreeableness; CO, conscientiousness; EX, extraversion; NE, neuroticism; OPE, openness; RTP, risk-taking propensity; EPI, experience of privacy invasion; IPC, information privacy concerns; ER, extrinsic rewards; WP, willingness to provide personal information. Diagonals are the square root value of average variance extracted and off-diagonals are the correlations. EPI was measured using a single item, such that the CR and AVE were 1.00
However, neither the relationship between information privacy concerns and willingness to provide personal information nor the moderating effect of extrinsic rewards on such a relationship was significant. Thus, \( H2 \) was supported, however, \( H1 \) and \( H3 \) were rejected.

The above results showed that privacy calculus is not a trade-off process between privacy concerns and extrinsic rewards; this is because intention to disclose personal information is solely affected by extrinsic rewards; furthermore, there was no interaction effects between privacy concerns and extrinsic rewards.

There was 18.04 percent variance in information privacy concerns. Of the antecedents, only experience of privacy invasion (\( \beta = 0.13, p < 0.01 \)) was found to have a positive effect on information privacy concerns as hypothesized. Thus, \( H10 \) was supported. The relationships between agreeableness and information privacy concerns (\( \beta = 0.40, p < 0.001 \)) and between risk-taking propensity and information privacy concerns (\( \beta = 0.21, p < 0.001 \)) were significant; however, their impact direction were positive and against our hypotheses. The rest of the antecedents (i.e. conscientiousness, extraversion, neuroticism and openness to experience) did not have a significant effect on information privacy concerns. As such, the results for \( H5, H6, H7, H8, \) and \( H9 \) were not supported. As the results on antecedents to information privacy concerns revealed, the Big Five model of personality in predicting information privacy concerns was not supported. Finally, the effect sizes of two dependent variables were calculated. Willingness to provide personal information (\( f^2 = 0.17 \)) and information privacy concerns (\( f^2 = 0.22 \)) were successfully predicted, respectively, with medium effect sizes (Cohen, 1988); this indicated that the overall structural model had validity.

**Discussion**

**Theoretical implications**

This study contributes to literature and theories on disclosing personal information by empirically testing a portion of Li's (2012) framework in web contexts. Specifically, the privacy calculus mechanism on how privacy concerns are determined by individual factors was empirically examined. Furthermore, this study contributes to the privacy calculus
theory by investigating whether privacy concerns and privacy benefits influence information disclosure intention independently or/and interactively.

The results showed that extrinsic rewards positively influenced internet users’ willingness to provide personal information while privacy concerns had no significant effect on willingness to provide personal information. Past studies had indicated that privacy calculus was the process of considering benefits and costs simultaneously (i.e. Xu et al., 2011); however, this research revealed that privacy calculus was largely determined by privacy benefits while privacy costs had no role to play. Such results challenged the underlying assumption of the privacy calculus theory. Kokolakis (2017) reviewed literature on the privacy paradox and found that about one-third of the studies on this topic had provided evidence that challenge the privacy paradox hypothesis; this included positive correlations that were in opposition to the definition of the information privacy paradox as well as had no correlations at all. Based on the above, the underlying assumption of privacy calculus is questionable; it is likely that privacy calculus is largely determined by benefit considerations.

The results indicated that extrinsic rewards were a dominant factor in facilitating website users’ willingness to provide personal information. Chelappa and Sin (2005) found that the effect of personalization benefits was nearly twice of that of privacy concerns; Min and Kim (2015) indicated that the effect of privacy concerns was greater in two of three social benefits (i.e. motivation of relationship management through SNS and perceived usefulness of SNS for self-presentation) in facilitating intention to disclose personal information. Hence, our finding that privacy benefits were more influential than privacy costs in privacy calculus was consistent with Chelappa and Sin (2005), however, was not in line with Min and Kim (2015). The different findings of the relative importance of privacy benefits vs costs could be ascribed to the different research contexts. Chelappa and Sin (2005) and our research were conducted in the contexts of online websites; meanwhile Min and Kim (2015) was conducted in the context of social networking sites.

To better understand the mechanism of privacy calculus, this study examined whether extrinsic rewards embodying the privacy benefits and privacy concerns embodying the privacy costs may facilitate user willingness to provide personal information independently or/and interactively. The results indicated that the interaction effect of extrinsic rewards and privacy concerns and the main effect of privacy concerns are insignificant; that is, the mechanism for privacy calculus functioned solely through privacy benefits. This suggests that online website users perform privacy calculus in a less complex manner. Past studies on heuristic-systematic model indicated that whether individuals’ process information using heuristic processing or systematic processing depended on the level of knowledge (Chaiken et al., 1989). When individuals have a higher level of knowledge, they tend to process information systematically; however, when individuals have a low level of knowledge, they tend to process information heuristically. It is evident that many online users are less knowledgeable on privacy issues; further, they tend to view the default option as the easiest choice when registering on a website. Many firms cash in on the inattention and cognitive laziness of website users to obtain personal information (Tezinde et al., 2002). Thus, it is likely that when website users perform privacy calculus, they tend to use heuristic processing and focus on privacy benefits mostly. An interesting future avenue of research is to explore different types of heuristics processing used in privacy calculus.

The finding that only one of Big Five personality traits (i.e. agreeableness) had a significant effect on information privacy concerns seems to suggest that personality traits are not strong antecedents; however, most past studies that investigated the effects of personality on privacy concerns have also been inconclusive (Smith et al., 2011). Such inconclusive results may be attributed to the omission of situational factors. The theory of situational strength postulates that implicit and explicit cues provided by external entities can hint to the desirability of potential behaviors; such hints or cues can influence an
individual to respond in a particular manner (Meyer et al., 2010). Specifically, the effects of global personality traits (i.e., Big Five personality traits) are likely to be attenuated under stronger situations, as indicated by situations which have clear consequences. As the results of this research indicated, because willingness to provide personal information was dominated by extrinsic rewards, this reflected a stronger situation in which providing personal information could lead to positive outcomes. Thus, it is likely that the effects of Big Five personality traits on privacy concerns are constrained. Future studies can adopt the perspective of situational strength and consider both personality and situational factors when examining the influence of personality traits on privacy concerns.

Another plausible explanation on the findings concerning the effects of personality traits on privacy concerns could be due to the fact that each personality has multiple facets (Korzaan and Boswell, 2008). For conscientiousness, self-discipline and deliberateness are the main facets. While a deliberate nature may increase the perception of privacy threats/costs, a self-disciplined nature may mitigate concerns at the same time; being self-disciplined means to carry out tasks “to completion despite boredom and other distractions” (Costa and McCrae, 1992, p. 18). On other words, the effects of some facets in a given personality on privacy concerns could be counterbalanced by the other facets. This is also the case in neuroticism, for which anxiety and impulsiveness are two main facets. While anxiety may raise an individual’s privacy concerns, impulsiveness may prompt individuals to feel less severe threats to privacy threats less severe (Badgaiyan and Verma, 2014; Costa and McCrae, 1992).

Regarding Big Five personality hypothesis, agreeableness was the only significant antecedent of privacy concerns. Such a result indicated that the effect of agreeableness was not constrained by the aforementioned situational strength; such was in accordance to Jensen-Campbell et al. (2010). Jensen-Campbell et al. (2010) reviewed past studies on agreeableness and indicated that more agreeable people were less likely to be influenced by social contexts. Furthermore, the effect of agreeableness was the opposite of what was hypothesized. This unexpected result could be attributed to the moderating effect of communication settings. Face-to-face interactions that feature more interactive dialogues and visual clues may cause individuals with a high level of agreeableness to be more comfortable in communicating boundaries of personal privacy than in computer-mediated interactions (Bansal et al., 2010); it is plausible that the effect of agreeableness on information privacy concerns is negative under face-to-face communication settings, however, is positive under computer-mediated communication settings; the reason for this may be the threatening nature of providing personal information on websites. Although individuals with a higher level of agreeableness tend to be more cooperative, this is not always the case. Past studies indicated that more agreeable people are more likely to be less cooperative; moreover, they have been found to have negative emotions or even to engage in retaliation within threatening social contexts (Jensen-Campbell et al., 2010). Since personal information on websites could easily be manipulated or misused, the threatening online context may make respondents less cooperative and have more privacy concerns.

Finally, the relationship between risk-taking propensity and information privacy concerns was confirmed, but in an unexpected direction. This may be because that high risk-taking people have a positive attitude toward risks (Bansal et al., 2010, p. 141). High rewards typically accompany high losses. Thus, individuals with a high risk-taking propensity are more likely willing to accept risks, not because they underestimate a privacy threat, but because they determine that potential rewards may be greater than possible losses (Skeel et al., 2007). Thus, it is reasonable that high risk-taking persons may have stronger privacy concerns and, at the same time, may believe their disclosure on personal information will yield greater benefits.
Practical implications
Managers are advised to offer attractive extrinsic rewards to elicit internet users' willingness to provide personal information. For example, users may earn reward points when they agree to share personal information with websites. These points, in turn, can be used in exchange for advanced service or gifts. Moreover, websites could develop different marketing campaigns for users with an agreeableness personality, a higher risk-taking propensity, or more experience of privacy invasion. Finally, more explanations for the benefits available and greater assurance for the mechanisms implemented to minimize the costs for collecting and storing personal information should be provided; such should be easily accessible on the websites so that the users can better evaluate the trade-offs and opt in for disclosing their personal information in exchange for a more personalized web experience.

Limitations and future research
This study had a number of limitations. First, individuals may have different concerns about various types of personal information. Acquisti and Grossklags (2005) claimed that certain profile information such as age, gender and physical characteristics was weak for identifying a given individual raised minimal concerns in users. Conversely, some profile information such as name, phone and bank account are good identifiers which lead individuals to perceive greater threats to privacy in their disclosure. Future studies are suggested to classify information in terms of identifiability. How individuals evaluate and respond to such more refined privacy concerns may be further explored. Second, we concentrated on the possible threats of privacy invasion while defining information privacy concerns. As researchers recently urged that individuals' abilities may impact their assessments of information privacy concerns and disclosure intention (Xu et al., 2011), future studies should consider privacy self-efficacy or privacy control in their analyses. Next, this study examined the effect of privacy benefits in the light of extrinsic rewards (i.e. financial incentives). As Smith et al. (2011) reported that privacy benefits may also be manifested in personalization and social benefits, future studies may need to analyze the effects of various forms of benefits which may be more detrimental. Lastly, this research used purposive sampling to collect data; this limits the external validity of the findings. In order to enhance external validity, future studies are suggested to test our research model using different samples, and contexts. Given the objectives of this research, a portion of Li's (2012) framework related to privacy calculus theory and personality theory is utilized to develop the research model. Since Li (2012) proposed an integrated theoretical framework, it is relatively difficult to include the entire framework in a single study. Future studies can develop and empirically examine research models using the other portions of Li's (2012) framework.

References
What drives willingness to provide information?


Costa, P.T. and McCrae, R. (1992), *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual*, Psychological Assessment Resources, Odessa, FL.


What drives willingness to provide information?


(The Appendix follows overleaf.)
Appendix

Agreeableness (AG) (John et al., 1991; John et al., 2008)
AG1 I am considerate and kind to almost everyone
AG2 I like to cooperate with others
AG3 I am helpful and unselfish with others
AG4 I am sometime rude to others ($R$)
AG5 I am generally trusting*

Conscientiousness (CO) (John et al., 1991; John et al., 2008)
CO1 I am easily distracted ($R$)*
CO2 I tend to be disorganized ($R$)
CO3 I do a thorough job
CO4 I persevere until the task in finished
CO5 I can be somewhat careless ($R$)

Extraversion (EX) (John et al., 1991; John et al., 2008)
EX1 I am outgoing, sociable
EX2 I have an assertive personality
EX3 I am sometimes shy, inhibited ($R$)*
EX4 I generate a lot of enthusiasm
EX5 I am full of energy

Neuroticism (NE) (John et al., 1991; John et al., 2008)
NE1 I am relaxed, handle stress well ($R$)
NE2 I remain calm in tense situations ($R$)
NE3 I am depressed, blue
NE4 I worry a lot
NE5 I am emotionally stable, not easily upset ($R$)

Openness (OPE) (John et al., 1991; John et al., 2008)
OPE1 I am original, come up with new ideas
OPE2 I am inventive
OPE3 I am curious about many different things
OPE4 I am ingenious, a deep thinker
OPE5 I like to reflect, play with ideas

Risk-taking propensity (RTP) (Weber et al., 2002)
RTP1 My tastes are different from those of my friends
RTP2 I disagree with my parents on a major issue*
RTP3 I argue with my friend about an issue on which he or she has a very different opinion
RTP4 I approach my boss to ask for a raise*
RTP5 I will tell my friend if his or her significant other had made a pass at me
RTP6 I wear provocative or unconventional clothes on occasion*
RTP7 I take a job that I enjoy over one that is prestigious but less enjoyable*
RTP8 I defend an unpopular issue that I believe in at a social occasion*

Information privacy concerns (IPC) (Smith et al., 1996)
IPC1 I am concerned that X website are collecting too much personal information about me
IPC2 X website should take more steps to make sure that the personal information in their files is accurate
IPC3 X website should not use personal information for any purpose unless it has been authorized by the individuals who provided the information
IPC4 X website should take more steps to make sure that unauthorized people cannot access personal information they collected

Table AI.
Measurement items used in this study

(continued)
Extrinsic rewards (ER) (Bock et al., 2005)

| ER1 | I will receive monetary rewards in return for providing my personal information at X website |
| ER2 | I will receive additional benefits (e.g. website functionality and points) in return for providing my personal information at X website |

Willingness to provide personal information (WPPI) (Lee and Cranage, 2011)

| WPPI1 | I will provide personal information such as name, address and phone numbers at X website |
| WPPI2 | I will provide financial information such as bank account numbers and credit card numbers at X website* |
| WPPI3 | I will provide ID numbers at X website |
| WPPI4 | I will provide personal information such as age, gender and birth date at X website |
| WPPI5 | I will provide family information such as marriage and family status at X website |
| WPPI6 | I will provide job status at X website |
| WPPI7 | I will provide financial details at X website |

**Notes:** Items marked “(R)” were scored reversely. Items marked “*” were excluded in the PLS analyses

Table AI.
Library marketing via
social media

The relationships between Facebook content
and user engagement in public libraries

Soohyung Joo
Department of Information Science, University of Kentucky,
Lexington, Kentucky, USA

Namjoo Choi
Department of Library and Information Science, University of Kentucky,
Lexington, Kentucky, USA, and

Tae Hyun Baek
Department of Integrated Strategic Communication, University of Kentucky,
Lexington, Kentucky, USA

Abstract
Purpose – The purpose of this paper is twofold: to explore what kinds of social media content public
libraries create to communicate with users online, and to examine the relationships between social media content
types and corresponding levels of user engagement.

Design/methodology/approach – The sample comprises 4,637 Facebook posts collected from 151 public
libraries across the USA. The authors identified ten types of Facebook posts based on the open coding, and
calculated the degrees of user engagement for each type of Facebook post, represented by the numbers of
likes, shares and comments. Also, The authors examined the effects of the inclusions of images or video clips
on user engagement.

Findings – The authors observed that the most frequent type of post was related to announcing
upcoming events held in libraries. This study also found that posts about community news or emotionally
inspiring messages elicited much engagement from users. Posts having an image or images tend to receive
more user engagement.

Practical implications – Based on the findings of this study, the authors discussed practical strategies for
public libraries to effectively use social media to better facilitate user engagement.

Originality/value – This study is one of a few attempts that examine the relationships between the types of
social media content and the degrees of user engagement in public library environments. Also, the authors
have proposed a coding scheme useful to analyze social media content in the context of public libraries.

Keywords Online marketing, Social media, Public library, Social media marketing, User engagement

Introduction
Social media has emerged as a popular marketing tool for public libraries; enabling them to
reach out and communicate with users online. Carr and Hayes (2015, p. 50) define social
media as “internet-based channels that allow users to opportunistically interact and
selectively self-present, either in real-time or asynchronously, with both broad and narrow
audiences who derive value from user-generated content and the perception of interaction
with others.” Social media enables users to create their own profiles and make connections
with others online, as well as allowing them to contribute and share content and
commentary on online networks (Boyd and Ellison, 2007; Steiner, 2012). Social media, which
is often equipped with various functions to share messages easily to a wider network of

This study was funded by Speed Dating For Researchers-V (SDR-V) Collaborative Grant, College of
Communication and Information, University of Kentucky.
online users, can serve as a compelling marketing tool at a low cost. Social media enables libraries to efficiently promote their services, programming and events online to their communities, and eventually to create a sense of community amongst users (Charnigo and Barnett-Ellis, 2007; Hendrix et al., 2009; De Rosa et al., 2007).

Among different platforms of social media available nowadays, Facebook has become one of the most popular social media channels for public libraries, and it is likely that a large portion of library patrons already have a Facebook account (Dryden, 2014; Tomlin, 2014). According to the recent survey by Greenwood et al. (2016), Facebook continues to be the most popular social media platform, and nearly 79 percent of internet users in the USA use Facebook. Facebook is also one of the most heavily used social media platforms in library communities (Xie and Stevenson, 2014). The strength of Facebook as a marketing tool lies in that it allows users to connect with others by simply uploading posts, and to easily respond to others’ posts through liking, sharing and commenting.

Recognizing the potential benefits of using social media in library marketing, researchers have investigated different aspects of social media use in public libraries. In particular, analyzing the content of social media has emerged as an important focus of research inquiry to offer insights that may help public librarians harness the effectiveness of social media marketing. Extant literature has identified various types of information posted on library social media based on content analysis (e.g. Chen et al., 2012, Aharony, 2012). In addition, several researchers have explored the best practices for effective deployment and use of social media in public libraries (e.g. Cahill, 2011; Steiner, 2012; Tomlin, 2014). However, little is known about the relationships between social media content and user engagement in public library environments.

Given the utility of social media as an effective platform for community engagement, the primary objectives of this study are twofold: to examine how public librarians utilize social media (i.e. Facebook pages) to facilitate user engagement and identify social media content that drives high levels of user engagement. To that end, this study analyzes a total of 4,637 Facebook posts from public libraries. We construct a coding scheme tailored to the public library context in an attempt to better understand the unique nature of Facebook content generated by libraries. Using the coding scheme, we investigate which types of Facebook content are most or least frequently generated by public libraries. More importantly, we attempt to reveal the relationships between Facebook post types and user engagement. This study is one of a few that examine the patterns of user engagement for different types of Facebook posts in public libraries. Practical implications for effective social media use are discussed based on the findings of the study.

**Literature review**

As social media has emerged as a key channel to communicate with patrons, library science researchers have paid much attention to social media use in libraries. Prior literature has investigated different aspects of social media use in public libraries, such as social media content, perceptions of stakeholders and strategies for effective social media use.

There have been several efforts that researchers analyzed types of social media content created by libraries. For example, Chen et al. (2012) analyzed five months of social media data, including Facebook and Twitter, collected from both public libraries and academic libraries. From their analysis, four types of interactions in library social media were identified: knowledge sharing, information dissemination, communication and knowledge gathering. Facebook was effective in knowledge sharing while Twitter was an efficient tool to facilitate communication. They also found that academic library patrons were more likely to engage in social media to communicate with librarians, while public library patrons used social media more for knowledge sharing. As part of her study, Aharony (2010) conducted a comparative content analysis of tweets produced by 15 public and 15 academic libraries,
and indicated that the two types of libraries exhibit differences regarding the content of their tweets. For example, public libraries mainly use the platform to promote their events with more informal language. These studies focused on the comparison between public libraries and academic libraries. Researchers also tried to analyze the content of social media messages and categorize them to understand how social media is used in library environments. Al-Daihani and AlAwadhi (2015), by analyzing the content of tweets from 17 academic libraries, presented a framework classifying the tweets into four main categories: news and announcements, library collection, library services and technology, along with several sub-categories within them. Al-Daihani and AlAwadhi’s (2015) categories served as a content analysis scheme in the analysis of Twitter messages in the academic library environment. Madge and Coserea (2014) analyzed three public library cases, and suggested that Facebook could be used as an intense marketing tool for advertisement of library services, as well as cultural and educational activities. They also emphasized the function of Facebook as a means of sharing diverse events in which people of all ages can engage. Aharony (2012) explored Facebook use in libraries, and defined its role as a tool for marketing activities. The study pointed out the limited use of advanced features and functions when libraries used Facebook; libraries used social media simply as a means to deliver news and information to their patrons, rather than as a venue for interactive, engaging discussion. AlKindi and Al-Suqri (2013) suggested a comprehensive scheme that categorizes types of social media content in libraries. They identified 11 types of content generated by public libraries on their Facebook walls, such as general news, announcements about new books, recommendations of books and other materials, announcements of workshops and programs. Shiri and Rathi (2013) explored the use of Twitter in a large public library system, focusing on the content of the tweets created by the library. They conducted a content analysis of tweets and identified in which ways Twitter was used. Their classification framework is one of the most comprehensive in the context of public library social media use. They created a scheme of categories for tweet messages, such as acknowledgment, advisory services, announcements, events, feedback seeking, informal conversation, information sharing, library operations and others. They also found that a significant number of tweets were related to information sharing, recommendations and advisory services.

