Online Information Review

The international journal of digital information research and use

Number 7

1009 Editorial advisory board

1010 Timely assessment of disaster and emergency response networks in the aftermath of superstorm Sandy, 2012
   Jungwon Yeo, Louise Comfort and Kyujin Jung

1024 Predicting the quality of health web documents using their characteristics
   Melinda Oroszlányová, Carla Teixeira Lopes, Sérgio Nunes and Cristina Ribeiro

1048 A study of antecedents influencing eWOM for online lecture website: personal interactivity as moderator
   Mi sook Lee and Hongbok An

1065 To get cited or get tweeted: a study of psychological academic articles
   Yingxin Estella Ye and Jin-Cheon Na

1082 Mediating impact of fan-page engagement on social media connectedness and followers purchase intention
   Zoha Rahman, Sedigheh Moghavvemmi, Kumaran Suberamanaian, Hasmah Zanuddin and Hairul Nizam Bin Md Nasir

1106 Challenges for an SNS-based public sphere in 2016
   Kyle John Lorenzano, Miles Sarl, Colin Harrell Storm, Samuel Rhodes and Porsmita Borah

1124 The perception of e-servicescape and its influence on perceived e-shopping value and customer loyalty
   Ana Cuic Tankovic and Dragan Benazic

1146 Social media and co-creative service innovation: an empirical study
   Bijoylaxmi Sarmah, Shampy Kamboj and Jay Kandampully

1180 Analysis of user-generated comments posted during live matches of the Cricket World Cup 2015
   Snehasish Banerjee
EDITORIAL ADVISORY BOARD

Dr Noa Aharony  
Bar-Ilan University, Israel

Dr Dan Albertson  
University of Alabama, USA

Chris Armstrong  
Information Automation Ltd, UK

Professor Judit Bar-Ilan  
Bar-Ilan University, Israel

Professor Theo Bothma  
University of Pretoria, South Africa

Dr Hing Kai Chan  
University of Nottingham Ningbo China, People’s Republic of China

Dr Hsin-liang (Oliver) Chen  
Long Island University, USA

Dr Javier De Andrés  
Universidad de Oviedo, Spain

Professor Helen S. Du  
Guangdong University of Technology, People’s Republic of China

Dr Jia Tina Du  
University of South Australia, Australia

Professor Per Flensburg  
University West, Sweden

Dr Francisco Flores-Munoz  
University of Huelva, Spain

Professor Ina Fourie  
University of Pretoria, South Africa

Dr Dion Goh  
Nanyang Technological University, Singapore

Dr Tor-Morten Grønli  
Westerdals, Oslo School of Arts, Norway

Dr Jutta Haider  
Associate Professor in Information Studies, Lund University, Sweden

Dr Wu He  
Old Dominion University, USA

Professor Bruce Chien-Ta Ho  
National Chung Hsing University, Taiwan

Dr Orland Hoeber  
University of Regina, Canada

Professor Andreas Holzinger  
Graz University of Technology and Medical University Graz, Austria

Dr Hong Huang  
University of South Florida, USA

Professor Peter Jacso  
University of Hawaii at Manoa, USA

Dr Frances Johnson  
Manchester Metropolitan University, UK

Professor Wing Lam  
GlobalNxt University, Malaysia

Dr Tracey Lauriault  
Assistant Professor of Critical Media and Big Data, Carleton University, Canada

Dr Chei Sian Lee  
Nanyang Technological University, Singapore

Professor Kun Chang Lee  
Sungkyunkwan University, Republic of Korea

Dr Chen Li Liew  
Victoria University of Wellington, New Zealand

Dr Yuwei Lin  
Course Leader of BA (Hons) Media & Communications, University for the Creative Arts, UK

Dr Ying-Hsang Liu  
Charles Sturt University, Australia

Dr Carla Ruiz Mafé  
University of Valencia, Valencia, Spain

Dr M. Shaheen Majid  
Nanyang Technological University, Singapore

Professor Michele Notari  
Bern University of Teacher Education, Switzerland

Professor Keng-Boon Ooi  
UCSI University, Malaysia

Professor Robert Opoku  
Red Deer College, Canada

Dr Jahna Otterbacher  
Assistant Professor, Open University of Cyprus, Cyprus

Professor Eun Park  
McGill University, Canada

Dr Steve Probets  
Loughborough University, UK

Professor Jennifer Rowley  
Manchester Metropolitan University, UK

Professor Silvia Schiaffino  
ISISTAN – CONICET, Argentina

Professor Kalpana Shankar  
Head of School; School of Information and Communication Studies, University College Dublin, Ireland