Researchers have been interested in different perceptions of social media use in the context of public libraries. For example, based on the survey method, Cavanagh (2015) attempted to understand how libraries would perceive the use of social media in public libraries. The survey results reveal that potential benefits of using Twitter include closer relationships and communication between users and libraries as well as expansion of library services and audiences. Some researchers specifically focused on young adult patrons in this line of research. Phillips (2015) investigated the impact of social media on young adult library users, most of whom use social media on a daily basis. Her research surveyed the perceptions and attitudes of librarians toward social media use in libraries, and then explored the roles and responsibilities of librarians in relation to young adult patrons. She found that public librarians well perceived the importance of their professional responsibility to engage with young adults via social media. She also emphasized the role of librarians as the co-creators of social media content in collaboration with young adults rather than the sole creators of content. Kim (2015) assessed the quality of social media services in library environments based on the investigation of user perception. She proposed and empirically validated an instrument to assess library social media services by measuring user perception of different aspects of such services, such as efficiency, availability, fulfillment and privacy.

Prior literature has also contributed to practical marketing plans, strategies and guidelines for effective social media use in public libraries. In her case study of social media
use in a public library system, Cahill (2011) analyzed several successful cases of social media use, including word-of-mouth marketing, emergency information broadcast, community engagement and soliciting feedback. She further suggested best practices of social media use in public libraries, such as maintaining a friendly, informal tone, updating at least once a day, posting varied, timely, inclusive content, etc. Steiner (2012) suggested strategies for social media implementation and marketing in libraries, which comprehensively covered specific instructions for analyzing the current status of social media use in libraries, segmenting the audience, setting up the mission and vision, and best practices. Similarly, Tomlin (2014) produced guidelines for effective social media marketing for public libraries, which emphasized the importance of regular updating and monitoring. She also proposed specific ways to make Facebook posts more user-friendly, such as interactive messaging, different format of media use beyond text alone and polls. Alman and Swanson (2014) suggested that social media marketing should include specific plans about media types, content, monitoring and updating and assessment. Also, they noted that it was important to set up target users in advance and address customized messages to them. They recommended that libraries generate a wide variety of content, such as news about recent acquisitions, upcoming programs, community events, trivia, recommendations for books and others.

Previous researchers and practitioners have acknowledged the importance of user engagement in determining the relationships between service providers and users (Bowden, 2009; Brodie et al., 2013; Gummerus et al., 2012), and have conceptualized it in varying perspectives. For example, engagement is viewed as “a desirable – even essential – human response to computer-mediated activities” (Laurel, 1993, p. 112). Calder and Malthouse (2008) have viewed user engagement as the sum of motivational experiences users have with the media. According to Van Doorn et al.’s (2010) view, we can consider user engagement a behavioral manifestation that drives motivation to attach to the brand of the service provider beyond using the service. Cho et al. (2014) suggested that likes, shares and comments reflect different levels of engagement between an organization and the public on Facebook. That is, “like” is used as the lowest level of engagement to let users know that they enjoy Facebook posts without verbal expression, “share” enables the public to voluntarily deliver organizational messages to their own social media communities and “comment” is deemed to be the highest level of engagement because it requires the public to expend more time and effort in responding to organizational messages.

As shown in the literature review, prior studies have contributed to the understanding of different aspects of social media use in library environments. Particularly, researchers have tried to investigate social media content generated by public libraries and suggest strategies for effective social media use. In addition, there has been discussion on user engagement in the social media context in general. However, little research has been done to analyze the relationships between social media content and user engagement in the context of public libraries. In particular, there were few studies that conducted a quantitative analysis to examine user engagement patterns by social media content type. Several prior studies identified different types of social media content in library environments, but few of them further quantitatively looked into how user engagement patterns would differ by content type. To the best of our knowledge, this study is one of a few attempts that quantitatively analyzed the degrees of user engagement by Facebook post type in the public library context. In addition, we found that little attention has paid to sentiment or emotion analysis in public library social media research. In this study, we intend to advance our understanding of Facebook content created by public libraries based on the analysis of the relationships between content types and behavioral aspects of user engagement, including likes, shares and comments.
Research questions
This study explores types of Facebook posts and their associations with user engagement in public libraries. Three research questions guided the investigation:

RQ1. What types of content are posted on Facebook by public libraries?

RQ2. What are the relationships between post types and user engagement?

RQ3. What are the relationships between the inclusion of images/videos and user engagement?

Methodology
Sample and research variables. In this study, we collected data from Facebook “pages” of 151 public libraries. Users can subscribe to those library pages by clicking the “like” button to follow the page. For this study, 151 public libraries were randomly selected from a directory of public libraries listed in the IMLS Public Library Statistics (IMLS, 2017). We used a random number generation function and assigned a random number for each library in the list. Then, we chose top 151 libraries after sorting the list by those numbers. Highlighted posts were collected from July to October of 2015 from each selected library. Facebook shows highlighted posts selectively on timelines when viewing old posts, while other posts become hidden over time. In total, 4,637 posts were collected from 151 library Pages.

For each post, the following information was collected: library name, posted text, posted date, the number of likes, the number of comments, the number of shares, the inclusion of images, the inclusion of video clips and others. A coding scheme consisting of ten categories was developed based on the open coding method (see Table I, Coding scheme). Then, the content of each post was manually coded into ten categories using content analysis.

In this study, user engagement is represented by the number of likes, comments and shares for each post. These variables can be used as signs of success on social media sites, showing users’ engagement and contribution (Alman and Swanson, 2014). The ways users engage with libraries’ social media include showing interest in the content (likes), providing feedback (comments), or sharing the content to others (shares). The number of likes, comments and shares has been used as a measure of user engagement on Facebook (Bonzanini, 2016). To control for the unwanted size effect, those user engagement indicators were standardized, including the number of likes, the number of comments and the number of shares. Larger libraries were likely to receive a higher amount of user engagement, as they had more followers on their pages. To represent a library size, we used both the population of the legal service area (PLSA) and the number of Facebook followers as of October 2015. IMLS provides nation-wide survey data of public libraries in the USA, and it includes the “PLSA.” PLSA refers to “the number of people in the geographic area for which a public library has been established to offer services and from which the library derives revenue, plus any areas served under contract for which the library is the primary service provider (IMLS-PLS, p. 6).” PLSA has been used in the library science field as an indicator to represent the size of libraries (Joo and Cahill, in press). In addition, the number of Facebook followers was taken into consideration. On Facebook, the number of followers implies the actual number of Facebook users who receive posting updates from the library. As shown in Figure 1, these two variables exhibit a typical Zipf’s law pattern. In this study, Libsize was used to control for unwanted size effect in data analysis:

\[ \text{Lib}_{\text{size}} = \log_{10}(\text{PLSA}) \times \log_{10}(\text{Follower}) \]

Data analysis. For data analysis, we adopted a mixed method: open coding and content analysis, descriptive statistics and inferential statistics. In this study, the unit of analysis is
an individual post as we analyzed the content type of each post. That is, the frequency analysis, descriptive analysis and inferential analysis were all conducted at the individual post level.

Open coding was employed for analysis, which is the part of the data analysis process concerning identifying, naming, categorizing and describing phenomena from unstructured text (Corbin and Strauss, 1990). As reviewed in the Literature Review section, prior studies contributed to the identification of categories for social media content in libraries.
However, there are relatively a small number of category frameworks tailored to the public library context. Shiri and Rathi (2013) suggested a comprehensive list of content categories, but those categories were derived from the analysis of Twitter messages, not Facebook. Thus, we decided to come up with a new set of categories, which specifically reflects Facebook content produced by public libraries. Ten categories were defined to classify all observed posts into specific content types, using open coding, as shown in Table I. Two coders who have expertise in advertising and social media marketing were recruited to conduct content analysis. Both coders took a series of courses related to marketing, social media and strategic communication. Also, both of them had years of experience in using Facebook, and were familiar with its interface and functionalities. In this study, we suggested several categories that were not identified in prior literature. First, we differentiated the event related posts into two groups; upcoming event announcements (C1) and completed events (C2). From the initial scanning of Facebook posts, we observed that users’ response patterns would be different between C1 and C2. Second, this study attempted to identify messages that drive users’ feeling or emotion by separately coding them. In the public library context, there was little discussion of emotion related messages. In the marketing field, it is considered an effective strategy to utilize social media content that inspires emotion (Dinesh, 2017). We wanted to see how such an emotion-inspiring post would work effectively in the context of public library social media. Third, we wanted to investigate storytime or children programming related posts separately (C7). Public libraries use a large portion of their resources for programing for children (Joo and Cahill, in press). We observed that libraries used social media to promote those children related services. The C7 category was defined to specifically investigate children related posts in Facebook, which are unique content in the public library context. Fourth, we further classified community news or local news as a distinct category (C8). From the initial observations, we found that public libraries often upload and share their local news via their Facebook page, and those posts were not perfectly fit with the other categories. Thus, we ended up with a distinct category dedicated to local/community news. Our initial observations also pointed out the frequent changes to Facebook profiles or cover photos (C9). Those infrequent post types that showed less than 0.5 percent of the sample, such as thank you messages or fundraising, had to be combined together as “Other (C10).” In this way, we made a coding scheme (Table I) that is specifically designed to content analyze Facebook content in public libraries.

Some posts were coded into multiple categories. For instance, posts about storytime or children’s programming could be coded into the categories of “storytime (C7)” and “upcoming event announcement (C1).” For example, the case of “Don’t forget to join us for Story Time at 10:30 this morning!” was coded into C1 and C7 as it was about an upcoming storytime event. Also, we observed multi-category coding cases in relation to library club activities.

![Figure 1. Distributions of PLSA and Facebook followers](image-url)
For example, the post, “The Family Book Club will discuss The Lion, the Witch, and the Wardrobe on July 25 at 10 AM, and The Book Thief will be the topic of a book discussion on August 6 at 5,” was categorized into C1 and C5, which was about the meeting announcement of a library book club. For checking inter-coder reliability, two coders independently coded randomly selected 287 posts from 10 libraries. The inter-coder reliability turned out to be adequate, CR = 0.927, according to Holsti’s (1969) index.

After the coding of all observed posts, we employed descriptive statistics to investigate the proportions of different types of posts, the proportions of multimedia-embedded posts and the degrees of user engagement for posts. In particular, RQ1 was answered by analyzing the proportions of types of posts as well as the inclusion of images or videos. In addition, RQ2 was answered by calculating the degrees of user engagement by post type. Third, to respond to RQ3, we conducted t-tests to examine whether the inclusion of multimedia (images or videos) would affect the degrees of user engagement. The dependent variables are the degrees of user engagement measured by the numbers of likes, shares and comments while the independent variables are the inclusion of an image(s) or a video(s). Here are specific hypotheses defined in relation to RQ3:

H1. There is a significant mean difference of the degree of likes between the image-embedded posts and the posts with no image.

H2. There is a significant mean difference of the degree of shares between the image-embedded posts and the posts with no image.

H3. There is a significant mean difference of the degree of comments between the image-embedded posts and the posts with no image.

H4. There is a significant mean difference of the degree of likes between the video-embedded posts and the posts with no video.

H5. There is a significant mean difference of the degree of shares between the video-embedded posts and the posts with no video.

H6. There is a significant mean difference of the degree of comments between the video-embedded posts and the posts with no video.

Results
Content types of Facebook posts
First, we analyzed types of posts uploaded on public library Facebook pages. Table II shows the frequency and percentage of each content type category. Not surprisingly, many of the
posts were related to events hosted by libraries, which are C1 and C2. Posts announcing
upcoming events (C1) were most frequently observed (2,056), accounting for 35.5 percent.
Sharing past library events (C2) was ranked second (735), 12.7 percent. This indicates that
public libraries widely utilize Facebook to announce their planned events or share the past
events they completed. Public libraries also often upload emotionally inspiring posts
(C3, 11.9 percent) as a way to get attention from their Facebook followers and potential
users. Approximately 12 percent of the observed posts involved a funny image, humor, or
an inspirational message. General announcements (C4) about the library, such as open
hours, library services and library instruction, made up about 11 percent. In particular,
many libraries posted any changes of open hours or holiday closings on Facebook. About
8.5 percent were related to library clubs (C5), such as book clubs, library-supporting groups
and children’s Lego clubs. Information about books, collections, or authors (C6) turned out to
be around 7 percent, followed by storyline-related posts (C7, 4.5 percent), profile or cover
photo updates (C9, 3.8 percent) and community news (C8, 2.6 percent). Approximately
2.4 percent of the observations were categorized as C10.

Use of images and videos
We calculated the proportion of posts that included an image(s) or a video clip(s). Table III
presents the frequency of posts that contained an image or video clip. Interestingly, images
were used most often in public library Facebook posts. We found that 3,924 out of all 4,637
posts included at least one image, accounting for 84.6 percent. Video clips were rarely used,
showing only 3.2 percent.

Numbers of likes, comments and shares
Users engaged in libraries’ Facebook activities by adding likes, leaving comments, or
sharing posts. Table IV presents descriptive statistics about user engagement indicators,
including likes, comments and shares, that library posts received. On average, each post
received 10.44 likes, 0.6 comments and 1.93 shares. This reveals that patrons mostly
engaged with the posts via likes, rather than comments or shares. Approximately 82 percent
of the posts had at least one like. Relatively, the proportion of the posts that received
comments (21.5 percent) or shares (26.2 percent) were small. The patterns of likes and shares
were very similar, showing a Pearson r coefficient of 0.852 (p < 0.01). This implies that
users who liked the posts tended to share them via their own Facebook accounts. We
observed relatively lower correlations between likes and comments (r = 0.472, p < 0.01) and
between comments and shares (r = 0.345, p < 0.01).

<table>
<thead>
<tr>
<th>Table III.</th>
<th>Frequencies of posts that include an image or a video clip</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Image</td>
</tr>
<tr>
<td>Included</td>
<td>3,924</td>
</tr>
<tr>
<td>Not included</td>
<td>713</td>
</tr>
<tr>
<td>Percent</td>
<td>84.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table IV.</th>
<th>Descriptive statistics of user engagement indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Likes</td>
</tr>
<tr>
<td>Average per post</td>
<td>10.44</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>43.74</td>
</tr>
<tr>
<td>Range</td>
<td>0–1,352</td>
</tr>
<tr>
<td>Proportion of posts that received at least one engagement</td>
<td>81.7</td>
</tr>
</tbody>
</table>
The post that received the most likes and shares, 1,352 likes and 604 shares, respectively, was about the Kansas City Royals’ Major League Baseball World Series: “Taking it on the road! Fantastic game, Royals. #TaketheCrown #WorldSeries.” As shown in this case, a post about a community’s big event could draw lots of attention from patrons (Table V).

Interestingly, as shown in Figure 2, the numbers of likes, comments and shares exhibit typical Zipf’s law patterns. All three measures showed similar patterns. This reveals that a smaller number of posts tend to get higher amounts of attention, while a larger number of posts receive little attention. In particular, the long tails in the comment or share distribution graphs indicate that lots of posts do not receive any comment or share at all.

Relationships between post content types and user engagement

We investigated the relationships between post types and user engagement measured by the numbers of likes, shares and comments. Larger libraries with more Facebook followers tended to receive more user engagement. We controlled for the size effect in this analysis by using the Libsize (see the Methods section above). Therefore, the degrees of user engagement in Table VI indicate the values normalized by Libsize. Using the normalized user engagement values, we analyzed which types of posts received more or less likes, comments, and shares from users. Table VI presents the mean values of three standardized user engagement indicators by post type.

<table>
<thead>
<tr>
<th>Likes</th>
<th>Comments</th>
<th>Shares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likes</td>
<td>–</td>
<td>0.472**</td>
</tr>
<tr>
<td>Comments</td>
<td>0.472**</td>
<td>–</td>
</tr>
<tr>
<td>Shares</td>
<td>0.852**</td>
<td>0.345**</td>
</tr>
</tbody>
</table>

Notes: *p < 0.05; **p < 0.01

Table V. Correlations among likes, comments and shares

Library marketing via social media

949
C8 was ranked first in all three user engagement indicators (likes $M = 2.751$; comments $M = 0.074$; shares $M = 0.936$), and it was partially because of some extreme cases: local community events observed during the Major League Baseball postseason in 2015. For example, Kansas City–Kansas Public Library uploaded several posts related to their regional MLB franchise’s games, and those posts received extremely high attention from users. Even after controlling for the library size, those numbers were extremely high and made C8 ranked as the most user engaged post type. The standard deviation values of likes, comments and shares of C8 were comparatively higher ($M = 9.212$; $M = 0.219$; $M = 4.057$, respectively) than those of the other categories, which reveals the existence of such extreme cases. C10 ranked at 2nd for all three engagement indicators (likes $M = 1.016$; comments $M = 0.071$; shares $M = 0.248$). The C10 category included many congratulations or thank you posts, which led users to join in those happy moments. Also, in the C10 category, posts regarding library fundraising activities earned a relatively high number of likes and shares.

C3 was ranked at 3rd in likes ($M = 0.916$) and shares ($M = 0.160$), and at 4th in comments ($M = 0.160$). C5 was ranked at 3rd in comments ($M = 0.060$), 4th in shares ($M = 0.072$) and 5th in likes ($M = 0.689$). Those library club users tended to be more loyal to their library and associated clubs, so they were more likely to leave comments or share the posts about such clubs. Interestingly, C9 (profile/cover photo update) also received a relatively large number of likes, ranked at 4th ($M = 0.749$). However, it was least shared ($M = 0.025$). When looking into library event related posts (C1 and C2), C2 was likely to receive more attention than C1: C2 likes $M = 0.618$, C2 comments $M = 0.050$ and C2 shares $M = 0.041$ vs C1 likes $M = 0.360$, C1 comments $M = 0.025$ and C1 shares $M = 0.061$. It can be interpreted that users enjoyed seeing the activities or events that they or others participated in. By contrast, we observed that C1 ($M = 0.061$) was more shared than C2 ($M = 0.041$). That is, users were likely to share upcoming library programs to their Facebook friends to let them know of those interesting events. The study shows that C7 received least attention from users (likes $M = 0.326$; comments $M = 0.017$; shares $M = 0.037$) (Figure 3).

Finally, we are interested in the impact of the inclusion of images or video clips on user engagement with public libraries on Facebook (RQ3). Table VII presents how user engagement patterns would differ by the inclusion of such multimedia items. We found that the inclusion of an image in a Facebook post increased user engagement. The $t$-test results indicate that users tended to extend more likes, comments and shares to those posts. $H1$–$H3$ are accepted. There were significant mean differences in user engagement degrees between the image-embedded posts (likes $M = 0.672$; comments $M = 0.042$; shares $M = 0.116$) and the no-image posts (likes $M = 0.284$; comments $M = 0.020$; shares $M = 0.033$) in all three types of engagement indicators at the $\alpha$ of 0.01: likes $t(4241.867) = 9.169$, $p < 0.01$; comments $t(1728.382) = 6.326$, $p < 0.01$; shares $t(4634.833) = 5.245$. From this result, it can be inferred that users liked to see visually

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
<th>C9</th>
<th>C10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likes</td>
<td>Mean</td>
<td>0.360</td>
<td>0.618</td>
<td>0.916</td>
<td>0.520</td>
<td>0.689</td>
<td>0.407</td>
<td>0.326</td>
<td>2.751</td>
<td>0.749</td>
</tr>
<tr>
<td>SD</td>
<td>1.041</td>
<td>0.845</td>
<td>1.898</td>
<td>1.094</td>
<td>1.453</td>
<td>0.898</td>
<td>0.584</td>
<td>9.212</td>
<td>1.356</td>
<td>2.369</td>
</tr>
<tr>
<td>Rank</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Comments</td>
<td>Mean</td>
<td>0.025</td>
<td>0.050</td>
<td>0.056</td>
<td>0.034</td>
<td>0.060</td>
<td>0.027</td>
<td>0.017</td>
<td>0.074</td>
<td>0.036</td>
</tr>
<tr>
<td>SD</td>
<td>0.091</td>
<td>0.138</td>
<td>0.163</td>
<td>0.092</td>
<td>0.197</td>
<td>0.108</td>
<td>0.071</td>
<td>0.219</td>
<td>0.134</td>
<td>0.162</td>
</tr>
<tr>
<td>Rank</td>
<td>9</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Shares</td>
<td>Mean</td>
<td>0.061</td>
<td>0.041</td>
<td>0.160</td>
<td>0.071</td>
<td>0.072</td>
<td>0.045</td>
<td>0.037</td>
<td>0.093</td>
<td>0.028</td>
</tr>
<tr>
<td>SD</td>
<td>0.236</td>
<td>0.165</td>
<td>0.753</td>
<td>0.277</td>
<td>0.352</td>
<td>0.220</td>
<td>0.084</td>
<td>4.057</td>
<td>0.108</td>
<td>0.869</td>
</tr>
<tr>
<td>Rank</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>10</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
<th>C9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likes</td>
<td>Mean</td>
<td>0.360</td>
<td>0.618</td>
<td>0.916</td>
<td>0.520</td>
<td>0.689</td>
<td>0.407</td>
<td>0.326</td>
<td>2.751</td>
</tr>
<tr>
<td>SD</td>
<td>1.041</td>
<td>0.845</td>
<td>1.898</td>
<td>1.094</td>
<td>1.453</td>
<td>0.898</td>
<td>0.584</td>
<td>9.212</td>
<td>1.356</td>
</tr>
<tr>
<td>Rank</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Comments</td>
<td>Mean</td>
<td>0.025</td>
<td>0.050</td>
<td>0.056</td>
<td>0.034</td>
<td>0.060</td>
<td>0.027</td>
<td>0.017</td>
<td>0.074</td>
</tr>
<tr>
<td>SD</td>
<td>0.091</td>
<td>0.138</td>
<td>0.163</td>
<td>0.092</td>
<td>0.197</td>
<td>0.108</td>
<td>0.071</td>
<td>0.219</td>
<td>0.134</td>
</tr>
<tr>
<td>Rank</td>
<td>9</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Shares</td>
<td>Mean</td>
<td>0.061</td>
<td>0.041</td>
<td>0.160</td>
<td>0.071</td>
<td>0.072</td>
<td>0.045</td>
<td>0.037</td>
<td>0.093</td>
</tr>
<tr>
<td>SD</td>
<td>0.236</td>
<td>0.165</td>
<td>0.753</td>
<td>0.277</td>
<td>0.352</td>
<td>0.220</td>
<td>0.084</td>
<td>4.057</td>
<td>0.108</td>
</tr>
<tr>
<td>Rank</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>10</td>
</tr>
</tbody>
</table>
attractive posts rather than simple text posts. However, H4–H6 were not accepted by the t-test results. That is, user engagement patterns were not significantly different when video clips were included in all three types of engagement indicators.

**Discussion**

This study explored what types of Facebook posts are created by public libraries, and examined the relationships between post types and user engagement. We identified ten types of Facebook posts, and calculated the degrees of user engagement for each type of post, represented by the numbers of likes, shares and comments. Also, we investigated the effects of post features, such as the inclusion of images and video clips, on user engagement.

The findings of this study affirm that the most common type of Facebook posts are concerned with library events, including upcoming library events (35.5 percent) or past library events (12.7 percent), which make up almost half of the observed posts. This shows that social media serves as an important channel to spread library events broadly to patron communities online. In addition, we observed various types of Facebook posts, ranging from funny images, inspirational messages, general announcements, announcements about books and collections, community news, among others. Diversifying social media content can expand the role of social media in public libraries beyond distribution of library events. Libraries need to market various services, programs, events happening at the library, collections, staff expertise and more (Dryden, 2014). To cover such diverse content in
library social media, it is necessary that several staff from different departments work together to upload various stories and messages on social media in public libraries. Tomlin (2014) suggests that multiple content creators from several departments can add diversity to social media posts.

In addition, this study examined the relationships between post types and user engagement with the Facebook pages of public libraries. Not surprisingly, clicking on the like button turned out to be the most common method for users to explicitly engage with the Facebook post, as it is relatively easier than sharing the post or leaving a comment. Interestingly, posts about community news were more likely to receive likes, comments and shares from users. This implies that expanding the scope of social media content to local community issues can be an effective way for public libraries to draw attention from community residents. In particular, public libraries can take advantage of big community events or news as a boon to connect themselves to community populations. We also observed that acknowledgment messages, such as thank you notes or congratulations, and fundraising messages received much engagement. This finding implies that people are likely to involve themselves in those emotionally positive messages, as well as spread fundraising messages to others to make them more visible in the online space. In addition, the results also indicate that Facebook posts containing funny images, short jokes, or inspirational messages were likely to receive much attention from users. Social media posts that inspire emotion, such as humor or inspirational content, are considered an effective method to create a connection with viewers and to drive much engagement (Dinesh, 2017). This study empirically affirms that content inspiring emotion can also be useful to raise engagement in the context of public library social media.

When looking into event related posts, past event posts tended to get more engagement than those announcing upcoming events. We can infer that those who attended an event might want to see photos of the event they participated in and consequently to react to those posts. Past event posts could also offer a glimpse of what the event was like to those who would be potentially interested in that kind of event in the library. Thus, uploading posts about finished events can be an effective strategy for public libraries to facilitate interactions with current and potential users. In particular, pictures of a great turnout at a library’s event can indicate the role the library plays in the community (Alman and Swanson, 2014). In addition, libraries need to exert more effort to create interactive messages that lead to more user engagement when announcing events. Librarians can ask for opinions or comments about the events they plan to hold to invite users to engage with their posts (Steiner, 2012).

Posts about general announcements (C4), books/collections/authors (C6), or storytimes (C7) yielded comparatively smaller engagement from users. The C4 category posts include library open hours, introduction to general services, location information and others. These general announcements are usually one-way, factual bits of information that elicit few responses from users. The C6 category includes many posts about new arrivals. Many of those new arrival messages are routine updates on library books and collections. Posts about storytimes are limitedly targeted to parents of young children, and most of those posts are announcements of regular weekly storytimes. Although these three types of posts receive a relatively small amount of user engagement, it is important that these messages be shared with the user community. Facebook should serve as one of the main channels to announce library services and news as well as updates on books/collections. Also, storytimes are one of the key services provided by public libraries (Joo and Cahill, in press), and it is necessary to continuously promote children’s services via library social media.

The results of the study found that posts containing an image or images received more user engagement. Use of different forms of media, including pictures, can make Facebook posts more user-friendly and lead to user engagement (Tomlin, 2014). We found that public libraries already have actively used image-embedded content. The findings of this study show that almost 85 percent of the observed posts created by public libraries...
included at least one image. Those images include various types, such as a flyer, a picture taken at a library event, or an image shared from another source. Pictures of library events or activities can indicate what kinds of services and programming the library offers to their patrons and the community. In addition, it is important to regularly update the profile or cover photo. For instance, we found that a profile update or cover photo update was one of the post types that received higher numbers of likes from users. Some libraries changed their cover photos on a regular basis to reflect seasons, holidays, or special events. Those posts with new cover photos draw user attention. Also, updating cover photos can remind users that the library continues to update their Facebook pages. However, there was no significant effect of the inclusion of video clips on user engagement. In Facebook, video is presented sound off by default, and users usually do not turn on sound when seeing posts with video (Patel, 2016), which makes video materials less interactive. In addition, watching a video takes time and requires more effort than viewing an image. This might be part of the reason why there was no significant effect of video content on user engagement.

In addition, the contribution of this study lies in that it employed a quantitative approach to investigate the relationships between the content type and user engagement in public library Facebook practice. This study also analyzed a relatively large sample, i.e., 151 libraries and 4,637 posts, compared with prior studies (e.g. Chen et al., 2012; Aharony, 2010; AlKindi and Al-Suqri, 2013; Shiri and Rathi, 2013). While previous studies focused on the understanding of social media content generated by libraries, this study tried to further examine the relationships between content and user engagement quantitatively. Most of prior relevant studies worked with relatively small samples and analyzed cases qualitatively (e.g. Chen et al., 2012; Madge and Coserea, 2014; etc.). The coding data were derived from the analysis of larger Facebook post observations, and we came up with a comprehensive set of categories tailored to Facebook use in public library contexts. For example, we divided the event related posts into two types, one as upcoming events and the other completed events. Also, this study identified the categories of storytimes, local/community news, and profiles/cover photos separately. The categorization framework suggested in this study can be useful to understand the nature of Facebook use in public libraries.

Conclusion
Facebook has been widely adopted as a communication tool to interact with patrons in public libraries. We identified different types of Facebook posts created by public libraries, and further investigated the relationships between post types and user engagement. We observed that around half of Facebook posts were directly related to library events. Also, we found that posts about community news or emotionally inspiring messages drove much engagement from users. The findings of this study yield insight into effective social media use in public libraries. The contribution of this study lies in that it defined ten different categories of post types in public library environments and examined user engagement patterns for those types.

This study has several limitations to be acknowledged. First, the sample size, 4,637 posts from 151 libraries, might not be enough to represent social media use in public libraries in the entire USA. Also, the data covers only four months of a year, mostly the summer and fall seasons. There might be different patterns of content types in other seasons. Additionally, the collected data covers only the cases of the USA. Obviously, there might be different patterns of social media content and use in different regions of the world. Second, the present study did not consider library level factors such as staff and resources dedicated to social media services. Third, as we rely on descriptive statistics and basic inferential statistics, the implications might not be particularly novel. Finally, this study did not look into other social media venues, such as Twitter, YouTube and Pinterest. These limitations illustrate future research that enlarges a sample size to better represent public libraries, and completely covers
different seasons and months in a year. Also, the future study will incorporate library level factors, especially resources relevant to social media services, to have an extended understanding of social media use in public libraries. More importantly, social media use on other types of platforms, including Twitter and Pinterest, will be investigated and compared to survey similarities and differences among different platforms of social media. As a way to better understand the underlying motivations of user engagement, a survey method can supplement this line of research. The future research can adopt a user survey to explore the nature of engagement from the analysis of library patrons’ perceptions.

References


Laurel, B. (1993), Computers as Theatre, Addison-Wesley, Reading, MA.


Steiner, S.K. (2012), Strategic Planning for Social Media in Libraries, ALA TechSource, Chicago, IL.


Further reading


Corresponding author

Namjoo Choi can be contacted at: namjoo.choi@uky.edu

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com
Do corporate websites’ changes reflect firms’ survival?

Desamparados Blazquez and Josep Domenech

Department of Economics and Social Sciences, Universitat Politècnica de València, Valencia, Spain, and Ana Debón

Department of Applied Statistics and Operational Research, and Quality, Universitat Politècnica de València, Valencia, Spain

Abstract

Purpose – The purpose of this paper is to analyze to what extent changes in corporate websites reflect firms’ survival. Since keeping a website online involves some costs, it is likely that firms would invest resources on it only when they are active and healthy. Therefore, when a firm dies, this event is likely to be manifested on its website as lacking updates or being down.

Design/methodology/approach – Changes in the corporate websites of a panel of Spanish firms were tracked between 2008 and 2014 in order to evaluate the approach. The status of websites, classified according to the type of change undergone, was used to infer firms’ activity status (active or inactive). Multi-period logistic regressions and a duration model were applied to study the relationship among the website status and the firm’s status.

Findings – Results showed that changes in website contents clearly reflect the firm’s status. Active firms were mainly associated with updated corporate websites, while inactive firms were more associated with down websites. In fact, results confirmed that the firms’ death hazard increases when the website activity lowers.

Originality/value – Although online information is increasingly being used to monitor the economy, this is the first study to connect online data to firms’ survival. The results revealed a new source of information about business demography and evidenced corporate websites as a fresh source of high granularity business data.

Keywords Corporate websites, Business demography statistics, Firms’ survival, Online economic information

PAPER TYPE Research paper

Introduction

Business demography is one of the economic aspects that attracts more attention from governments and policy makers. Indeed most official statistics institutions (e.g. Eurostat, Office for National Statistics of Spain and Australian Bureau of Statistics) carry out detailed surveys to monitor the active population of firms, their birth, survival and death. The interest shown in business demography statistics relies on the important role they play in economic growth, productivity and employment (Eurostat and OECD, 2007).

In the digital era, the prominent role of the internet in economy and society, along with the development of advanced computer systems and architectures, opens up new ways of monitoring economic activities (Blazquez and Domenech, 2014; Vaughan, 2014) and, therefore, business demography. The internet and the World Wide Web (WWW) have become basic tools for the daily activities of individuals and companies, whose importance is increasing in both developed and developing countries. For consumers, the WWW is a convenient instrument to find information on products and services, and if available, to purchase them online. For companies, the WWW is an inexpensive channel to not only offer information about their products, services and activities, but to also make transactions with customers more quickly and more flexibly. In this context, companies have massively
developed their websites in order to be present on the digital channels. For instance, 75 percent of companies are present on the WWW in Spain (INE, 2016), which is the country on which our empirical analysis focuses.

Corporate websites constitute the most formal and official representation of firms on the internet. Generally, firms describe their main activities, products and intended strategies on their websites. Therefore, corporate website contents are necessarily connected to business activity to some extent, which has been recently studied from different perspectives. For instance, evidence has been found for the relation of website contents to technology adoption (Youtie et al., 2012; Arora et al., 2013), innovation activities (Arora et al., 2016; Gök et al., 2015), firms’ growth (Li et al., 2016), and firms’ export orientation (Blazquez and Domenech, 2014, 2018).

Once it has been proved that firms’ activities emerge on their corporate website, the question as to whether firms’ inactivity is also manifested on their website arises. Keeping a website online involves some costs, such as fees for domain name registrations or server hosting upkeep. Furthermore, costs increase when companies wish to keep website contents and related technologies up-to-date. Active companies are expected to regularly modify their website to include new products or services, renew its design and offer additional functionalities, or to inform potential customers about new offers or promotions. Since keeping the website updated requires firms to mobilize some resources (financial, working or both), it is plausible that only active and healthy firms would invest their resources to that end. Therefore if a company dies, this event is likely to be manifested on its website as lacking updates or simply as being down.

Most academic research conducted on firms’ survival has focused on the factors that contribute to keep companies alive. The firm’s age, size, productivity or profitability have been widely studied as determinant factors that contribute to firms’ survival. Despite the important role that the WWW plays in today’s business, there are no studies relating the WWW to firms’ survival. Corporate websites are a fresh source of business information as they are publicly accessible and provide access to high granularity (company level) data which are generally updated regularly. For these reasons, they have been used to analyze some company behaviors or strategies. Nevertheless, the approach of employing corporate websites to analyze firms’ survival is novel.

This paper analyzes to what extent changes in corporate websites reflect firms’ survival. This work hypothesizes that if a firm dies, it is very likely that its website goes down, which could happen shortly before or after the firm’s death. If this relationship exists, then the corporate website status (down, unchanged or updated), whose retrieval and tracking are inexpensive, could help monitor firms’ survival. To evaluate our proposal, the corporate website changes and firm activity status of a panel of Spanish firms were monitored for seven years and analyzed. Multi-period logistic regressions and survival analysis were run to infer the firms’ activity status. The results showed that the corporate website status clearly reflects the firm’s status.

The rest of the paper is organized as follows. Next section reviews the literature on firms’ survival and the detection of economic information through web analysis. The subsequent section describes the data used and the methodology applied for the empirical analysis. The following section describes the results, including a data overview and a comprehensive analysis of model estimations. Finally, the last section draws some concluding remarks.

Theoretical background
This section provides background on firms’ survival analysis and the detection of economic activities on the WWW. First, the related literature on firms’ survival analysis is reviewed, reporting the firm-related variables which researchers have paid more attention to. Second, a review on the detection of firms’ activities through web and online data is provided, motivating the exploration of whether firms’ inactivity could also be detected through their websites.
Firms’ survival analysis

Firms’ survival is a hot topic for researchers because of its implications for business success, economic stability and growth. However, it was not until the 1990s, promoted by the increasing economic globalization, when the academic community started to focus on analyzing firms’ survival. Firms started to face new challenges in a more complex and turbulent environment, which is the reason why determining which characteristics or actions could help them survive was more necessary than ever before.

The seminal work of Evans (1987), and other later ones like Audretsch (1991), Mata and Portugal (1994) and Geroski (1995), helped expand the field of firms’ survival, focusing on a systematic analysis of which industry-specific and firm-specific factors affected companies’ survival, and in which direction.

Regarding firm-related structural variables, firm size and age have been widely explored since they have been considered stylized facts related to firm survival (Geroski, 1995). Generally, firm size has been found to increase the likelihood of a firm’s survival, especially for new entrants (Agarwal and Audretsch, 2001; Cefis and Marsili, 2005; Geroski et al., 2010). Larger firms usually have more financial and human resources available as well as a solid structure after reaching a certain production level. These factors could help reduce the risk of mortality.

A similar pattern has been exhibited by firm age, which has been found to principally increase the likelihood of survival (Audretsch et al., 2000; Manjón-Antolín and Arauzo-Carod, 2008). Older firms have had the possibility of acquiring experience in how the market works and which strategies are more profitable for them. This could help them survive compared to newcomers. That is, the effect of experience on firms’ survival is generally positive.

Other structural variables whose relationship with firms’ survival has been consistently studied by researchers include the firm’s debt structure, its level of productivity and its level of profitability (Audretsch et al., 2000; Delmar et al., 2013; Görg and Spaliara, 2014). These variables are closely related to the firm’s level of success, stability and health, and are thus potentially influential for the likelihood of a firm to survive.

The technological intensity of the activity sector in which the company operates has also been considered in other firms’ survival studies (Esteve-Pérez and Mañez-Castillejo, 2008; Giovannetti et al., 2011). The first findings pointed out that firms had more difficulties to survive in high-technological sectors. However, an opposite pattern was found later; providing highly technological products and services requires firms to develop sophisticated skills, to focus on innovation and knowledge, and these factors are potential contributors to firms’ survival particularly within today’s complex economic frame.

More recent studies continue providing insights into how the classical firm structural factors, e.g., size, age or financial ratios, and environmental factors like financial crises, location or the specific business life cycle, contribute to increase or decrease the likelihood of firms’ survival (Basile et al., 2017; Gémar et al., 2016; Guariglia et al., 2016). The results of most reviewed literature works aim to serve mainly as guidelines in managers’ decision-making processes, who could use this information to run or promote strategies that can contribute to firms’ survival.

However, none of the studies in the literature has analyzed the relationship between firms’ survival and corporate websites. The role of corporate websites in the firms’ strategies is basic in today’s digital society, and is expected to gain importance in the future. While accounting data have been useful for predicting firm’s failure, they are not perfect measures of a firm’s operational and financial status (Astebro and Winter, 2012), so complementing them with online data could offer a better idea of a firm’s health.

For these reasons, it is relevant to confirm to what extent the corporate website status is related to the firm’s status of activity, and to explore if the information provided on the website can be used for monitoring a firm’s survival. The next subsection reviews the literature on the suitability of the WWW to reflect business activity, which motivated our study on checking whether it also reflects a firm’s inactivity, i.e., a firm’s death.
Capturing firms' economic activities through web data

Every minute of the day, thousands of individuals, companies and public organisms generate, post and share information through the internet. These online activities leave a digital footprint behind that can be tracked and, if properly processed and analyzed, can help describe their economic and social behavior.

The detection of behavioral and consumer patterns, and economic and business activities, through online data is an incipient research field whose importance is starting to increase at the same time as the adoption of the internet is expanding worldwide. This generalized expansion in internet use is affecting the way companies do business, which are being enforced by the current digital context to go online. To do so, firms generally start by implementing websites, which are the most official representation of their image and could, at the same time, be used as a commercial channel.

Indeed websites are relevant sources of online data whose potential for detecting and monitoring economic activities has remained unexplored until quite recently. Websites have a complex structure that differs from one case to another, making the process of extracting, processing and analyzing information difficult to standardize and automate to allow massive data exploitation compared to traditional databases. However, websites also present many advantages, such as: they are publicly accessible, provide fresh information and can be analyzed at any time, which traditional databases generally do not. In particular, corporate websites are attracting more attention because they are being increasingly adopted by firms, which normally use them to reflect their characteristics, products and intended strategies. Therefore, websites have become rich sources of business information. For these reasons, specific technologies and methodologies for extracting and analyzing web data are being developed (Munzert et al., 2015).

The first works about detecting economic or business information on corporate websites were published a decade ago. Following a non-automated approach, Overbeeke and Snizek (2005) captured different corporate culture dimensions by analyzing the text and images available on a set of corporate websites, while Meroño-Cerdán and Soto-Acosta (2007) found that external web content related positively to firm performance.

Firms’ corporate social responsibility and sustainability strategies, and their levels of adoption, have also been successfully detected in corporate websites contents (Gallego Álvarez et al., 2008; Tagesson et al., 2009; Tang et al., 2015). This has been done, for instance, by detecting the occurrence number of keywords related to green products (Albino et al., 2009). This measure has been extended and successfully used in other studies that have focused on novel technology industries. In their work, Libaers et al. (2010) found six types of business models for commercializing novel technology by automatically analyzing the frequency with which specific keywords were present on the corporate websites of the firms under study.

Following an automatic approach, Youtie et al. (2012) and Arora et al. (2013) applied web scraping and content analysis techniques to corporate websites, including the count of keywords, in order to track the technology adoption strategies of firms on emerging technology sectors. Innovation is another relevant topic which has been recently detected through web mining techniques. Gök et al. (2015) and Arora et al. (2016) successfully detected firms’ R&D activities by analyzing corporate websites contents. For their part, Li et al. (2016) tracked firms’ sales growth in a triple helix context.

The first attempt to generalize the automatic analysis of corporate websites to discover economic information was introduced by Domenech et al. (2012). This work presents a web data mining system architecture that manages the process of crawling and analyzing corporate websites, which was successfully tested for finding web-based indicators for firms’ size. This system was adapted by Blazquez and Domenech (2018) to detect firms’ export orientation by automatically analyzing their corporate websites since a previous manual analysis found that websites potentially reflect such business activity Blazquez and Domenech (2014).
Based on previous research, in which corporate websites were demonstrated to reflect economic information and business activities, this paper hypothesizes that detecting firms’ inactivity by analyzing the data retrieved from corporate websites is also possible.

Data and methodology
This section first describes the structure of the data used herein and how it was obtained. Second, it reviews the methodology employed, which relies on multi-period logistic regression models to detect the ability of website status to predict firm’s activity status, and a duration model to provide a deeper understanding of how the web status relates to a firm’s survival.

Data
The initial study sample included 780 companies[1] established in Spain from manufacturing, services and other sectors (NACE Rev.2[2] codes 10-95), all of which were active and had a website in 2008. The sample was retrieved by a simple random sampling design from the SABI database (Bureau van Dijk, 2010), being eligible all firms in the database that met four criteria: being active, being located in Spain, belonging to any of the mentioned activity sectors, and having a website; all of them referred to year 2008, in which this study begins. The data set consists of a panel of economic and online data for these firms for years 2008 to 2014. The economic information was retrieved from company financial records by accessing a more recent version of the SABI database in January 2016, and 2014 was the last year for which complete company economic records were found.