Professor Alan D. Smith  
Robert Morris University, USA

Professor Pedro Soto Acosta  
University of Murcia, Spain

Professor Ulrike Spree  
Hamburg University of Applied Sciences, Germany

Dr David Stuart  
King’s College London, UK

Dr I-Hsien Ting  
National University of Kaohsuing, Taiwan

Dr Kristina Voigt  
Helmholtz Zentrum Munchen (German Research Center for Environmental Health), Germany

Dr Geoff Walton  
Northumbria University, UK

Dr June Wei  
University of West Florida, USA

Professor Melius Weideman  
Cape Peninsula University of Technology, South Africa

Dr Yejun Wu  
Louisiana State University, USA

Dr Tao Zhang  
Purdue University, USA
Abstract

Purpose – The purpose of this paper is to elaborate pros and cons of two coding methods: the rapid network assessment (RNA) and the manual content analysis (MCA). In particular, it focuses on the applicability of a new rapid data extraction and utilization method, which can contribute to the timely coordination of disaster and emergency response operations.

Design/methodology/approach – Utilizing the data set of textual information on the Superstorm Sandy response in 2012, retrieved from the LexisNexis Academic news archive, the two coding methods, MCA and RNA, are subjected to social network analysis.

Findings – The analysis results indicate a significant level of similarity between the data collected using these two methods. The findings indicate that the RNA method could be effectively used to extract megabytes of electronic data, characterize the emerging disaster response network and suggest timely policy implications for managers and practitioners during actual emergency response operations and coordination processes.

Originality/value – Considering the growing needs for the timely assessment of real-time disaster response systems and the emerging doubts regarding the effectiveness of the RNA method, this study contributes to uncovering the potential of the RNA method to extract relevant data from the megabytes of digitally available information. Also, this research illustrates the applicability of MCA for assessing real-time disaster response networks by comparing network analysis results from data sets built by both the RNA and the MCA.

Keywords Disaster response network analysis, Manual content analysis, Rapid network assessment, Superstorm Sandy

Introduction

Disaster and emergency management studies have employed social network analysis, which allows for superior understanding of the roles and functions of emergent disaster response systems (Jung et al., 2017; Yeo and Comfort, 2017; Kim et al., 2017; Kim and Hossain, 2013; Comfort et al., 2004, 2012; Kapucu et al., 2010; Provan and Kenis, 2007; Comfort and Haase, 2006; Comfort, 1994). Social network analysis supports operations of disaster response and recovery by identifying key agents and by highlighting any relationships, functional or

The research was funded by the National Association of Workforce Boards in 2013. Authors, also, acknowledge the researchers in the Center for Disaster Management at the University of Pittsburgh: Brian Chalfant, Jee Eun Song and Mengyao Chen for MCA data coding; and Mark Voortman, PhD for his technical support for RNA data coding.
dysfunctional, among those agents (Kim and Hossain, 2013; Comfort et al., 2012; Kapucu et al., 2010). The analysis results help managers understand emergent disaster response systems and develop better strategies for managing disasters (Jung and Song, 2015; Comfort et al., 2012; Kapucu et al., 2010; Provan and Kenis, 2007; Comfort, 1994).

Studies of disaster response networks have utilized textual data from multiple sources, such as newspapers, magazines, situation reports, whitepapers and inquiry transcripts (Yeo and Comfort, 2017; Jung and Park, 2016; Kim and Hossain, 2013; Diesner, 2012; Pfeffer and Carley, 2012; Isett et al., 2011; Moyrhh, 2009; Van Atteveldt, 2008; Comfort and Haase, 2006; Carley, 1993). Textual information has the advantage of providing a comparatively objective overview of events based on actual occurrences. It, thus, overcomes the retrospection and recollection bias prevalent in the results of traditional research tools, such as surveys, focus groups and interviews (Diesner, 2012; Pfeffer and Carley, 2012; Isett et al., 2011; Carley, 1993; Gephart, 1993).

In particular, many disaster network scholars have recently become conscious of the utility of online textual information (Yeo et al., 2018; Kim and Hossain, 2013; Isett et al., 2011). Diverse online sources, such as news websites, cloud sources and social media platforms, provide an ample amount of real-time disaster information, which could enhance the understanding of emergent disaster response systems as well as support informed decision-making (Jung and Park, 2016; Kim and Hossain, 2013; Isett et al., 2011; Carley, 1993).