The online information was collected by accessing the corporate websites with the Wayback Machine of the internet archive (Kahle and Gilliat, 2016), which is a public and free repository of snapshots of about 484bn web pages. The internet archive captures and stores websites on a daily basis, allowing users to access them and track their history and evolution over time. However, there are some limitations as to its use: its inability to capture websites that prevented themselves from being explored by web crawlers by means of the robots exclusion standard (i.e. robots.txt); its limited ability to capture Flash content; the fact that it does not crawl the whole WWW, so some websites are not captured and, therefore, their evolution over time cannot be tracked; and that not every website is frequently captured, even some of them less than once a year. These limitations prevented us from tracking the evolution of some corporate websites.

For these reasons, the firms whose websites were not found in the Wayback Machine were removed from the initial sample. This gave a final sample of 720 companies to be included in the study, of which 674 survived the whole time period, while the remaining 46 died at some time. Only the years from 2010 to 2014 were included in the data analysis presented below in order to track any website changes compared to the previous year and to align the website status to the time at which financial information is available. In order to take into account the different moments of time at which company data are available, information from the financial statements was lagged two periods in the empirical analyses. That is, it is possible to know the corporate website status at time $t$, but at this time the most recent financial statements available correspond to $t-2$.

Some website captures in a particular year $t$ of the study period were not available in the Wayback Machine. This resulted in an unbalanced panel with 3,254 observations, of which 3,152 corresponded to the firms that survived to 2014, and the remaining 102 to those that died during the study period.

To account for changes in the corporate websites, the procedure followed consisted in querying the Wayback Machine with the URL of each company’s website and checking the homepage for each year studied. The observed changes were coded into the variable
Web_status, which could take five different values depending on the status or type of change experimented each year. These five levels are defined as:

1. Code 1: the website is down. This includes the websites that do not work (e.g. HTTP Error 404 Not Found) or whose domain name has expired or is for sale.
2. Code 2: the website remains unchanged. This includes the cases in which the website remains exactly the same as its previous year’s version.
3. Code 3: the website has undergone minor changes. These changes include the removal or addition of sections, options, pictures and contents.
4. Code 4: the website has undergone major changes. These changes refer to a new website design, so that it completely differs from the previous year’s version; this may imply a change in the technology used to build the website.
5. Code 5: the website has not been captured by the Wayback Machine. These cases were processed as missing data and were removed from the final sample as it was impossible to determine the website status.

The data set also included the economic variables classically related to firms’ survival according to the reviewed literature. These variables, together with the firm's status (active or inactive), were retrieved from the SABI database and complemented with the information taken from the Official Gazette of the Commercial Registry[3] to account for merges and acquisitions. The following economic variables were retrieved:

- Active_i,t: dichotomous variable that takes a value of 1 if firm i is active in year t, and 0 otherwise[4].
- Size_i,t: quantitative variable measured as the logarithm of the number of employees of firm i in year t. It is a proxy to firm size.
- Age_i,t: quantitative variable measured as the number of years since firm i was established up to year t. It is a proxy to the firm’s experience.
- Debt_i,t: quantitative variable measured as the percentage of debt of firm i in year t.
- Productivity_i,t: quantitative variable measured as the value added per employee (in millions of euros) of firm i in year t.
- Profitability_i,t: quantitative variable measured as the ratio of economic profitability of firm i in year t. This ratio, known as “Return on Assets (ROA),” is obtained from dividing the operating profit by total assets.
- High_tech_i,t: dichotomous variable that takes a value of 1 when the economic activity of firm i in year t is considered of high or medium-high technological intensity according to the Eurostat classification (Eurostat, 2014), and 0 otherwise.

Multi-period logistic regression
In a first approach, firms’ survival was studied by multi-period logistic regression models. These models are useful for examining how some independent variables are related to a dependent variable when the data used as input include individuals observed over time, which was the case of the present study, and have been applied successfully in existing firms’ survival studies (Bridges and Guariglia, 2008; Jacobson and von Schedvin, 2015).

The dependent variable in this research is whether or not the firm’s status is active (Active_i,t), which is a dichotomous variable that makes logistic regression suitable for analyzing the relation with covariates. The models used include fixed-time effects to account
Survival analysis

The relation of the firm's website status with its duration, the latter defined as the time elapsed (during the observed period) until a firm fails, was analyzed through survival models (also known as duration models (Lancaster, 1990)). These models are useful for predicting events like failures or deaths on a subject (e.g. firm, machine, system, product or patient). Specifically, time and other predictive variables are considered to estimate the hazard of failure or death during a particular time period.

In survival analysis, the hazard function \( h(t) \) is the one used for conducting regressions. In this study, the hazard function was estimated through a cloglog generalized linear model, which is the equivalent to the discrete-time version of the cox proportional hazard model (Jenkins, 1995). It has been successfully applied in previous firms' survival studies for data

\[
\theta_{i,t} = \ln \left( \frac{P(\text{Active}_{i,t} = 1)}{1-P(\text{Active}_{i,t} = 1)} \right) = \beta'X_{i,t} + \gamma_t,
\]

where \( \theta_{i,t} \) is the logit, \( P(\text{Active}_{i,t} = 1) \) is the probability of occurrence of status “1” of the dependent variable \( y_{i,t} \), \( \beta' \) the vector of regression coefficients, \( X_{i,t} \) the vector of covariates for firm \( i \) in year \( t \) and \( \gamma_t \), the time specific parameters that reflect the unobservable events that affect all firms each year.

This model is used to first assess the relation between the WWW and firms' status, as it estimates the probability of a firm being active given its website status in a first specification, and given this website status and a number of economic variables in a second specification. Both model specifications controlled for the economic juncture or period effect by including dummies for each year considered in the study. Accordingly, the first model was defined as follows:

\[
\theta_{i,t} = \ln \left( \frac{P(\text{Active}_{i,t} = 1)}{1-P(\text{Active}_{i,t} = 1)} \right) = \beta_0 + \alpha \text{Web}_\text{status}_{i,t} + \gamma_t,
\]

where \( P(\text{Active}_{i,t} = 1) \) is the probability that firm \( i \) is active in year \( t \), and the logit, \( \theta_{i,t} \) is regressed on the explanatory variable \( \text{Web}_\text{status}_{i,t} \) and the fixed-effect of time, captured by \( \gamma_t \).

An extended model was specified by including also the firms' economic variables that can affect firm survival according to the literature. The variables that were finally selected were those that varied with an admissible level of significance (\( p < 0.05 \)) between both groups of firms and did not highly correlate. This second specification was defined as follows:

\[
\theta_{i,t} = \ln \left( \frac{P(\text{Active}_{i,t} = 1)}{1-P(\text{Active}_{i,t} = 1)} \right) = \beta_0 + \alpha \text{Web}_\text{status}_{i,t} + \beta'Z_{i,t-2} + \rho \text{High}_\text{tech}_{i,t-2} + \gamma_t,
\]

where \( P(\text{Active}_{i,t} = 1) \) is the probability that firm \( i \) is active in year \( t \), and the logit, \( \theta_{i,t} \) is regressed on the variable based on corporate website, \( \text{Web}_\text{status}_{i,t} \), the vector of economic quantitative variables \( Z_{i,t-2} \) which includes \( \text{Size}_{i,t-2} \), \( \text{Debt}_{i,t-2} \), \( \text{Productivity}_{i,t-2} \) and \( \text{Profitability}_{i,t-2} \), the economic categorical variable \( \text{High}_\text{tech}_{i,t-2} \) and the fixed-effect of time, captured by \( \gamma_t \).

Having confirmed the relation between the WWW and firms' status with both regressions, a duration model is applied to estimate the firm's probability of surviving one time period or more given the corporate website status.

Survival analysis

The relation of the firm's website status with its duration, the latter defined as the time elapsed (during the observed period) until a firm fails, was analyzed through survival models (also known as duration models (Lancaster, 1990)). These models are useful for predicting events like failures or deaths on a subject (e.g. firm, machine, system, product or patient). Specifically, time and other predictive variables are considered to estimate the hazard of failure or death during a particular time period.

In survival analysis, the hazard function \( h(t) \) is the one used for conducting regressions. In this study, the hazard function was estimated through a cloglog generalized linear model, which is the equivalent to the discrete-time version of the cox proportional hazard model (Jenkins, 1995). It has been successfully applied in previous firms' survival studies for data
collected on an annual basis (Tsoukas, 2011; Görg and Spaliara, 2014; Guariglia et al., 2016), which is this case. The proportional hazard model assumes that the hazard rate depends only on the time at risk, \( h_0(t) \) (the baseline hazard) and on the vector of explanatory variables, \( X \). This is the rate at which firms die in year \( t \), provided they survived the previous year, \( t-1 \). It is expressed as:

\[
h(t, X) = h_0(t) \exp(\beta'X).
\]

(4)

Particularly, the discrete-time hazard function (with period-specific effects) takes the following specification:

\[
h(t, X) = \frac{1}{C_0(\exp(b_0X) + g_t)}
\]

(5)

where \( b_0 \) is the regression coefficient vector that describes how the hazard varies in response to explanatory vector \( X \) of the covariates, and \( g_t \) captures the period-specific effects on the hazard.

For this study, this duration model was specified as follows:

\[
h(t, Web \_status) = \frac{1}{C_0(\exp(b_0 + a_{Web \_status}) + g_t)}
\]

(6)

where \( h(t, Web \_status) \) is the hazard rate; that is, the rate at which firms become inactive at time \( t \) provided they were active in year \( t-1 \), which is modeled through the explanatory variable \( Web \_status \) and the period-specific effect, \( g_t \).

Results

This section first shows some descriptive statistics and group comparisons to provide a data overview. Second, two multi-period logistic regressions are built and compared to evaluate to what extent the \( Web \_status \) variable captures the company’s activity status. Finally, these results are complemented with a duration model.

Descriptive statistics and group comparisons

The descriptive statistics of the whole data set are reported in Table I. As we can see, the sample is dominated by active firms (96.9 percent of the sample) which operate in a low-technology sector (81 percent) and that have a moderate level of debt (61 percent). This table evidences the absence of high correlations among variables, which means that there was no high risk of information redundancy and multicollinearity when estimating the regression models.

The first column of Table II summarizes the behavior of corporate websites by showing the distribution of the \( Web \_status \) variable across the sample. It indicates that most websites remained unchanged (37.7 percent) or underwent a moderate change (36.4 percent).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Active</td>
<td>0.969</td>
<td>0.174</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Web _status</td>
<td>2.618</td>
<td>0.868</td>
<td>0.272</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Size</td>
<td>3.855</td>
<td>1.241</td>
<td>0.046</td>
<td>0.149</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Age</td>
<td>24.823</td>
<td>14.318</td>
<td>0.007</td>
<td>0.022</td>
<td>0.258</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Debt</td>
<td>60.470</td>
<td>31.608</td>
<td>−0.127</td>
<td>−0.009</td>
<td>0.016</td>
<td>−0.140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Productivity</td>
<td>0.270</td>
<td>0.071</td>
<td>0.007</td>
<td>0.021</td>
<td>−0.129</td>
<td>−0.042</td>
<td>0.046</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Profitability</td>
<td>1.099</td>
<td>22.526</td>
<td>0.112</td>
<td>0.046</td>
<td>−0.005</td>
<td>0.005</td>
<td>−0.291</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td>8. High _tech</td>
<td>0.199</td>
<td>0.399</td>
<td>0.076</td>
<td>0.143</td>
<td>0.207</td>
<td>−0.062</td>
<td>−0.048</td>
<td>−0.021</td>
<td>−0.005</td>
</tr>
</tbody>
</table>

Notes: Procedures employed: Pearson’s \( r \) coefficient for pairs of continuous variables; Point-biserial coefficient for pairs of a continuous and a binary variable; \( \phi \) coefficient for pairs of binary variables; and \( \eta \) for pairs of a continuous and a categorical variable with more than two levels (Cohen et al., 2002).

Table I. Global descriptive statistics and correlation matrix

Websites and firm survival
compared with the previous year. Only 8.8 percent of the observations presented a down website, while 17.1 percent of them had totally changed. To illustrate the association with the other variables, the numeric value of Web status is also included in Table I.

In order to test whether the variables behaved differently depending on the firm’s status (active or inactive), statistical techniques of group differences were employed. Pearson’s $\chi^2$ test was applied to the categorical variables, whose results are reported in Table II. Statistically significant differences were found for the technological intensity, being active firms more associated with technology-intensive sectors than those that became inactive (20.5 vs 2.9 percent). For the website status, statistically significant differences were found between active and inactive firms. For the latter, most websites were down (50 percent) or remained unchanged (40.2 percent), while minor or major changes were found only in the remaining 9.8 percent.

In contrast, content changes were found in more than half of the active firms’ websites, mainly minor changes (37.3 percent). This was expected because website design forms part of the firm’s corporate image, which is not renewed yearly by most companies. Instead, minor changes to keep information up-to-date are frequently made by active firms. Moreover, down websites are not common among active companies (7.5 percent). The presence of unchanged websites (37.6 percent) was similar to the case of inactive firms, so this website status is not as indicative of the firm’s status as the cases in which changes were made.

With the quantitative variables, normality and homogeneity of variance were checked both graphically and numerically. As none of the variables fulfilled both assumptions, the nonparametric Mann–Whitney $U$-test was employed, which is based on the median (Anderson et al., 2014). These results are reflected in Table III. Most economic variables showed statistically different values for the active firms compared to the firms that had died during the observed period. The log number of employees was statistically higher for the active firms (3.761 vs 3.401), so firm size relates to the firm’s duration to some extent.

### Table II.
Descriptive statistics of qualitative variables and group comparisons

<table>
<thead>
<tr>
<th>Web_status(1)</th>
<th>All (%)</th>
<th>Active(0) (%)</th>
<th>Active(1) (%)</th>
<th>$\chi^2 p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 3,254)</td>
<td>(n = 102)</td>
<td>(n = 3,152)</td>
<td></td>
</tr>
<tr>
<td>Active(0)</td>
<td>3.1</td>
<td>96.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High_tech(0)</td>
<td>80.1</td>
<td>2.9</td>
<td>20.5</td>
<td>0.000</td>
</tr>
<tr>
<td>High_tech(1)</td>
<td>19.9</td>
<td>97.1</td>
<td>79.5</td>
<td></td>
</tr>
<tr>
<td>Web_status(1)</td>
<td>8.8</td>
<td>50.0</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Web_status(2)</td>
<td>37.7</td>
<td>40.2</td>
<td>37.6</td>
<td></td>
</tr>
<tr>
<td>Web_status(3)</td>
<td>36.4</td>
<td>7.8</td>
<td>37.3</td>
<td></td>
</tr>
<tr>
<td>Web_status(4)</td>
<td>17.1</td>
<td>2.0</td>
<td>17.6</td>
<td>0.000</td>
</tr>
</tbody>
</table>


### Table III.
Descriptive statistics of quantitative variables and group comparisons

<table>
<thead>
<tr>
<th></th>
<th>Active(0)</th>
<th>Active(1)</th>
<th>Mann–Whitney $U p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 102)</td>
<td>(n = 3,152)</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>3.761</td>
<td>3.401</td>
<td>0.007</td>
</tr>
<tr>
<td>Age</td>
<td>21.501</td>
<td>22.815</td>
<td>0.875</td>
</tr>
<tr>
<td>Debt</td>
<td>88.990</td>
<td>60.810</td>
<td>0.000</td>
</tr>
<tr>
<td>Productivity</td>
<td>30.947</td>
<td>48.787</td>
<td>0.000</td>
</tr>
<tr>
<td>Profitability</td>
<td>3.775</td>
<td>2.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
About the firm’s age, no statistically significant differences were found, so firms seem to die with the same probability regardless of their age. The debt value was much higher for inactive firms (88.99 vs 60.81 percent), which is indicative of the detrimental effect that high levels of debt have on firms’ health, and thus on their continuity. Active firms were associated with statistically higher levels of productivity and profitability than inactive ones. High productivity levels are connected to overall better firm performance, which would contribute to having a higher profitability. Both these measures are related to the company’s health so as expected, healthier companies continue with their activities more frequently.

**Multi-period logistic regression models**

In this section, we shed light on the role played by corporate websites status on firms’ probability of being active. First, a multi-period logistic regression model based on the Web_status variable was built, as specified in Equation (2).

Table IV provides the estimation results for this model, including the estimated regression coefficients ($\beta$), odds ratios (OR), standard errors (SE), z-values and $p$-values. The OR is a measure of the association between the different website statuses and the firm’s status, and is calculated as the exponent of the coefficients. Thus, an OR over 1 indicates that the probability that a firm is active increases with a given independent variable (in this case, each particular website status). If it is lower than 1, it indicates that this probability decreases, while if it equals 1 then there is no association between the independent and dependent variable.

For this web-based model, the results show that the observed web statuses have a statistically significant effect on the probability of a company being active. As website activity increases, the probability of a firm being active also increases. The estimate that corresponds to the website status “Unchanged” (Code 2) is positive, which means that having a working website, even if its contents or look are not changed compared to the previous year, increases the probability of a firm being active with respect to having a down website (Code 1, which was taken as the baseline level). Indeed, the probability of a firm with an unchanged website being active is 5 times (or 409.4 percent higher) than that of a firm whose website is down, as indicated by the OR.

Updating websites to a minor (Code 3) or major (Code 4) extent increases the probability of a firm being active, as expected. Furthermore, the increase found is dramatically high in both cases. The probability of a firm being active when it moderately changes its website is 30-fold higher compared to a firm whose website is down, while it is more than 50-fold higher when a website has been completely renewed. These results are in line with what was hypothesized: healthy firms invest more in maintaining and updating their websites. Hence, the more activity they evidence on their website, the more likely they are active.

<table>
<thead>
<tr>
<th>Variables</th>
<th>$\beta$</th>
<th>OR</th>
<th>SE</th>
<th>z-value</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.483</td>
<td>32.557</td>
<td>0.598</td>
<td>5.827</td>
<td>0.000</td>
</tr>
<tr>
<td>Web_status(2)</td>
<td>1.628</td>
<td>5.094</td>
<td>0.227</td>
<td>7.158</td>
<td>0.000</td>
</tr>
<tr>
<td>Web_status(3)</td>
<td>3.423</td>
<td>30.661</td>
<td>0.390</td>
<td>8.786</td>
<td>0.000</td>
</tr>
<tr>
<td>Web_status(4)</td>
<td>3.970</td>
<td>52.985</td>
<td>0.727</td>
<td>5.460</td>
<td>0.000</td>
</tr>
<tr>
<td>Observations</td>
<td>3,254</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>−349.349</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Web_status(2): Unchanged; Web_status(3): Minor change; Web_status(4): Major change. Time dummies were included. Dependent variable: Active
It is noteworthy that this does not mean that updating websites help firms remain active, but it strongly reflects firm’s active status.

Once the relationship between the corporate websites’ status and the firms’ status was evidenced, the extended specification given by Equation (3) was estimated. It included the website status variable and the structural variables selected for their potential relation to firms’ survival, and for the significant variation across active and inactive firms.

As reported in Table V, the effect of each website status on the probability of being active remains positive, and statistically significant and high. Regarding the economic variables, only the firm’s debt structure shows a statistically significant effect. Its negative coefficient indicates that as the amount of a firm’s debt increases, its probability of being active decreases. Indeed, it decreases by 1.8 percent for each percentage point increase in debt.

Although the remaining economic variables showed differences in the univariate level, they do not help explain the firm’s status at the multivariate level. On the one hand, such economic variables are related to a firm’s status, but only to a limited extent as a large number of other factors, such as the firm’s strategic decisions or specific market situations, can contribute to the death of firms with a wide range of characteristics (small or large, more or less productive, from any activity sector, etc.). On the other hand, website status has been revealed to be a clear indicator of a firm’s status. So these economic variables were unable to complement the information displayed on the web.

Given that the relationship between corporate website status and firms’ status was demonstrated, the next section went a step further to complement this analysis and to confer the study a different point of view. To do so, a survival analysis was conducted.

**Survival analysis**

This section describes the survival analysis conducted for modeling the hazard of a firm’s death at certain times depending on website status. Since the data in this study were collected on an annual basis, a time-discrete duration model was built, as specified in Equation (6).

Table VI offers the estimation results for this model, including the estimated regression coefficients (β), hazard ratios (HR), SE, z-values and p-values. The HR, calculated as the exponent of coefficients, are a measure of how often an event happens in one group compared to how often it happens in another group over time. In this case, they measure how often the different website statuses happen in the groups of active and inactive firms. Thus an HR above 1 indicates that the hazard of death increases with the corresponding website status. If it is lower than 1, it means that this hazard decreases, while if it equals 1 then there is no difference in survival between the two groups being compared.

<table>
<thead>
<tr>
<th>Variables</th>
<th>β</th>
<th>OR</th>
<th>SE</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.917</td>
<td>50.249</td>
<td>0.974</td>
<td>4.023</td>
<td>0.000</td>
</tr>
<tr>
<td>Web_status(2)</td>
<td>1.579</td>
<td>4.850</td>
<td>0.285</td>
<td>4.001</td>
<td>0.000</td>
</tr>
<tr>
<td>Web_status(3)</td>
<td>2.490</td>
<td>12.061</td>
<td>0.522</td>
<td>4.768</td>
<td>0.000</td>
</tr>
<tr>
<td>Web_status(4)</td>
<td>3.242</td>
<td>25.585</td>
<td>0.951</td>
<td>3.410</td>
<td>0.001</td>
</tr>
<tr>
<td>Size</td>
<td>0.206</td>
<td>1.229</td>
<td>0.169</td>
<td>1.218</td>
<td>0.223</td>
</tr>
<tr>
<td>Debt</td>
<td>−0.018</td>
<td>0.982</td>
<td>0.006</td>
<td>−3.053</td>
<td>0.002</td>
</tr>
<tr>
<td>Productivity</td>
<td>0.034</td>
<td>1.034</td>
<td>1.034</td>
<td>0.283</td>
<td>0.777</td>
</tr>
<tr>
<td>Profitability</td>
<td>0.005</td>
<td>1.005</td>
<td>0.011</td>
<td>0.443</td>
<td>0.658</td>
</tr>
<tr>
<td>High_tech</td>
<td>10.811</td>
<td>49,563.01</td>
<td>960.4</td>
<td>0.019</td>
<td>0.985</td>
</tr>
</tbody>
</table>

Table V. Multi-period logistic regression with web and structural variables

**Notes:** Web_status(2): Unchanged; Web_status(3): Minor change; Web_status(4): Major change. Time dummies were included. Dependent variable: Active
The negative and statistically significant coefficient estimates indicate that the firms whose websites are unchanged (Code 2), or undergo minor (Code 3) or major changes (Code 4) compared to the previous year, are exposed to a significantly lower hazard than the firms whose corporate websites are down (Code 1, which is the baseline web status). Specifically, the HR for the “Unchanged” website status (Code 2) indicates that the firms whose website contents are the same as the previous year have 0.301 times the hazard of death of the firms whose website is down; that is, their death hazard is 69.9 percent lower. The death hazard for firms which made minor changes in their websites (Code 3) is 91.6 percent lower than for those with down websites, while the percentage reaches 93.7 percent when the changes made are major (Code 4). As we can observe, the death hazard lowers at the same time as website activity increases. These results are consistent with those of the multi-period logistic regressions, and confirm the strong relationship between corporate website status and a firm’s survival.

Conclusions

Business demography is a major area of interest for researchers and policy makers because the creation and failure of companies have a huge impact on the production and employment in all the economies. In the current context, in which digital communications and internet contents reflect society’s main behavior, a new challenge arises: that of relating business demography with the WWW evolution.

This paper analyzed and confirmed the connection of a company’s activity status to the corporate website’s activity status. This was done by tracking corporate websites and statuses of firms for seven years, and then analyzing their relation with logistic regressions and a survival model. Logistic regression estimates that major changes in the corporate website increase the odds of a firm being active by more than 50-fold compared to a down website. In terms of survival, corporate website changes are related to a predicted death hazard more than 90 percent lower. Since both methods gave similar results, this means that the corporate website is a robust indicator of a firm’s activity status.

These results open up new possibilities to monitor business demography. Web data capture a firm’s status, while access to corporate websites is open and inexpensive. This means that it is possible to build online indicators to nowcast and monitor business death rates. Unlike traditional official statistics methods, which rely on surveys done on a population sample that take time to be processed, monitoring the WWW could fast reach the entire population of companies with a website. This can be done in a very short period thanks to the digital nature of the WWW, which allows firms’ information to be automatically retrieved and analyzed. This, in turn, allows policy makers and other consumers of official statistics to obtain short-term estimates of the business demography, which would eventually turn into more informed decisions.

<table>
<thead>
<tr>
<th>Variables</th>
<th>β</th>
<th>HR</th>
<th>SE</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−4.942</td>
<td>0.007</td>
<td>1.029</td>
<td>−4.802</td>
<td>0.000</td>
</tr>
<tr>
<td>Web_status(2)</td>
<td>−1.202</td>
<td>0.301</td>
<td>0.348</td>
<td>−3.454</td>
<td>0.001</td>
</tr>
<tr>
<td>Web_status(3)</td>
<td>−2.480</td>
<td>0.084</td>
<td>0.484</td>
<td>−5.128</td>
<td>0.000</td>
</tr>
<tr>
<td>Web_status(4)</td>
<td>−2.764</td>
<td>0.063</td>
<td>0.753</td>
<td>−3.668</td>
<td>0.000</td>
</tr>
<tr>
<td>Observations</td>
<td>3,194</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>−195.262</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Web_status(2): Unchanged; Web_status(3): Minor change; Web_status(4): Major change. Time dummies were included. Dependent variable: 1−Active

Table VI. Discrete-time duration model
Among the limitations of this study, first, it is worth noting that only the homepage of the website was analyzed; that is, no changes in inner sections were taken into account. Second, the sample only includes companies based in Spain, so generalizing the results to different countries must be done cautiously. Finally, we must emphasize that we describe how website status correlates with a company’s activity status, without causal analysis. Although this is useful for monitoring purposes, our results do not indicate that managers should continuously change corporate websites to increase company survival.

Notes
1. From the total sample of 780 firms, 92 percent were small and medium-sized (SMEs), in line with the prevailing productive structure in Spain (DGIPYME, 2017).
3. BORME, from their initials in Spanish, “Boletín Oficial del Registro Mercantil.”
4. We considered inactive the following firm status: in extinction; in dissolution; in liquidation; in a finished receivership where dissolution or liquidation has been ordered, but is not yet done; or in receivership in progress (if it is the most recent status and no additional information is available), except when the firm had been taken over or had merged (Eurostat and OECD, 2007).

References


Eurostat (2014), High-tech Aggregation by NACE Rev. 2. Eurostat Indicators of High-tech Industry and Knowledge-intensive Services, Office for Official Publications of the European Communities.


**Corresponding author**
Josep Domenech can be contacted at: jdomenech@upvnet.upv.es

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: permissions@emeraldinsight.com
A comparative study of disaster management information systems

Muhammad Masood Rafi
Department of Earthquake Engineering, NED University of Engineering and Technology, Karachi, Pakistan, and
Tariq Aziz and Sarosh Hashmat Lodi
Department of Civil Engineering, NED University of Engineering and Technology, Karachi, Pakistan

Abstract

Purpose – Disaster management information systems (DMISs) have been proposed in different parts of the world for effective response to a disaster. The purpose of this paper is to compare design approaches of these DMISs; examine similarities in the design of databases and communication infrastructure; and draw conclusions. Based on the examination of the studies, future opportunities have been identified and discussed.

Design/methodology/approach – The studies in the available literature on the designs of automated DMISs have been reviewed in the presented paper to identify similarities in design premise, conceptual design and design considerations.

Findings – The examination of the available studies indicates that the research on DMIS has increased significantly in different countries of the world since 2004. Data of baseline information and available resources are required by most of the presented studies, as these data are necessary for effective response to a disaster. The communication infrastructures suggested include local area network, wide area network and satellite communication for better coordination between the responders and different relief agencies at different locations. The connectivity to these networks is possible through Ethernet, Wi-Fi, general packet radio service or satellite.

Research limitations/implications – Although the research on DMIS has increased significantly over the last one decade, the studies are still few in numbers. Similarly, only few of the proposed systems have been developed and tested during a real disaster.

Practical implications – The presented review of available studies provides a holistic view of the proposed DMISs which could be useful to the disaster management authorities.

Originality/value – The paper provides valuable information on the differences in the proposed DMISs. This can help in identifying the gaps for future improvements for increased effectiveness of a DMIS. The future opportunities have also been identified in the presented paper and are discussed.

Keywords Communication, Disaster, Mitigation, Information system, Response, Hurricane

Introduction

Human response to natural hazards has been a subject of intense investigation and study. These hazards have proved to be the most difficult enemy of mankind as they are able to cause destruction on a large scale close to human settlements. The events of natural disasters may be identified by excessive magnitude, frequency or duration (Arey and Baumann, 1971; Bolt et al., 1975). Human activities, in some cases, also influence the frequency and severity of natural hazards (Shield, 2004). Natural hazards are of either geological or atmospheric origin. These include hurricanes, volcanoes, floods, tornados, typhoons, famine, fires, landslides and earthquakes. Some of these hazards may be interconnected with each other. For example, earthquakes may cause landslides.

The study of human history indicates that the ability of natural hazards to cause destruction is partly due to lack of preparedness of human beings to mitigate the effects of these hazards. A hazard needs not to become a disaster. Disaster preparedness and mitigation prevent a hazard from turning into disaster which can reduce the efforts for reconstruction and rehabilitation of the affectees. Nevertheless, in many cases, this is not possible to avoid disasters completely. As a result, disasters happen and cause loss of life.
Disaster management (DM) deals with both the mitigation and consequences of disasters. The former refers to pre-disaster preparations whereas the latter is related to post-disaster response. Aspects of planning to deal with a hazard and to avoid associated risks (Haddow and Bullock, 2004) are parts of pre-disaster mitigation. Post-disaster response is required to assure prompt and appropriate assistance to the victims of disaster, and to achieve rapid and effective recovery from a disaster.

Post-DM operations are complex (Bigley and Roberts, 2001) and are information-intensive (Davenport and Prusak, 1998; Meissner et al., 2002; De Bruijn, 2006) as these involve different stakeholders (Comfort and Kapucu, 2006). An effective response to a disaster requires not only a large variety of information for the decision makers but also its rapid flow and better coordination of activities. Good quality information improves the effectiveness of DM operations and avoids dangers to the responders and victims (Helsloot, 2005; Fisher and Kingma, 2001; Turoff, Chumer, Hiltz, Klashner, Alles, Vasarhelyi, Kogan, 2004; Turoff, Chumer Walle and Yao, 2004; Hassan et al., 2011). The coordination is needed both within the DM agency and between different agencies at several hierarchical levels (Auf der Heide, 1989). Since these operations involve severe time pressure and high uncertainty (Argote, 1982; Smith and Hayne, 1997; Ganeshkumar and Ramesh, 2010), improved coordination between relief agencies and workers can help in optimising the resources to carry out relief activities efficiently.

The use of computing technology has increased in DM similar to other fields (Borkulo et al., 2005; Dorasamy et al., 2011; Dorasamy and Raman, 2011; McEntire, 2007; Raman et al., 2011; Turoff, Chumer, Hiltz, Klashner, Alles, Vasarhelyi, Kogan, 2004; Turoff, Chumer Walle and Yao, 2004; Wattegama, 2007; Ariyabandu, 2009; Tad and Janardhanan, 2014). The availability of advanced hardware and software capabilities allows a swift response to any emergency situation in reasonable time and cost. The use of an integrated system to connect relief agencies together and to provide them ways to exchange and process information in real-time can enhance effectiveness of emergency response. The advancement in technology has made it possible to achieve both these objectives by transferring its benefits in the area of DM. As a result, the need of development and use of automated systems to manage disasters have attracted the attention of researchers all over the world (Stephenson and Anderson, 1997). Efforts have been carried out to provide reliable and coordinated response to a disaster using support networks and physical facilities which could be kept functional in a disaster situation with the help of modern technology (Kunreuther and Lerner-Lam, 2002; Mork, 2002).

DM refers to the process of acquisition, management and utilisation of disaster information in order to carry out disaster relief operations (Zhang et al., 2002). The modern technologies provide an effective tool for the development of a DM information system (DMIS). A DMIS is a computer database that enables the responders to share and use real-time information during a disaster (Atteih et al., 2010; Lee et al., 2012; Murphy and Jennex, 2006). Considering the advantages that an automated DMIS can offer, these systems have been proposed in different countries of the world. Data integration, data mining and multi-criteria decision making are the essential components of an automated DMIS (Peng et al., 2011). The data integration provides a link between the modules of data sources and data analysis. The data mining assists the users in the management of pre- and post-disaster information. The multi-criteria decision making is used to dispatch emergency resources and to evaluate effective alternative solutions. This paper reviews the literature on the strategies to design these components. The approach used by different researchers to design these aforementioned components, and the system development and management schemes have been compared. Based on this comparison, gaps in the proposals of DMISs...
have been identified and future opportunities for improved effectiveness of DMISs in DM operations have been discussed. Although the examined studies differ by region, the technological advancements are universal which makes the comparison possible. Valuable information can be obtained where differences are found.

Strategies for disaster risk reduction (DRR)
DRR refers to the identification and assessment of risk of life losses and property damages and the efforts to protect people, infrastructure and economic and social resources from the effects of disasters. The notion encompasses all activities related to DM, preparedness and mitigation. DRR provides a pathway to sustainable development by increasing resilience of individuals and communities. The factors responsible to changing a hazard into disaster include poor built environment, poor socio-economic conditions of a society and governance issues. As a result, DRR demands a multi-prong approach which requires involvement of different stakeholders and sections of a society such as individuals, communities, government, non-governmental organisations (NGOs), civil society and private sector. In view of the disastrous consequences to the sustainable development due to a disaster, DRR has been made a policy priority for the governments all over the world and a number of international policies and frameworks have been developed. These include Yokohama Strategy and Plan of Action (1999–2000), Hyogo Framework of Action: Building the Resilience of Nations and Communities (2005–2015) and African Regional Disaster Risk Reduction Framework and its Plan of Action.

The role of government is crucial in providing and/or implementing effective institutional framework for DRR to ensure sustainable development. The framework provides a basis for the development of policies for DRR programme which improve the effectiveness of DRR agenda with the help of the aforementioned stakeholders and multi-disciplinary approaches. The overall DRR strategy is based on vulnerability and risk assessment for different disasters and capacity building to mitigate these disasters. Carter (1991) identified the following means for disaster mitigation: provision of building codes, land-use planning, government grants and incentives for disaster resistant construction, training and education, public awareness and strengthening of social structure. In addition to these, an early warning system and considerations of gender behaviour in the DRR agenda (Ariyabandu, 2009) are the effective tools of risk reduction.

History of DMIS
The research on the development of automated information systems in different fields started during the 1950s (Lewin, 1958). The computing technology has also been in use for DM applications for some time. For example, Office of Emergency Preparedness (OEP) in USA developed Emergency Management Information System and Reference Index (EMISARI) during the 1980s to deal with the natural and man-made disasters (Macon and McKendree, 1974; Macon et al., 1975; McKendree, 1977, 1978). The design of EMISARI was based on a group communication process (Hiltz and Turoff, 1978/1994, 1993; Ruben, 1992; Turoff, 1991). It followed the concepts of Delphi method (Linstone and Turoff, 1975) to integrate people and data into a single database which could be dynamically changed by non-technical administrators (Turoff, Chumer, Hiltz, Klashner, Alles, Vasarhelyi, Kogan, 2004; Turoff, Chumer Walle and Yao, 2004). EMISARI allowed up to 300 users throughout USA to coordinate during an emergency and to utilise the available data urgently (Turoff, Chumer, Hiltz, Klashner, Alles, Vasarhelyi, Kogan, 2004; Turoff, Chumer Walle and Yao, 2004). Nevertheless, this was a static information system (Hiltz and Turoff, 1978/1994) in that it was unable to allow management of resources in real time during a disaster.

The need of an integrated communication system was realised long time ago (Smith and Hayne, 1997). A technical cooperation project in Latin America and the Caribbean entitled A comparative study of DMISs
SUMA (PAHO, 1998) was started in 1992 to increase the capacities in these countries to effectively manage information of the incoming relief goods. A computer software entitled Disaster Inventory System (DesInventar) (Corporación, 1994) was developed to connect nine countries of Latin America together for building databases of loss and damage, and the effects of disasters (using existing data, newspaper and institutional reports) in these countries.

Lakovou and Douligeris (2001) proposed information management system for hurricane disasters (IMASH) which was implemented using Oyster (Collins et al., 1997). IMASH was similar to oil spill information management system proposed by Douligeris et al. (1995). Nevertheless, IMASH was designed specifically for hurricane DM, contingency planning and preparedness and post-disaster restoration activities.

Based on the experiences learnt from the use of EMISARI and the response failure events as reported in the literature, Turoff, Chumer, Hiltz, Klashner, Alles, Vasarhelyi, Kogan (2004), Turoff, Chumer Walle and Yao (2004) carried out a study to propose the development of a virtual command and control centre to conduct disaster response operations with the help of a multi-level human network. These authors proposed a framework for the design and development of a comprehensive Dynamic Emergency Response Management Information System using the design basis of EMISARI and design concepts reported in the technical literature. Turoff, Chumer, Hiltz, Klashner, Alles, Vasarhelyi, Kogan (2004), Turoff, Chumer Walle and Yao, (2004) considered an exhaustive set of design premises and principles in the proposed framework of DMIS based on five essential concepts of a group communication system; these include system metaphor, human role identification, notification, context visibility and hypertext. The year 2004 marks a rapid increase in the research on the development of DMISs across the globe, as discussed in the forthcoming sections.

The first application of DMIS in South Asia started with the development of Sahana in Sri Lanka in the aftermath of the 2004 tsunami in the Indian Ocean (Ariyabandu, 2009). Sahana is a free and open source computer programme. The use of hypertext preprocessor (PHP) as a programming language in Sahana provides compatibility with the geographically referenced data on the geographical information system (GIS) platform which greatly facilitates response to a disaster. This software has been recognised internationally and has received several awards (Careem et al., 2006). The details of a number of available and proposed systems in different countries have been discussed in the forthcoming sections.

**DM information systems**

The increased importance of a comprehensive DMIS to meet the functionality requirements of planning and conducting DM functions was realised after the events of 9/11 attacks on twin towers in USA (Turoff, Chumer, Hiltz, Klashner, Alles, Vasarhelyi, Kogan, 2004; Turoff, Chumer Walle and Yao, 2004). As a result, there has been a notable increase of research in DMISs since 2004 which appears to be occurring in parallel in different countries. This is owing to the reason that it is not possible to design a generic system which could satisfy needs of all countries for all disasters (Dorasamy and Raman, 2011).

A DMIS is a software package which can perform networking, scheduling, and data resource analysis and management during a disaster to enable accelerated response and recovery of the affectees. The overall design principle of a DMIS is based on development of software that provides the required functionality. The structure of DMIS software involves coordination of humans involved in the response by allowing various levels of command and control using a communication system. The developed application requires the use of computing technology for the operation of software which can easily be met with the availability of present day hardware such as personal computers, laptops, personal digital assistants (PDAs), smart phones and wireless gadgets. The system can be applied in all disaster situations such as post-earthquake disaster, fires, bomb blasts, hazardous material
spill, transportation interruptions, etc. It, nevertheless, requires that adequate training and experience is provided to the involved personnel.

The available studies on DMIS from different parts of the world that are examined in the presented paper are summarised in Table I. The studies in Table I have been organised in the same chronological order as they appeared in the literature. The method of analysing these systems includes a comparison of available features and their functionality. The analysis of these studies indicates large variations in their scope and functionality requirements. Different modules and their components which are found from the analysis of studies in Table I are illustrated in Figure 1. The availability of any of these is dependent on the functionality requirements of the proposed system and varies from one system to the other. It is seen in Figure 1 that DMISs can be divided in two categories: disaster inventory system (DIS) and information management system. DIS is comprised only a data inventory module in which a disaster database is stored. This database consists of baseline data, information of resources, historical records of disasters, missing person’s registry, information of response agencies and response facilities.

A maximum of five modules have been suggested for a typical DMIS. These include data inventory, mitigation, administrative control, response and recovery and communication infrastructure. Of these (except mitigation), four have been proposed in the majority of the studies included in Table I. A review of the studies in Table I indicates that most of these focus response and recovery (R and R) phase of an on-going disaster; only few consider pre-disaster mitigation phase. Mitigation process may start with hazard scenario analysis using available simulation models for different hazards. Hazard and risk maps may be developed for an area based on the simulation results. Accordingly contingency plans and DRR strategies may be adopted to reduce the damages associated with different disasters.

Administrative control module is made available only to authorised managers or administrators. This module provides facility of viewing the situation of on-going disaster using monitoring system to an administrator who may declare an emergency.

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Author</th>
<th>DM Component</th>
<th>Origin</th>
<th>S. No.</th>
<th>Author</th>
<th>DM Component</th>
<th>Origin</th>
</tr>
</thead>
</table>

Table I. Summary of examined studies on DMISs
Accounting and finance management is also carried out using this module. Request for financial aid and resources may be forwarded by the administrator based on a need analysis. A message/chat window may be used for the sharing situational updates among disaster managers.

R and R module is the main module of DMIS which is used during the disaster as it manages the entire response to the disaster. Resource allocation is carried out using this module according to the requests received from the field operators. Missing persons' management and shelter management is carried out according to the data entered by the field operators. Early warning system is also a part of this module which is used to disseminate information on an emerging dangerous situation to enable action in advance to reduce the risk.

Table I also provides the details of main focus of the proposed systems. It is seen in Table I that, while all the studies focus on the response to the needs of the affectees, the studies in the reference Eraslan et al. (2004) and EmerGeo (2016) provide modules for pre-disaster mitigation activities. These activities include development of hazard and risk maps of the target areas, and providing education material for DRR (Figure 1). In this respect, EmerGeo (2016) provides the advanced capability of computer simulations to conduct scenario analysis for different types of hazards. Based on these simulations, design of protection against a hazard (where applicable) is carried out by the programme and its cost is calculated. The preparedness module in reference Lakovou and Douligeris (2001) provides contingency plans to deal with hurricane, and information on research and technology development to reduce the level of this hazard.