Despite the advantage, there is a certain challenge in the use of online information in disaster network studies, namely, the timely extraction of valid data. After the occurrence of a disaster, the volume of information available online expands exponentially. However, methods of traditional manual content analysis (MCA) coding cannot keep up with the amount of online information generated daily (Diesner, 2012; Pfeffer and Carley, 2012; Novak and Cañas, 2008). MCA coding usually requires a great deal of time and human resources as researchers must review all the textual information, screen the focusing events, and hand code relevant information in spreadsheets to build analyzable network data sets (Diesner, 2012; Novak and Cañas, 2008; Gephart, 1993; Carley, 1993). Given the nature of the coding process, MCA usually delays timely assessment of emerging disaster response coordination (Pfeffer and Carley, 2012; Carley, 1993).

In response to the limitations of MCA, computational sociology researchers (Morstatter et al., 2013; Martin et al., 2013; Diesner, 2012; Pfeffer and Carley, 2012; Tambayong and Carley, 2012; Mihalcea and Radev, 2011) have developed a simple method for rapid online data extraction, termed the rapid network assessment (RNA) coding method. Linking the coding of network data with modeling processes, RNA helps to promptly identify and assess emerging social networks from the massive volume of digital texts, improving time and cost efficiency in data collection (Diesner, 2012; Pfeffer and Carley, 2012; Martin et al., 2013).

Setting aside rapidity, the prominent advantage of RNA, the question remains whether the output generated by RNA would be comparable to that of MCA. However, there has been few comparison studies of RNA and MCA. With respect to the growing need for timely assessment of real-time disaster response systems yet emerging doubts regarding the comparability of the RNA to MCA, this study performs a cross comparison between two methods, examining similarities and differences of the findings from the data generated by RNA and MCA. The results will help to settle remaining questions around RNA by providing comprehensive information on the advantages and disadvantages of RNA relative to MCA of online information in assessing real-time disaster response networks.

In the following sections, this study describes the coding choices for RNA and MCA; provides background information on the research context, Superstorm Sandy; presents the detailed data-modeling processes of RNA and MCA; discusses findings; and concludes with a brief summary and policy implications.
Common and distinctive coding choice of RNA and MCA

In this section, we describe coding choices in both RNA and MCA. Network coding choices include units of observation and assumptions of nodes, links and identification of other attributes (Kim et al., 2017; Yeo and Comfort, 2017; Jung and Park, 2016; Diesner, 2012; Carley, 1993, 1994). Understanding this is important because such choices determine the content of data sets, as well as affecting the results of analysis and findings (Boréus and Bergstrom, 2017; Elo et al., 2014; Diesner, 2012; Carley, 1993, 1994). We first describe shared coding choices in both methods and then examine the distinctive coding choices in each method.

Common coding choices: identification of nodes and ties

Social network theory guides the essential coding choices of both RNA and MCA. In this theory, there are two fundamental concepts: node and tie. A node is a social entity, i.e., person, group, organizations, or nation, acting within an identified event (Wasserman and Faust, 1994). Ties are social relationships, occurring between any two nodes (Wasserman and Faust, 1994). The range or types of ties vary greatly, depending on the contents of communication, interactions, transactions, or affiliations. Guided by social network theory, both RNA and MCA identify entities and relationships between them in response to a given disaster situation.

Discrete coding choices: identification of nodes and ties

Specific coding choices are made in each method in the course of the identification of social entities and relationships in emerging disaster response networks (Zhavoronkov et al., 2017; Saldaña, 2015). In MCA, researchers first download and screen all available textual information, reading through all of it that is potentially relevant to identify node populations that have the same types of social entities. In addition, MCA defines the ties found among the nodes, using identification of the actual occurrences of actions, communications, interactions, or transactions, relative to disaster response. For example, let us say that researchers have identified three documents (Doc 1, Doc 2, and Doc 3 in Figure 1) that contain information about a disaster response. If the first document (Doc 1) indicates that organization A donated money to disaster relief organizations B and C, the researcher codes organization A as an initiating node and organizations B and C as organization A’s responding organizations (partners). At the same time, in document 2 (Doc 2 in Figure 1), the researchers identify that organization B sent water to other disaster relief organizations D and C. Here, the researcher codes the ties as outgoing from organization B to organizations D and C. Finally, if document 3 (Doc 3 in Figure 1) reports that organization D provided manpower to relief organizations C, E and F, the researchers code outgoing ties from organization D to organizations C, E and F. MCA add nodes incrementally and

![Figure 1. MCA data to modeling process](image-url)
links information. The structure of the network can be identified by the subsequent extension of the bilateral relationships across all nodes. The overarching network structure cannot therefore be comprehended until all the available textual information has been reviewed. Figure 1 illustrates the coding process for MCA.