Further, the systems proposed in references Corporación (1994), Kitamoto (2000) and Tata (2009) are DISs, as opposed to DMIS. As mentioned earlier, a DIS is a database of past and on-going disasters without management functionality (Figure 1). This information is made available publicly through a website or a web portal. The system suggested by
Ganeshkumar and Ramesh (2010) is standalone software which contains databases of response agencies and facilities; the system is limited to finding out the shortest route for the nearest response facility. Saeed et al. (2013) proposed a GIS-based web application suitable for Android smart phones which provides the location of response facilities in the affected areas. Although these aforementioned systems are limited in their applications, these have been included here to compare the design of databases in these with those in DMISs.

Finally, the systems proposed in references Corporación (1994), Kitamoto (2000), Tata (2009), Eraslan et al. (2004), EmerGeo (2016), PAHO (1998), International (2005) and Lakovou and Douligeris (2001) have been developed and tested during the real disasters. These are available with the trade names of DesInventar, Digital Typhoon, India Disaster Knowledge Network, AFAYBIS, EmerGeo Fusionpoint, SUMA, PHOENIX and IMASH, respectively. A comparison of design of data inventory and communication infrastructure suggested in the studies in Table I has been presented in the forthcoming sections.

Design of data inventory

Data inventory is a basic and essential module of any DMIS. The data required for DM can be divided into pre-disaster data and post-disaster data. The hierarchical structure of the data for the access and presentation of the information in both these categories is illustrated in Figure 2. It is seen in Figure 2 that pre-disaster data consist of baseline data, past records of disasters and the available resources. Post-disaster data consist of available resources, temporarily dislocated people (TDPs) and missing persons. Therefore, the data of available resources are overlapped by both the pre-disaster and post-disaster stages, as these data are continuously updated after a disaster.

The resources may comprise of human and material and equipment resources. Human resources are provided by both the response organisations (civic agencies, NGOs, etc.)
and volunteers. Material and equipment resources may consist of logistical support such as vehicles, tools, supplies, etc. Resource management is carried out by managing demand and supply for both types of resources using the data available in the data inventory. The demands are those requests for the emergency supplies that are received from the disaster sites.

The design of data inventory is based on the following essential elements.

Data input template: a data input template is a form that is used to collect different data sets as shown in Figure 1. The template provides a user interface whose contents are derived from the logical data structure; this is identified during the analysis stage of the system development using flow diagrams.

Data authentication: data authentication refers to the confirmation of the origin and integrity of data as reliable data are essential in a disaster situation. The reliability of data is safeguarded by authenticating the data entry into the system to the authorised entities.

Data format standardisation: data format standardisation transforms the stored data and information into a well-defined and consistent form which is compatible with the scheme of database development. Disaster information may be collected in different forms and order, and spelled differently. It may become out-dated or may contain errors. For example, if data are collected over telephone for entry into the system, spelling variations of names cannot be ruled out. Typing errors are also possible at the time of data entry. The data format standardisation attempts to deal with these issues. Conversion of the input data into a well-defined form and segmenting it into many smaller output fields, allows the linkage process to become more accurate.

Data category analysis: the aim of data category analysis is to organise the data into different categories for their effective and efficient use. A well-planned data classification system facilitates data search and its retrieval. Data classification procedures are used to define those categories and criteria that DMIS will use for data classification. Based on the data classification, security standards (that specify appropriate handling practices for each category) and storage standards (that define the data lifecycle requirements) are provided.

Data redundancy: data redundancy refers to a condition in which the same data may be stored in two separate spaces in a single database. This can mean two different fields within the same database, or two different spots in multiple software environments. Data redundancy can occur by accident or deliberately for backup and recovery purposes.

The schematic of data flow in a DMIS is illustrated in Figure 3. It is seen in Figure 3 that different types of data, such as geographically referenced data, demographic distribution in an area, data related to response organisations, available resources and situational updates, are collected independently and are stored in local servers by the authorised personnel using the data input templates. Data authentication is also ensured at this stage. These data are fed to the main server through internet after the data format standardisation is completed. The information collected from various sources is filtered to retain that information which is likely to satisfy the user needs. The filtered information is organised in different categories through indexing, categorisation and linking tools. Of these, indexing plays a critical role in achieving fast and accurate search of correct data at an appropriate time. Linking connects the relevant information with knowledge base related to a specific disaster event. The extracted information can be shown in different forms such as utility maps, evacuation routes from the disaster site, medical and shelter facilities, demographic distribution maps and real-time situation of disaster related damages at a site.

Table II presents the details of data and database type considered in the studies examined in the presented paper. It is noted in Table II that database types have been mentioned in only few studies. In addition, no particular trend appears in Table II in this regard. It is also noted in Table II that most of the studies consider the data of geographically referenced baseline data and available resources which increases the effectiveness of response to a disaster by providing information about alternate routes, distribution of supply lines and buried facilities.
Only few studies have included the modules of historical disasters and natural hazards which are needed for obtaining essential information of the earlier disasters, such as affected area, magnitude, number of casualties and estimated loss. Only the studies in reference Eraslan et al. (2004), Ariyabandu (2009) and Tata (2009) have suggested data for shelters. With the exceptions of Ariyabandu (2009), Meissner et al. (2002) and Murphy and Jennex (2006), no other study considers the module for missing persons registry. Finally, Ariyabandu (2009) is the only study which requires TDPs registration.

**Communication infrastructure design**

Communication works as a backbone in DM as it allows quick and effective implementation of the decisions (Eraslan et al., 2004). Reliable and protected communication network is needed for an effective DM system. The network should be inter-operable with those of other organisations. A committed two-way communication system is required for swift response to a disaster which should be cost effective, reliable, flexible and accessible at all locations.

The revolutionary changes in the communication network in the twenty-first century have enabled the technology-based communication more swift and reliable. Commonly used modern communication infrastructure systems include local area network (LAN) and wide area network (WAN). LAN refers to a group of computers and associated devices (printers, scanners, etc.) that are connected to a server within a small geographic area such as an
The office building. Ethernet and Wi-Fi and are the most commonly used LAN technologies. WAN consists of two or more LANs and allows a computer network to span over a relatively large geographical area. Public networks, such as the telephone system, are used to connect computers to WAN. WAN could be operational through Ethernet, Wi-Fi and general packet radio service (GPRS). The access to the remote computer is done through an internet browser. Internet cable network or public network becomes vulnerable to the damages during a disaster. As a result, wireless LAN, wireless WAN and satellite communications are considered to be reliable communication infrastructures in these situations. In addition to these, wireless personal area network (WPAN) can also be used to connect devices in an individual's limited workspace such as Bluetooth.

A schematic of the overall communication demand during a disaster is illustrated in Figure 4. It is seen in Figure 4 that a site response centre is established by the first responders after reaching the disaster site to send updates to the central command unit. This unit, in turn, provides these updates to the local government offices responsible for dealing with the disaster. The information is analysed by the public decision makers and resource allocations are carried out using the administrative control module of DMIS (Figure 1). Thereafter, the relief agencies are mobilised by the government to deliver the supplies to the affectees. The data and information at local government offices are also stored as backup in remote data servers for protection.

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Author</th>
<th>Baseline data</th>
<th>Historical records of disasters</th>
<th>Resource registry</th>
<th>TDPs registry</th>
<th>Missing persons registry</th>
<th>Database type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Corporación (1994)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Online database</td>
</tr>
<tr>
<td>2.</td>
<td>PAHO (1998)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>3.</td>
<td>Kitamoto (2000)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>PostgreSQL</td>
</tr>
<tr>
<td>4.</td>
<td>Lakovou and Douligeris (2001)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Oracle</td>
</tr>
<tr>
<td>5.</td>
<td>Meissner et al. (2002)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>6.</td>
<td>Eraslan et al. (2004)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>8.</td>
<td>International (2005)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>9.</td>
<td>Murphy and Jennex (2006)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Online database</td>
</tr>
<tr>
<td>10.</td>
<td>Ariyabandu (2009)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>PostgreSQL (LAMP)</td>
</tr>
<tr>
<td>11.</td>
<td>Tata (2009)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>12.</td>
<td>Atteh et al. (2010)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Oracle</td>
</tr>
<tr>
<td>13.</td>
<td>Ganeshkumar and Ramesh (2010)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>GIS Server</td>
</tr>
<tr>
<td>14.</td>
<td>Hassan et al. (2011)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>15.</td>
<td>Dorasamy et al. (2012)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>16.</td>
<td>Lee et al. (2012)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>17.</td>
<td>Saeed et al. (2013)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>GIS Server</td>
</tr>
<tr>
<td>18.</td>
<td>Li et al. (2014)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
<tr>
<td>19.</td>
<td>EmerGeo (2016)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>NA</td>
</tr>
</tbody>
</table>

Notes: n/a, not available. Includes data of material and equipment resources only; b includes data of shelter only; c includes data of human resources only

Table II. Summary of data inventory module in examined studies
The communication infrastructure proposed by the researchers for DMISs in the examined studies is summarised in Table III. The details of infrastructure design for internal communication within an organisation and externally between different organisations are given in Table III. A difference in the proposed communication infrastructure with the development of technology is noted in Table III and the increased use of wireless LAN and WAN is evident in the studies presented over the last one decade. Similarly, use of GPRS is also introduced to facilitate external communication in the absence of IT infrastructure network. It is noted in Table III that most of the studies have suggested WAN to connect site response centre with both the central unit and the local government. On the other hand, LAN has been proposed to be a communication channel within organisations. Few studies have also considered satellite communications which improve the reliability and survivability of services during a disaster. An exception from the remaining studies is that of Li et al. (2014) where a peer-to-peer (P2P) architecture has been proposed to manage distributed data sets of the affected community. A P2P network is created when two or more computers are connected and share their resources without going through a separate server.

The studies carried out by Turoff, Chumer, Hiltz, Klashner, Alles, Vasarhelyi, Kogan (2004), Turoff, Chumer Walle and Yao (2004), Hassan et al. (2011), Dorasamy et al. (2012) and Murphy and Jennex (2006) did not provide specific details regarding the communication infrastructure although these authors have mentioned use of devices for communication such as personal computers, servers, laptops, PDAs and smart phones. Communication infrastructure is not required in DISs, as mentioned earlier.

Management of resources
A link is provided between the data inventory and communication infrastructure during the management of resources. An overview of resource management cycle is illustrated in Figure 5. It is seen in Figure 5 that resource management cycle starts at field operators who provide disaster situational updates to the site response centre established at the disaster sites. This communication could be through either voice communication (using wireless or
satellite phones) or e-mails. These data are entered into DMIS and are sent to the central command unit through the employed communication infrastructure (Table III). The data are analysed at the central command unit and reports are generated which are sent to local government offices or DM authorities in the form of demand for aid supplies. The local government offices maintain the record of available resources. The demands received from the central command unit are compared with those available and decisions for the transactions of resources are made by the public decision makers. Damage statistics and situational updates are also prepared by the local government/DM authority offices. Note that the details of resources suggested by different studies are given in Table II.

**System development**

The available information related to the development of DMIS is limited in the literature. Eraslan et al. (2004) used unified process (Arlow and Neustadt, 2008) which is an iterative and incremental software development process framework. Unified modelling language (UML) was used to describe the system structure and behaviour. The interaction between different actors in the system was specified with use cases (Cockburn, 2008) which are part of UML diagram. Sahana (Ariyabandu, 2009) is open source software developed using Linux-Apache-MySQL-PHP/Postgres (Teter, 2006) web development platform; the code of the programme was written in PHP scripting language. Ganeshkumar and Ramesh (2010) used ArcGIS (ESRI, 2011), MapObjects (ESRI, 2011) and the programming language platform VB.NET (Vick, 2004) to develop the application for finding the shortest route to a relief centre. EmerGeo Fusionpoint (EmerGeo, 2016) employs different software libraries

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Author</th>
<th>Internal communication</th>
<th>External communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Corporación (1994)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>3.</td>
<td>Kitamoto (2000)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>10.</td>
<td>Ariyabandu (2009)</td>
<td>LAN</td>
<td>WAN</td>
</tr>
<tr>
<td>11.</td>
<td>Tata (2009)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>13.</td>
<td>Ganeshkumar and Ramesh (2010)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>14.</td>
<td>Hassan et al. (2011)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>16.</td>
<td>Lee et al. (2012)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>17.</td>
<td>Saeed et al. (2013)</td>
<td>Peer-to-Peer</td>
<td>Peer-to-Peer (P2P)</td>
</tr>
</tbody>
</table>

**Note:** n/a, not applicable

Table III. Summary of communication infrastructure design in examined studies
such as ETeam (NC4, 2016), WebEOC (Intermedix, 2016), Intergraph (Intergraph, 1976), Versaterm, Google Maps, Microsoft Bing and Open Street Map to provide functionality and services. Generic mapping tools (Wessel and Smith, 1998), advanced visualisation systems (AVS) (Upson et al., 1989), Java (Arnold et al., 2006), oyster, oracle (Oracle, 2016) and hypertext markup language interface supported through a set of UNIX (Raymond, 2004) script have been employed in the development of IMASH (Lakovou and Douligeris, 2001).

Future opportunities
The review of the studies presented in the aforementioned sections indicates a number of future opportunities which can increase the effectiveness of DMISs in managing response to a disaster. Firstly, baseline data (demographic distributions, building typology, transportation network, critical facilities and supply lines) is important information for the post-disaster relief operations. Therefore, geographically referenced baseline data should be collected and stored in a separate remote server. The information specific to a disaster event can be pushed to DMIS from this server through a communication mechanism. These data can be beneficial for the response agencies for effective response to a disaster in a particular area. As a result, availability of geographically referenced baseline data could improve the response in combination with DMIS. The provision of these data can be beneficial for the response agencies to make an effective action plan to respond to a disaster in a particular area.

Effective DRR reduces the need of response to the needs of affectees to a large extent. As a result, more attention is paid to mitigate the disaster risk. The addition of a module for
hazard analysis would appear a logical step for a DMIS. The obtained results could be combined to provide integrated multi-hazard reports and maps which can be used by disaster managers to develop DRR strategies for the reduction of risk of damages.

Of all the studies examined in the presented paper, only EmerGeo (2016) includes a module of early warning. This module is used for alerting the agencies and communities of an emerging dangerous condition which provides an opportunity to take actions in advance to reduce the associated risks. In addition, sufficient warning time can help in making better preparations to deal with post-disaster situations. The longer the preparation time the greater the savings in human life, household assets, livestock and stored provisions. In view of these benefits, a module of early warning system should be provided in DMIS. Different agencies such as meteorological department, flood authorities, ocean authorities and other agencies dealing with different natural hazards may be connected to the central command unit to provide regular situational updates. The central command unit can send situational warning messages to the local government officials, response agencies and media to alert the communities.

Accurate and instantaneous data over large areas can be obtained using satellites. Therefore, remote sensing could be a useful tool for DM as it can provide a clear picture of a disaster stricken region immediately. These data can also be used for risk modelling, early warning alerts and for identifying the escape routes.

Finally, the role of media is crucial during a disaster and can be included in an automated DMIS. This can help in sending early warning alerts, in keeping people updated about the situation and relief activities, and in sending appeals for aid supplies during a disaster.

These aforementioned suggestions have been made part of DMIS (Rafi and Lodi, 2015) that the authors have proposed to National Disaster Management Authority in Pakistan.

Concluding remarks

An effective response to a disaster requires swift flow of information and integrated response activities. The use of computing technology in DM can help in achieving these objectives. Automated DMISs have been suggested and developed in different parts of the world. A significant increase in the research in the field of DMIS is noted since about 2004. This paper reviews the studies related to the proposed and developed DMISs with a view to examine their similarities. In addition to DMIS, systems providing disaster inventory only (DIS) have also been proposed in the literature. These are the databases of past and on-going disasters without any response management functionality.

Database inventory and communication infrastructure are the two major components of a DMIS. Variations in the type of data proposed for different DMISs exist based on the functionality requirements. No particular trend appears from the reviewed studies for the design of databases as this is dependent on the functionality requirements of the system which were found to vary from one study to the other. Only few of the examined studies provide management of missing persons and TDPs. The trend for communication infrastructure design is towards satellite communication owing to its higher reliability and survivability after a disaster. The future opportunities identified, based on the review of the presented studies, include creating geographically referenced databases, use of remote sensing data, and modules for DRR, early warning system and media communication. Lastly, understanding of the areas of improvement in DMIS can increase by testing the systems in real situations.

References


A comparative study of DMISs


About the authors
Muhammad Masood Rafi is Professor and Chairman in the Department of Earthquake Engineering at the NED University of Engineering and Technology, Pakistan. He is the Editor-in-Chief of the NED University Journal of Research. His current research interests include behaviour of FRP reinforced concrete (RC) structures, fire resistance of RC structures, experimental and finite element analyses of RC structures, recycling of concrete, seismic analysis and retrofitting of RC and masonry structures, and disaster management and mitigation. Muhammad Masood Rafi is the corresponding author can be contacted at: rafi-m@neduet.edu.pk.

Tariq Aziz is Lecturer in the Department of Civil Engineering at the NED University of Engineering and Technology, Pakistan. He received his Bachelor's from the Department of Urban and Infrastructure Engineering and Masters from the Department of Earthquake Engineering at NED University of Engineering and Technology, Pakistan.

Sarosh Hashmat Lodi is Professor in the Department of Civil Engineering and Vice Chancellor at the NED University of Engineering and Technology, Pakistan. He possesses nearly 30 years of experience in teaching and research, specializing in seismic design and earthquake related research. He has been actively involved in the development of bridge data management system to monitor and assess the state of bridges in Pakistan.
Understanding social networking sites continuance
The perspectives of gratifications, interactivity and network externalities

Chun-Ming Chang
Department of International Business, Ming Chuan University, Taipei, Taiwan

Abstract
Purpose – The purpose of this paper is to develop a theoretical model to investigate the determinants of continuance intention toward social networking sites (SNSs) by integrating the perspectives of the uses and gratifications theory, perceived interactivity and network externalities.
Design/methodology/approach – Data collected from 255 Facebook users in Taiwan were used to test the proposed model. The partial least squares method was used to test the measurement model and the structural model.
Findings – The findings reveal that emotional gratifications and social gratifications are the key predictors of users’ continuance intention toward SNSs. Further, the results indicate that perceived network size, perceived complementarity, machine interactivity and person interaction influence information gratifications significantly, while perceived complementarity, machine interactivity and person interactivity exert positive effects on emotional gratifications. Finally, the results show that machine interactivity and person interactivity impact social gratifications positively, whereas perceived network size and perceived complementarity affect machine interactivity and person interactivity significantly.
Originality/value – This study is one of the earliest research inquiries to examine the effects of various types of gratifications on continuance intention. It is also one of the earliest studies to identify the antecedents of gratifications from social factors and technological attributes simultaneously.

Keywords Continuance intention, Social networking sites, Uses and gratifications theory, Network externalities, Perceived interactivity

Paper type Research paper

1. Introduction
Social networking sites (SNSs) (e.g. Facebook and Twitter) have attracted millions of internet users to create their profiles and connect these profiles to others’ profiles and thus build their personal network (Cheung and Lee, 2010; Ku et al., 2013). However, the competition among SNSs is intense (Zhao and Lu, 2012), and thus users can easily switch from one website to another one that provides similar services (Li et al., 2006). In fact, researchers have found that users have reduced active usage of SNSs and even abandoned SNSs (Huang et al., 2014; Ku et al., 2013). Because the success of information systems depends on users’ continued usage (Bhattacharjee, 2001), it is important to understand what may enhance users’ intention to continue using SNSs.

In the existing literature, one frequently endorsed theoretical base for exploring the determinants of users’ continuance intention toward SNSs is the uses and gratifications (U&G) theory (Hsu et al., 2015; Huang et al., 2014; Ku et al., 2013). Researchers agreed that fulfilling one’s gratification needs is crucial in fostering users’ continuance intention (Ku et al., 2013). However, few studies have been conducted to examine how various types of gratifications affect continuance intention. To further investigate the links between different types of gratifications and continuance, this study divides gratifications in three categories: information gratifications, emotional gratifications and social gratifications, based on the U&G theory (Chiang, 2013; Wei et al., 2015).

This research was funded by the grant of the Ministry of Science and Technology of Taiwan under Contract No. NSC 102-2410-H-156-007.
Furthermore, given that gratifications are the crucial factors affecting continuance intention, some researchers have begun to investigate the antecedents of gratifications. For example, Ku et al. (2013) and Wei and Lu (2014) suggested that social factors (e.g. perceived critical mass, subjective norms and perceived number of users) may impact users’ gratifications. Nonetheless, there is still much to learn about the antecedent factors determining gratifications from various perspectives. Prior research noted that SNSs are social computing tools (Zhao and Lu, 2012), and the outcome of system usage derives from the interactions between social and technical factors (Hong et al., 2013). Hence, it may be appropriate to examine the predictors of gratifications from the aspects of social factors and technological attributes. In fact, researchers have found that the utility derived from the use of SNSs is related to social factors (i.e. the number of users and perceived complementarity) (Lin and Bhattacharjee, 2008; Lin and Lu, 2011) and technological attributes (i.e. features of interactivity) (Wei et al., 2015). However, few studies have been carried out to identify the antecedents of gratifications from social factors and technological attributes simultaneously. To fill this knowledge gap, this study considers network externalities (perceived network size and perceived complementarity) as social factors and perceived interactivity (machine interactivity and person interactivity) as technological attributes to test their effects on perceived gratifications.

Overall, the purpose of this study is to address the following questions:

RQ1. Do perceived gratifications influence continuance intention?

RQ2. How do perceived interactivity and network externalities influence perceived gratifications?

RQ3. To what extent do network externalities affect perceived interactivity?

The findings of this study may help both academicians and practitioners gain insights as to how continuance intention toward SNSs may be promoted.

2. Theoretical background

2.1 U&G theory

The U&G theory is a psychological theory that attempts to explain why people select a specific medium to fulfill their social and psychological needs (Currás-Pérez et al., 2013; Ku et al., 2013; Lee and Ma, 2012). According to the U&G theory, users are active and goal-oriented on their chosen medium, and they select a medium to best fulfill their needs (Hsu et al., 2015; Ku et al., 2013; Lee and Ma, 2012). In this regard, the U&G theory has been considered as one of the effective paradigms for understanding the motivations bringing users to a specific medium (Lee and Ma, 2012; Stafford et al., 2004), including SNSs (Lee and Ma, 2012).

In addition, the U&G theory has been employed to identify the factors affecting users’ continuance intention toward SNSs. For example, Ku et al. (2013) suggested that the fulfillment of gratification-related needs is an important predictor of continuance intention toward SNSs. The analytical results of Wei et al. (2015) revealed that informational needs and social needs are the two critical factors motivating users to interact with others, which, in turn, affect users’ attitude toward SNSs and their stickiness intention toward these SNSs. On the other hand, Chiang (2013) reported that users may continue using SNSs to satisfy three types of needs: informativeness, social interactivity and playfulness. Based on the above arguments, this study considers that information gratifications (fulfillment of information needs), emotional gratifications (positive emotion or enjoyment) and social gratifications (fulfillment of social needs) have the potential to affect users’ continuance intention.

Moreover, given the importance of gratifications with respect to continuance intention, it is important to explore the factors that enhance users’ gratifications needs. Although Ku et al. (2013) and Wei and Lu (2014) have tested the influences of social factors
(e.g. perceived critical mass, subjective norms, and perceived number of users) on gratifications, the spirit of SNSs emphasizes user’s interaction and involvement using information technologies (Lin and Lu, 2011). Hence, technological attributes may also impact users’ need for gratifications. However, little attention has been given to the relationships between social factors, technological attributes and gratifications simultaneously. To fill this gap, this study aims to employ the perspectives of network externalities and perceived interactivity to identify the antecedents of gratifications from social factors and technological attributes simultaneously.

2.2 The perspective of network externality
Network externalities refer to the utility or value that users obtain from consuming a product or service as enhanced as the users and complementary product/service increase (Katz and Shapiro, 1985). Thus, information technologies such as SNSs are valuable to users when many other people also use them and when some compatible products are available for users that make technology usage easier and convenient (Brynjolfsson and Kemerer, 1996). In general, the enhanced benefit of technology usage due to the continued growth in the number of users and the availability of complementary goods is called network effects, while the products or services that exhibit such effects are called network goods (Lin and Bhattacherjee, 2008). SNSs, as popular interactive technologies, usually display strong network effects (Lin and Bhattacherjee, 2008). Thus, network externalities aptly explain the relationship between social network effects and user gratifications derived from SNS usage (Zhao and Lu, 2012).

Network externalities can be divided in two types: direct network externalities and indirect externalities (Lin and Bhattacherjee, 2008; Zhao and Lu, 2012). Direct network externalities refer to the number of participants in a given network (Lin and Bhattacherjee, 2008). Generally speaking, direct network externalities are derived from the increase in the number of users of a specific product or service (Lin and Lu, 2011). In this sense, researchers generally use perceived network size to reflect direct externalities (Lin and Bhattacherjee, 2008; Zhao and Lu, 2012). On the other hand, indirect network externalities display the increased sense of benefits from using a product or service due to the development of related/complementary products or services (Lin and Bhattacherjee, 2008; Lin and Lu, 2011). Indirect network externalities are generally represented by perceived complementarity (Lin and Bhattacherjee, 2008; Lin and Lu, 2011; Zhao and Lu, 2012). In this study, perceived network size reflects the users’ perception that other users are using the SNSs (Zhao and Lu, 2012), while perceived complementarity refers to the complementary goods or services provided by information technology providers or third-party vendors to support the use of SNSs (Lin and Bhattacherjee, 2008).

In previous studies, network externalities have been considered as factors that may impact personal benefits and gratifications. For example, in a study focusing on the user adoption of instant messages, Lin and Bhattacherjee (2008) found that perceived network size and perceived complementarity have positive impacts on network benefits. In the study exploring the factors affecting users’ continuance intention toward SNSs, Lin and Lu (2011) noted that the number of peers, the number of members and perceived complementarity have significant effects on perceived usefulness, while the number of peers and perceived complementarity exert a positive influence on enjoyment. In addition, similar findings were revealed by Zhao and Lu (2012). Their research revealed that perceived network size and perceived complementarity have significant effects on enjoyment and connectedness. Wei and Lu (2014) reported that network externalities have positive influences on gratifications in the setting of online games. By synthesizing the above arguments, this study recognizes that network externalities could be considered as social factors that may impact perceived values and, thus, satisfy users’ needs.
2.3 The role of perceived interactivity

Perceived interactivity has been considered as a critical feature of modern media (Wu and Chang, 2005). Perceived interactivity could be defined from various perspectives, such as the communication process (e.g. interchange and responsiveness), user perception (i.e. perception of interactivity) and technology features (e.g. control and two-way communication) (Lowry et al., 2009; McMillan and Hwang, 2002). However, there is little agreement on a common definition for it (Ko et al., 2005; Lowry et al., 2009). According to Zhao and Lu (2012), this study adopts the viewpoint of user perception to investigate how users perceive and experience interactivity after using SNSs. Drawing on Thorson and Rodgers (2006), perceived interactivity is defined as “the extent to which users perceive their experiences of interpersonal interactions and the sense that they are in the presence of a social other” (p. 36).

Perceived interactivity is also a multidimensional concept that can be divided in various constructs from different perspectives (Ko et al., 2005; Lu et al., 2010; Zhao and Lu, 2012). For example, Hoffman and Novak (1996) divided interactivity in two dimensions: machine interactivity and person interactivity. Several researchers considered that interactivity can be categorized into three types: human–human interaction, human–message interaction and human–machine interaction (Hsu et al., 2015; Lu et al., 2010). Chen and Yen (2004) offered five dimensions of interactivity: playfulness, choice, connectedness, reciprocal communication and information collection. Following Zhao and Lu (2012), the viewpoint of Hoffman and Novak (1996) is employed to reflect the interactivity occurring between machine (machine interactivity) and human (person interactivity) to access content and the interactivity occurring between humans because interpersonal communication among users in SNSs is generally composed of social and technical interactions.

Machine interactivity refers to the extent to which a user can participate in modifying content using a communication technology in real time (Wu and Chang, 2005). Zhao and Lu (2012) considered that machine interactivity reflects the extent to which a user feels in control of his/her interaction through SNSs with others. In this study, machine interactivity is defined as the degree to which a user believes that he/she can easily interact with others using the functions of SNSs (Lu et al., 2010). Moreover, person interactivity is generally defined as the interactions between people occurring over the communication technology (Wu and Chang, 2005). Zhao and Lu (2012) argued that person interaction is the extent to which an individual perceives how fast and frequently other users of social media respond to his/her message. Thus, person interactivity is conceptualized as the degree to which a user believes that he/she can easily communicate with others using the functions of SNSs (Lu et al., 2010).

Prior research posited that motivations would enable users to use SNSs to interact with others to satisfy their needs (Hsu et al., 2015; Wei et al., 2015). This implies that interactions with information technologies and users determines users’ gratifications. However, the influences of perceived interactivity on gratifications have received little attention from researchers. Hence, this study aims to treat machine interactivity and person interactivity as technological attributes to test their impacts on gratifications. In addition, Zhao and Lu (2012) postulated that network externalities are useful predictors of perceived interactivity. Therefore, perceived network size and perceived complementarity are considered as predictors of both machine interactivity and person interactivity.

3. Research model and hypotheses

The research model of this study is presented in Figure 1. The model depicts that perceived gratifications (information gratifications, emotional gratifications and social gratifications) are the predictors of continuance intention, while network externalities (perceived network size and perceived complementarity) and perceived interactivity (machine interactivity and
person interactivity) are the critical factors fostering perceived gratifications. Finally, the model asserts that network externalities impact perceived interactivity. The reminder of this section will define the constructs and justify the hypotheses.

3.1 Perceived gratifications and continuance intention
Perceived gratifications refer to one's perception that his/her personal needs can be fulfilled using SNSs (Ku et al., 2013). From the viewpoint of utility maximization (Li, 2012), people tend to portray a specific behavior to better satisfy their needs. The social cognitive theory also asserts that positive outcome expectations motivate an individual to portray a specific behavior (Bandura, 1997). Ku et al. (2013) found that perceived gratifications are strong determinants of users' continuance intention toward SNSs. Hence, this study expects that gratifications of information, entertainment and socialization will positively impact continuance intention. We propose the following hypotheses:

\[ H1a. \] Continuance intention is positively associated with information gratifications.

\[ H1b. \] Continuance intention is positively associated with emotional gratifications.

\[ H1c. \] Continuance intention is positively associated with social gratifications.

3.2 Perceived network size and perceived gratifications
As mentioned earlier, perceived network size represents one's perception about the number of users that are using the same SNS. As a general rule, when the network size increases, users may believe that they can communicate with others using the same SNSs easily. Users' perception of enjoyment will be enhanced when users perceive more people joining the SNSs (Lin and Lu, 2011). Frequent interpersonal interactions will expand their connections with others and may, thus, enhance their perception of being connected with others and getting responses from others quickly (Zhao and Lu, 2012). Based on the above arguments, this study proposes that perceived network size could facilitate gratifications of information, emotion and socialization. Wei and Lu (2014) also provided empirical evidence
to support the linkage between perceived network size and gratifications. Therefore, we propose the following hypotheses:

\[ H2a. \text{ Information gratifications are positively associated with perceived network size.} \]
\[ H2b. \text{ Emotional gratifications are positively associated with perceived network size.} \]
\[ H2c. \text{ Social gratifications are positively associated with perceived network size.} \]

3.3 Perceived complementarity and perceived gratifications
Perceived complementarity is expected to facilitate users’ perception of gratifications (Zhao and Lu, 2012). According to Lin and Bhattacharjee (2008) and Lin and Lu (2011), the high level of complementarity helps users conduct interactions with others frequently. This may increase users’ perception that they can obtain more information they need from others using SNSs (Lin and Bhattacharjee, 2008). Moreover, researchers postulated that complementary products provide ubiquitous access to users to share interesting photos and videos with others, and, thus, affects users’ perception of playfulness and interconnectivity (Zhao and Lu, 2012). Based on the above studies, this study expects that perceived complementarity may increase users’ gratifications of information, emotion and socialization. Therefore, we propose the following hypotheses:

\[ H3a. \text{ Information gratifications are positively associated with perceived complementarity.} \]
\[ H3b. \text{ Emotional gratifications are positively associated with perceived complementarity.} \]
\[ H3c. \text{ Social gratifications are positively associated with perceived complementarity.} \]

3.4 Machine interactivity and perceived gratifications
According to Lee and Ma (2012) and Lu et al. (2010), SNSs provide a relatively high level of interactivity for users to communicate with others and fulfill their needs of socialization, information exchange, and entertainment. From the standpoint of the media synchronicity theory (Dennis et al., 2008), media capability plays a vital role in determining the ability of users to interact with others. The better media capability supporting interactions between users and others will make users believe that using SNSs could satisfy their needs. Because machine interactivity describes the interactive features of technology and reflects the perception of its usability (Zhao and Lu, 2012), this study may reasonably believe that machine interactivity impacts gratifications positively. Therefore, we propose the following hypotheses:

\[ H4a. \text{ Information gratifications are positively associated with machine interactivity.} \]
\[ H4b. \text{ Emotional gratifications are positively associated with machine interactivity.} \]
\[ H4c. \text{ Social gratifications are positively associated with machine interactivity.} \]

3.5 Person interactivity and perceived gratifications
In general, users tend to interact with others via SNSs to fulfill their needs (Hsu et al., 2015), implying that person interactivity may impact perceived gratifications. According to Panteli and Sockalingam (2005), sustained and lasting online interactions with others enable users to build mutual trust, which, in turn, facilitates trust-related behaviors, such as information sharing and cooperation. This implies that frequent interactions over SNSs may enable users to acquire the useful information they need and may increase social network density. Hence, person interactivity is expected to have positive influences on information and social gratifications. In addition, Lu et al. (2010) and Wu and Chang (2005) found that users feel
enjoyment once they have intensive interpersonal interactions via SNSs, indicating that
person interaction is a useful factor determining emotional gratifications. By synthesizing
the above arguments, this study proposes the following hypotheses:

\[ H5a. \] Information gratifications are positively associated with person interactivity.

\[ H5b. \] Emotional gratifications are positively associated with person interactivity.

\[ H5c. \] Social gratifications are positively associated with person interactivity.

3.6 Perceived network size and perceived interactivity

Researchers assert that one’s decision of using an interactive technology generally depends
on how many of his/her friends, colleagues and others in the social network use the same
technology (Lin and Bhattacherjee, 2008). Generally, a relatively large network size may
enable users to believe that they can communicate with others and meet people who have
similar interests and opinions easily using the same technology (Zhao and Lu, 2012). A large
network size also makes users feel that SNSs are easy to use because more facilitating
conditions could be provided (Zhao and Lu, 2012). According to the above arguments, this
study recognizes that perceived network size may make users feel that they can use social
media easily, and thus, they will be more likely to interact with other users using SNSs.
Therefore, this study proposes the following hypotheses:

\[ H6a. \] Machine interactivity is positively associated with perceived network size.

\[ H6b. \] Person interactivity is positively associated with perceived network size.

3.7 Perceived complementarity and perceived interactivity

Prior research has found that perceived complementarity is positively related to machine
interactivity and person interactivity (Zhao and Lu, 2012). In general, users tend to use an
interactive technology when they perceive high complementarity between the technology
and its complementary goods (Lin and Bhattacherjee, 2008). This is because users believe
that technology is easy to use when more complementary applications are provided and
they can use this interactive technology to post messages and respond to others’ messages
quickly (Zhao and Lu, 2012). Accordingly, this study expects that perceived
complementarity impacts machine interactivity and person interactivity significantly.
Therefore, this study proposes the following hypotheses:

\[ H7a. \] Machine interactivity is positively associated with perceived complementarity.

\[ H7b. \] Person interactivity is positively associated with perceived complementarity.

4. Research methodology

4.1 Survey administration

The data collected from Facebook users in Taiwan were used to test the proposed
model. Facebook was chosen because it is a well-known SNS in Taiwan. Its features allow
users to build and expand their own networks through the participation of social
groups. It could, therefore, be considered as a good example of a highly interactive SNS
(Cheung and Lee, 2010). To target respondents, an invitation message with a hyperlink to
an online survey questionnaire was posted on a number of bulletin board systems
and Facebook websites. The first page of the web questionnaire stated the purpose
of this study and assured confidentiality to respondents. To increase the response
rate, 15 randomly selected participants were offered a gift certificate of a convenience
store amounting to US$6 as an incentive. At the end of data collection, a total of 255 questionnaires were collected for further data analysis. Table I lists the demographics of the respondents.

4.2 Measurement development
The questionnaire was developed by adopting measures that have been validated by prior studies wherever possible. Two experts in the information systems filed were invited to assess logical consistencies, the ease of understanding, the sequence of items and the contextual relevance of the questionnaire. The comments from these experts were used to conduct several minor modifications of the wording in the questionnaire. Furthermore, an online pilot test was carried out by involving 15 undergraduate students who had experience in using SNSs. The instrument was then modified slightly in accordance with their comments. For all measures, a five-point Likert scale was used with anchors ranging from strongly disagree (1) to strongly agree (5). Appendix lists the questionnaire items of this study.

5. Data analysis
The partial least squares (PLS) method was used for data analysis because PLS has minimum restrictions on measurement scales, sample size and residual distribution (Pavlou and Fygenson, 2006). According to Chin (1998), the minimum sample size required is calculated by identifying the endogenous construct with the maximum paths leading to it.

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Number of Responses</th>
<th>Percentage of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>169</td>
</tr>
<tr>
<td>Age</td>
<td>&lt; 18</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>18–24</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>25–34</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>35–44</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>45–54</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>&gt; 55</td>
<td>11</td>
</tr>
<tr>
<td>Education</td>
<td>High school or below</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>College (2 years)</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>University</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>Graduate school or above</td>
<td>34</td>
</tr>
<tr>
<td>Frequency of use</td>
<td>Every day</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td>5–6 days</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>3–4 days</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>1–2 days</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>5</td>
</tr>
<tr>
<td>Experience in using Facebook (in year)</td>
<td>&lt; 1</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>1–3</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>4–6</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>7–9</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>&gt; 9</td>
<td>42</td>
</tr>
</tbody>
</table>

Table I. Demographics of samples
and the minimum sample size is ten times the number of paths leading to this construct. The sample size of this study is similar to that used in the literature on SNSs (Sun et al., 2016). Thus, the sample size taken in this study is adequate. Following the two-step approach (Anderson and Gerbing, 1988), the measurement model was tested first followed by the assessment of the structural model. In this study, SmartPLS Version 2.0 M3 was used for data analysis (Ringle et al., 2005).

5.1 Measurement model test
The criteria of reliability, convergent validity and discriminant validity were used to evaluate the adequacy of the measurement model. Reliability was examined based on the composite reliability (CR) values. Table II shows that all the CR values are above 0.7 (Gefen et al., 2000), indicating that the scales have adequate CR. Moreover, Table II shows that all the factor loadings exceed the threshold value of 0.7, and the average variance extracted (AVE) for every construct is greater than 0.5. The results reveal that convergent validity is acceptable (Fornell and Larcker, 1981). Finally, Table III shows that all the values of the square root of the AVE exceed the correlation shared between the construct and other constructs in the model, demonstrating the accepted discriminant validity (Fornell and Larcker, 1981).

In addition, the variance inflation factor (VIF), the effect that the other independent variables have on the standard error of a regression coefficient, was used to assess the degree of multicollinearity (Hair et al., 2006). This study conducted regression analysis by modeling continuance intention as the dependent variable and the other seven variables as independent variables. As shown in Table III, the VIFs for perceived network size, perceived complementarity, machine interactivity, person interactivity, information gratifications, emotioal gratifications, social gratifications, and continuance intention are all less than 5, indicating that multicollinearity is not a concern.

<table>
<thead>
<tr>
<th>Construct Item Mean SD Factor loading</th>
<th>CR</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived network size (PNS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNS1 4.34 0.72 0.90 0.95 0.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNS2 4.22 0.74 0.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNS3 4.34 0.75 0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNS4 4.24 0.79 0.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived complementarity (PC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC1 3.76 0.82 0.83 0.90 0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC2 3.76 0.86 0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC3 3.64 0.85 0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machine interactivity (MI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI1 3.95 0.79 0.86 0.89 0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI2 3.61 0.80 0.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI3 3.70 0.77 0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person interactivity (PI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI1 3.81 0.73 0.87 0.92 0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI2 3.60 0.78 0.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI3 3.47 0.83 0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information gratifications (IG)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UV1 4.00 0.79 0.93 0.94 0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UV2 4.00 0.74 0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UV3 3.88 0.79 0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional gratifications (EG)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HV1 3.28 0.90 0.85 0.92 0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HV2 3.44 0.95 0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HV3 3.74 0.80 0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HV4 3.67 0.81 0.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social gratifications (SG)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV1 3.54 0.85 0.86 0.90 0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV2 3.34 0.85 0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV3 3.46 0.88 0.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuance intention (CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI1 3.77 0.80 0.90 0.87 0.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI2 3.20 0.97 0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI3 3.91 0.75 0.88</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table II. Measurement items statistics
emotional gratifications and social gratifications are 1.55, 2.33, 2.79, 2.69, 2.28, 3.13 and 2.36, respectively, which are well below the suggested threshold of 10 (Hair et al., 2006). Thus, the problem of multicollinearity does not influence the results. Furthermore, because the data was collected through self-report measures, the common method bias was tested controlling for the effects of an unmeasured latent method factor based on the technique described by Podsakoff et al. (2003), following Liang et al. (2007). Thus, a latent method factor was added to the structural model to assess the method variance; the significance of the structural parameters was examined with and without the latent common method variance factor in the model. The results show that only four paths from the common method factor to the single indicator constructs are significant and the average indicators' substantive variance is 0.008, while the average method-based variance is 0.772. The results suggest that common method bias is not a significant problem in the data.

5.2 Structural model test
The theoretical model was estimated using the bootstrap approach with a sample size of 500 to generate t-values for determining the significance of paths in the structural model. The results of the structural model test are presented in Figure 2 and the hypothesis tests are summarized in Table IV. As expected, emotional gratifications and social gratifications have positive impacts on continuance intention at $p < 0.001$ and $p < 0.01$, respectively ($\beta = 0.49, 0.22; t = 6.80, 3.24$, respectively), whereas information gratifications do not exert a positive effect on continuance intention ($\beta = 0.10; t = 1.37$), indicating that $H1b$ and $H1c$ are supported, whereas $H1a$ is not supported. Furthermore, perceived network size has a significant influence on information gratifications at $p < 0.01$ ($\beta = 0.22; t = 2.81$), while it does not exert significant influences on emotional gratifications and social gratifications ($\beta = -0.06, -0.01; t = 1.35, 0.21$, respectively). The results indicate that $H2a$ is supported, whereas $H2b$ and $H2c$ are not supported. Perceived complementarity has significant influences on information gratifications and emotional gratifications at $p < 0.05$ and $p < 0.01$, respectively ($\beta = 0.21, 0.22; t = 2.46, 2.89$, respectively), while it does not impact social gratifications significantly ($\beta = 0.14; t = 1.74$). Thus, $H3a$ and $H3b$ are supported, whereas $H3c$ is not supported.

Moreover, machine interactivity affects information gratifications, emotional gratifications and social gratifications positively at $p < 0.01$, $p < 0.001$ and $p < 0.001$, respectively ($\beta = 0.24, 0.39, 0.36; t = 2.98, 5.43, 4.67$, respectively). These results reveal that $H4a, H4b$ and $H4c$ are supported. Person interactivity, as expected, influences information gratifications, emotional gratifications and social gratifications positively at $p < 0.01$.
Perceived Network Size

Information Gratifications

0.22**

Emotional Gratifications

0.24***

Continuance Intention

0.20**

$R^2=0.51$

Social Gratifications

0.30***

$R^2=0.51$

Notes: *p<0.05; **p<0.01; ***p<0.001;

Figure 2.
Results of PLS analysis

Table IV. Summary of hypotheses testing results

<table>
<thead>
<tr>
<th>Structural path</th>
<th>Path coefficient</th>
<th>t-value</th>
<th>p-value</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a: IG → CI</td>
<td>0.10</td>
<td>1.37</td>
<td>0.172</td>
<td>Not supported</td>
</tr>
<tr>
<td>H1b: EG → CI</td>
<td>0.49</td>
<td>6.80</td>
<td>0.000</td>
<td>Supported</td>
</tr>
<tr>
<td>H1c: SG → CI</td>
<td>0.22</td>
<td>3.24</td>
<td>0.001</td>
<td>Supported</td>
</tr>
<tr>
<td>H2a: PNS → IG</td>
<td>0.22</td>
<td>2.81</td>
<td>0.005</td>
<td>Supported</td>
</tr>
<tr>
<td>H2b: PNS → EG</td>
<td>-0.06</td>
<td>1.35</td>
<td>0.178</td>
<td>Not supported</td>
</tr>
<tr>
<td>H2c: PNS → SG</td>
<td>-0.01</td>
<td>0.21</td>
<td>0.834</td>
<td>Not supported</td>
</tr>
<tr>
<td>H3a: PC → IG</td>
<td>0.21</td>
<td>2.46</td>
<td>0.015</td>
<td>Supported</td>
</tr>
<tr>
<td>H3b: PC → EG</td>
<td>0.22</td>
<td>2.89</td>
<td>0.004</td>
<td>Supported</td>
</tr>
<tr>
<td>H3c: PC → SG</td>
<td>0.14</td>
<td>1.74</td>
<td>0.083</td>
<td>Not supported</td>
</tr>
<tr>
<td>H4a: MI → IG</td>
<td>0.24</td>
<td>2.98</td>
<td>0.003</td>
<td>Supported</td>
</tr>
<tr>
<td>H4b: MI → EG</td>
<td>0.39</td>
<td>5.43</td>
<td>0.000</td>
<td>Supported</td>
</tr>
<tr>
<td>H4c: MI → SG</td>
<td>0.36</td>
<td>4.67</td>
<td>0.000</td>
<td>Supported</td>
</tr>
<tr>
<td>H5a: PI → IG</td>
<td>0.20</td>
<td>2.68</td>
<td>0.006</td>
<td>Supported</td>
</tr>
<tr>
<td>H5b: PI → EG</td>
<td>0.30</td>
<td>4.09</td>
<td>0.000</td>
<td>Supported</td>
</tr>
<tr>
<td>H5c: PI → SG</td>
<td>0.30</td>
<td>3.91</td>
<td>0.000</td>
<td>Supported</td>
</tr>
<tr>
<td>H6a: PNS → MI</td>
<td>0.25</td>
<td>3.75</td>
<td>0.000</td>
<td>Supported</td>
</tr>
<tr>
<td>H6b: PNS → PI</td>
<td>0.15</td>
<td>2.64</td>
<td>0.009</td>
<td>Supported</td>
</tr>
<tr>
<td>H7a: PC → MI</td>
<td>0.53</td>
<td>7.48</td>
<td>0.000</td>
<td>Supported</td>
</tr>
<tr>
<td>H7b: PC → PI</td>
<td>0.61</td>
<td>11.15</td>
<td>0.000</td>
<td>Supported</td>
</tr>
</tbody>
</table>

Notes: IG: information gratifications; EG: emotional gratifications; SG: social gratifications; PNS: perceived network size; PC: perceived complementarity; MI: machine interactivity; PI: person interactivity; CI: continuance intention

$p<0.001$, and $p<0.001$, respectively ($\beta=0.20, 0.30, 0.30; t=2.68, 4.09, 3.91$, respectively); thus, $H3a$, $H5b$ and $H5c$ are supported. Perceived network size has significant effects on machine interactivity and person interactivity at $p<0.001$ and $p<0.01$, respectively ($\beta=0.25, 0.15; t=3.75, 2.64$, respectively). These results indicate that $H6a$ and $H6b$ are supported. Finally, perceived complementarity has significant influences on machine
interactivity and person interactivity at $p < 0.001$ and $p < 0.001$, respectively ($\beta = 0.53, 0.61$; $t = 7.48, 11.15$, respectively); thus, $H7a$ and $H7b$ are supported.

The results indicate that emotional gratifications and social gratifications account for 56 percent of the variance of continuance intention. The $R^2$ values also show that machine interactivity, person interactivity, perceived network size and perceived complementarity account for 51 percent of the variance, 60 percent of the variance and 51 percent of the variance in information gratifications, emotional gratifications and social gratifications, respectively. In addition, perceived network size and perceived complementarity account for 47 percent of the variance of machine interactivity and 49 percent of the variance of person interactivity. All the $R^2$ values exceed 10 percent, indicating the acceptable explanation power (Bock et al., 2006; Falk and Miller, 1992).

6. Discussion

6.1 Key findings

The empirical results show that emotional gratifications and social gratifications are the predictors of continuance intention. These results are in line with the findings of some previous studies (Hsu et al., 2015; Ku et al., 2013; Wei et al., 2015), indicating that the fulfillment of personal needs, such as entertainment and socialization, facilitates users’ continuance intention toward SNSs. Contrary to the expectation, the results show that information gratifications do not affect continuance intention positively. A possible reason for this is that providing a good-quality system for interpersonal communications is not costly for online service providers. Hence, it is easy for users to use another interactive technology to receive information from others. In this sense, information gratifications may not be a critical factor in motivating users to stick to a specific SNS. Additionally, this study performed an additional PLS analysis to investigate the suppressor effect by removing constructs from the models. The results indicate that information gratifications exert a significant influence on continuance intention ($\beta = 0.30$, $t = 4.53$) when emotional gratifications were removed from the model. This result indicates that the influence of information gratifications declines once emotional gratifications are taken into account. This finding, in fact, is consistent with prior studies in that SNSs are pleasure-oriented information systems, and thus, enjoyment is a strong predictor of users’ continuance intention toward SNSs (Lin and Lu, 2011).

Further, the results reveal that perceived network size influences information gratifications significantly. This finding is in line with prior studies (Lin and Bhattacherjee, 2008), indicating that users’ value will facilitate as the number of users increases. However, contrary to the expectations, perceived network size has insignificant impacts on emotional gratifications and social gratifications. This study performed an additional PLS analysis, which indicates that perceived network size exerts significant influences on emotional and social gratifications ($\beta = 0.37, 0.37$, $t = 5.48, 5.51$) by removing machine interactivity, person interactivity and perceived complementarity. One possible explanation for the findings is that SNSs are platforms that allow users to interact with others, and thus, users are more concerned about their interaction quality and how they can communicate with others using complementary goods or services. Accordingly, the influences of perceived network size on emotional gratifications and social gratifications decline when machine interactivity, person interactivity and perceived complementarity are taken into account.

On the other hand, perceived complementarity has significant effects on information gratifications, emotional gratifications and social gratifications. The findings are similar to those in Lin and Bhattacherjee (2008), Lin and Lu (2011) and Wei and Lu (2014), providing additional evidence to support that users obtain benefits from indirect network externalities in the form of product standardization and complementary services. Moreover, the results
show that machine interactivity and person interactivity have significant effects on gratifications of information, emotion and socialization. These findings are similar to the standpoint of Dennis et al. (2008) that media capability will impact communication quality, and thus, affect users' perception of performance. In addition, these findings are in line with Lu et al. (2010) and Wu and Chang (2005) that users' perception of enjoyment increases once they have intensive interpersonal interactions. Finally, the results reveal that perceived network size and perceived complementarity affect machine interactivity and person interactivity positively. These findings support the findings of Zhao and Lu (2012) that network externalities impact perceived interactivity positively.

6.2 Implications for theory
This study offers several contributions and implications for research. First, although the U&G theory is a useful theoretical lens for understanding the determinants of continuance intention, research for testing the influences of various types of gratifications on continuance intention is still scant. By dividing perceived gratifications in three dimensions (i.e. information gratifications, emotional gratifications and social gratifications), this study extends the existing literature by proving a more complete viewpoint to explain the linkages between different types of gratifications and continuance intention.

Furthermore, while the existing literature has examined the influence of network externalities on perceived benefits and gratifications (Lin and Bhattacharyee, 2008; Lin and Lu, 2011; Zhao and Lu, 2012), few studies have tested the effects of network externalities on different types of gratifications. This study indicates that perceived network size is a useful predictor of information gratification, while perceived complementarity affects information gratifications, emotional gratifications and social gratifications positively. The results of this study reveal that Facebook is a popular integrative technology that creates user benefits (i.e. information gratifications, emotional gratifications and social gratifications) from the size of personal networks and the availability of complementary goods and services. This study extends the understanding of how network externalities affect users' gratifications in the context of SNSs.

Prior studies have emphasized the importance of interactivity in the context of SNSs (Lu et al., 2010). However, our understanding on the role of interactivity in SNSs has been insufficient. By examining the effects of interactivity on perceived gratifications, this study finds that machine interactivity and person interactivity exert positive effects on the three types of gratifications. This study extends the research on perceived interactivity to the context of SNSs. In addition, similar to some Eastern counties (e.g. Thailand, Japan and South Korea), Taiwan is a collectivist society that emphasizes interdependence and interpersonal relationships among people (Hofstede et al., 2010). People in such society are more likely to build and maintain close relationships with others to obtain social support (Kim et al., 2011; Ng, 2013). The technical features of SNSs are helpful in this area (Lowry et al., 2011). The results of this study confirm that the interactive features of SNSs could facilitate interpersonal interactions to fulfill users' social and emotional needs in a collectivist society.

Another differentiating element of this study is that it may be one of the earliest studies to examine the antecedents of gratifications from social factors (i.e. network externalities) and technological attributes (i.e. perceived interactivity) simultaneously. The results of this study reveal that machine interactivity and person interactivity mediate the linkages between perceived network size, perceived complementarity and the three types of gratifications. These findings provide a possible explanation for the insignificant relationship between perceived number of members and enjoyment reported by Lin and Lu (2011). Finally, the results reveal that perceived interactivity appears to be shaped more by the availability of complementary goods and services than by the size of personal networks on Facebook.
This may imply that users are more concerned about the availability of complementary goods and services than the network size when they want to use Facebook to interact with others. This finding contributes to existing studies on SNSs by demonstrating the relationships between network externalities and perceived interactivity in the context of Facebook.

6.3 Implications for practice
This study also provides several implications for the management of SNSs to maintain users’ continuance intention. First, given that perceived interactivity plays a critical role in shaping perceived gratifications, which, in turn, impacts users’ continuance intention, managers of SNSs should deploy strategies to improve users’ perception of the SNSs’ interactive features. The management may improve the functionality of SNSs and provide new applications periodically to satisfy users’ demands to enhance their perception of machine interactivity. In addition, providing an easy-to-use website and adding new applications to SNSs may facilitate user’ willingness to interact with others through SNSs (Wei et al., 2015). Moreover, managers of SNSs may provide some incentive mechanisms for users to reinforce users’ intention to continue using SNSs. For example, they could remain users who actively share opinions and respond to others’ posts by rewarding them with virtual badges or titles (Lee and Ma, 2012).

Furthermore, the influences of network externalities on perceived interactivity and perceived gratifications demonstrate the importance of network externalities in the context of SNSs. Thus, managers of SNSs should implement some strategies to enhance network externalities. For example, managers could track the network size of SNSs frequently (Lin and Bhattacharjee, 2008) and deploy loyalty programs to provide incentives (e.g. bonuses and prizes) to enable current users to use SNSs frequently (Hsu et al., 2015; Limayem et al., 2007). Moreover, they could encourage the current users to invite their friends to use SNSs to increase the network size (Zhao and Lu, 2012). Furthermore, to increase the number of complementary goods and services, managers can track the availability of complementary goods and services (Lin and Bhattacharjee, 2008) and invite service providers (e.g. web portals or online shopping sites) to integrate their services with SNSs. They may also encourage service providers to provide other complementary services to users (Zhao and Lu, 2012).

6.4 Limitations
There are several limitations to this study. First, the data used in this study were collected from Facebook users in Taiwan. Future research should examine the findings of our study in the context of other countries (e.g. countries with highly individualistic cultures) to test the generalizability of this study. Further, future research should examine the generalizability of this study for various types of SNSs because different social networking platforms may impact users’ activities and lead to different perceptions of technology features. Second, the results may have been impacted by self-selection bias because the sample comprises only active participants. Further studies should be conducted to explore the factors that cause users to cease using SNSs. Moreover, the study defines perceived interactivity from the viewpoint of user perception. However, interactivity has different definitions from various perspectives. Future studies may test the influences of interactivity on users’ continuance intention from different perspectives to obtain greater insight into this stream of research.

7. Conclusion
This study develops and tests a theoretical model to discover the antecedents of users’ continuance intention toward SNSs. The results reveal the influences of perceived gratifications on continuance intention and how interactive features and network
externalities impact users’ gratifications. This study demonstrates the value of using the perspectives of the U&G theory, perceived interactivity and network externalities to account for the continued usage intention of SNSs. The results of this study may provide useful implications for practitioners in the SNS domain regarding the strategies to be used for enhancing people’s continued usage behavior.

References


(The Appendix follows overleaf.)
## Appendix

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Item wording and code</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Network Size (PNS)</td>
<td>I think a good number of people use Facebook (PNS1)</td>
<td>Adapted from Lin and Lu (2011)</td>
</tr>
<tr>
<td></td>
<td>I think there will be still many people joining Facebook (PNS2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I think many friends around me use Facebook (PNS3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I anticipate many friends will use Facebook in the future (PNS4)</td>
<td></td>
</tr>
<tr>
<td>Perceived complementarity (PC)</td>
<td>Facebook is well compatible with my mobile devices</td>
<td>Adapted from Lin and Lu (2011) and Zhao and Lu (2012)</td>
</tr>
<tr>
<td></td>
<td>Facebook is well compatible with the website I usually visit</td>
<td></td>
</tr>
<tr>
<td>Machine interactivity (MI)</td>
<td>It is easy for me to use Facebook to do what I want it to do (MI1)</td>
<td>Adapted from Zhao and Lu (2012) and Wu and Chang (2005)</td>
</tr>
<tr>
<td></td>
<td>The process of using Facebook is clear and understandable (MI2)</td>
<td></td>
</tr>
<tr>
<td>Person interactivity (PI)</td>
<td>Learning to use Facebook is easy for me (PI3)</td>
<td>Adapted from Zhao and Lu (2012) and Wu and Chang (2005)</td>
</tr>
<tr>
<td></td>
<td>When I am using Facebook, other users are very responsive to my posts (PI1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>When I am using Facebook, I can always count on getting a lot of responses to my posts (PI2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>When I am using Facebook, I can always count on getting responses to my posts fairly quickly (PI3)</td>
<td></td>
</tr>
<tr>
<td>Information gratifications (IG)</td>
<td>Using Facebook allows me to find the latest news and events easily (IG1)</td>
<td>Adapted from Hsu et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>I can use Facebook to find others’ messages (IG2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I can use Facebook to exchange information with others (IG3)</td>
<td></td>
</tr>
<tr>
<td>Emotional gratifications (EG)</td>
<td>Using Facebook helps me feel exciting (EG1)</td>
<td>Adapted from Hsu et al. (2015) and Zhao and Lu (2012)</td>
</tr>
<tr>
<td></td>
<td>Using Facebook makes me feel relaxed stress (EG2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Using Facebook makes me feel fun (EG3)</td>
<td></td>
</tr>
<tr>
<td>Social Gratifications (SG)</td>
<td>Using Facebook is an easy way to connect with people (SG1)</td>
<td>Adapted from Hsu et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>Using Facebook helps me make new friends (SG2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Using Facebook helps me maintain the relationship with my friends and family (SG3)</td>
<td></td>
</tr>
<tr>
<td>Continuance intention (CI)</td>
<td>I intend to keep using Facebook in the future</td>
<td>Adapted from Hsu et al. (2015), Lin and Lu (2011) and Zhao and Lu (2012)</td>
</tr>
<tr>
<td></td>
<td>If possible, I would like to continue my use of Facebook</td>
<td></td>
</tr>
<tr>
<td></td>
<td>It is likely that I will continue using Facebook in the future</td>
<td></td>
</tr>
</tbody>
</table>

**Table AI.** Survey items

---

**About the author**

Chun-Ming Chang is currently Associate Professor at the Department of International Business, Ming Chuan University, Taiwan. He holds a PhD degree in Management from National Kaohsiung First University of Science and Technology, Taiwan. His current research interests include electronic commerce and knowledge management. He has published articles in *Decision Support Systems, Information Systems Journal, International Journal of Human–Computer Studies, Behaviour and Information Technology* and others. Chun-Ming Chang can be contacted at: chunming@mail.mcu.edu.tw

---

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com
Emerald is excited to announce a recent partnership with Peerwith, a platform that provides authors with a variety of services.

The Emerald Peerwith site can be found here: https://authorservices.emeraldpublishing.com/

Peerwith connects academics seeking support for their work with a relevant expert to get their research submission-ready. Peerwith experts can help with the following: language editing, copy editing, scientific editing, translation services, statistical support, funding application support, visuals, video, publication support, literature search, peer review and indexing services. Authors post their assignments on the Peerwith site, experts provide a quote, and the fee and conditions are then agreed upon directly between the author and the expert.

While we are not, of course, guaranteeing publication upon use of Peerwith, we hope that being able to direct academics to this resource either before submission or during the peer review process will help authors further improve the quality of their papers and increase their chances of positive reviews and acceptance.

Academics with relevant expertise can sign up as an expert on the Peerwith system here: https://www.peerwith.com/services/offer
Preserving over 100 years of management research online

A lifetime investment for your institution, Emerald Backfiles will significantly enhance your library’s offering by providing access to over 125,000 articles from more than 260 journals dating back to 1898.

Visit emeraldisight.com

Get Backfiles Collections for your library

Recommend Backfiles to your librarian today.
Find out more: emeraldpublishing.com/backfilescollections
Online Information Review
The international journal of digital information research and use

Number 6

733 Editorial advisory board

734 Protecting privacy on the web: a study of HTTPS and Google Analytics implementation in academic library websites
Patrick O'Brien, Scott W.H. Young, Kenning Arlitsch and Karl Benedict

752 Image search and retrieval problems in web search engines: a case study of Persian language writing style challenges
Yaghoub Norouzi and Hoda Homavandi

768 Google as a political subject: the right to be forgotten debate 2014–2016
Linnea Lindsköld

784 Social media users' share intention and subjective well-being: an empirical study based on WeChat
Liang Ma, Xin Zhang and Xiao Yan Ding

802 Factors driving continued use of online health promotion competitions: evidence from an online weight-loss community
Xiaolong Song, Yi-Hung Liu, Jiahua Jin and Jianguo Zhao

821 Topical evolution patterns and temporal trends of microblogs on public health emergencies: an exploratory study of Ebola on Twitter and Weibo
Lu An, Chuanming Yu, Xia Lin, Tingyao Du, Liqin Zhou and Gang Li

847 Context congruity effects of online product recommendations: an eye-tracking study
Jing Luan, Zhong Yao, Yongchao Shen and Jie Xiao

864 Mood metadata on Chinese music websites: an exploratory study with user feedback
Xiao Hu, Christy W.L. Cheong, Siwei Zhang and J. Stephen Downie

880 Sexual health information-seeking behavior on a social media site: predictors of best answer selection
Yong Jeong Yi

923 What drives internet users' willingness to provide personal information?
Ching-Hsuan Yeh, Yi-Shun Wang, Shin-Jeng Lin, Timmy H. Tseng, Hsin-Hui Lin, Ying-Wei Shih and Yi-Hsuan Lai

940 Library marketing via social media: the relationships between Facebook content and user engagement in public libraries
Soohyung Joo, Namjoo Choi and Tae Hyun Bae

956 Do corporate websites’ changes reflect firms’ survival?
Desamparados Blazquez, Josep Domenech and Ana Debon

971 A comparative study of disaster management information systems
Muhammad Masood Rall, Tariq Aziz and Sarosh Hashmat Lodi

999 Understanding social networking sites continuance: the perspectives of gratifications, interactivity and network externalities
Chun-Ming Chang