By contrast, RNA first screens social entities from each available document, using pre-structured indices (Diesner, 2012; Pfeffer and Carley, 2012); then, two-mode networks are created between each document and the actors identified within that document (Diesner, 2012; Pfeffer and Carley, 2012). This step is based on the assumption of co-occurrence, which states all actors reported in the same document are connected to each other. RNA folds all two-mode networks of actors and documents into a one-mode actor-based network (Diesner, 2012; Martin et al., 2013; Pfeffer and Carley, 2012). For example, if organization A, B and C are screened in document 1 (Doc 1 in Figure 2), then RNA assumes that organizations A, B and C worked with each other. In the same way, organizations B, E and F from document 2 (Doc 2 in Figure 2) are also identified as partners. The relationships among actors in documents 3 and 4 (Docs 3 and 4 in Figure 2) are likewise identified. Each node’s extended linkages are automatically retrieved from the folding processes, and RNA identifies the overarching disaster response network in this way (Martin et al., 2013; Diesner, 2012; Pfeffer and Carley, 2012). Where RNA is utilized, node populations and relationships among nodes easily can be identified over a relatively short period due to its assumption of co-occurrence and its automated node screening and folding. Figure 2 illustrates RNA at work.

Background for Superstorm Sandy

Brief introduction to the disaster

In late October 2012, Tropical Storm Sandy formed in the southwestern Caribbean. By the time it made landfall in Jamaica on October 24, it had increased in intensity to form a Category 1 hurricane. The storm again increased in severity as it moved toward Cuba, making a second landfall as a Category 3 hurricane on October 25. It then turned north, moving slowly over the Atlantic Ocean toward the coast of New Jersey and New York, weakening to a post-tropical storm as it made a third landfall near Brigantine, New Jersey, on the US Atlantic Coast at around 8:00 p.m. on October 29, 2012. The storm was well tracked: the US National Weather Service continuously projected simulations of the direction and strength of the storm for three days before its third landfall. This allowed emergency services and the residents of coastal communities to make preparations for the storm.

At the same time, a major cold front was moving from the Midwest toward the East Coast. It collided with Tropical Storm Sandy’s warm front, creating Superstorm Sandy;
this event unleashed a cascade of damaging effects; these went rippling through coastal communities, disrupting business operations in at least five states: New York, New Jersey, Connecticut, Delaware and Maryland (Benfield Report, 2013). Superstorm Sandy then coincided with an unusually high tide in New York City, creating a storm surge of 14 feet (4.2 meters) on Manhattan Island. This unusual coincidence of meteorological events hit one of the most densely populated regions of the East Coast, with New York City and the coastal communities of New Jersey bearing the brunt of the storm. Public agencies, business organizations and households endured significant damage and upheaval from the severe impact of this rapidly changing set of extreme events.

The total economic losses caused by Superstorm Sandy in 2012 are estimated to have reached $72bn, including approximately $30bn of insured losses and roughly $7.2bn in payments made by the US National Flood Insurance Program (Benfield Report, 2013). Roughly 60m people, many of whom lost work or suffered damage to their homes and businesses, were directly affected by Sandy, across 24 states. The economic losses to New Jersey and New York alone were estimated to be $66bn. Sandy was thus second only to Hurricane Katrina in amount of losses generated by a disaster in the USA. The need to conduct rapid data collection, analysis and interactive exchanges of information among the communities, organizations and households affected by this unusually severe storm can be seen in the evidence of the reported losses, some of which could have been reduced by informed, coordinated action.

Why Superstorm Sandy?
Due to its size and impact, Superstorm Sandy received extensive coverage by the media, leading to a gigantic amount of textual information being produced by a wide range of sources even before Sandy’s formation as a hurricane or landfall on the US East Coast. If this written information had been organized and analyzed promptly, effectiveness of disaster response could have been increased and accelerated. Yet, using MCA, it was almost impossible to collect data from the exponentially increasing mass of textual information or to conduct timely analyses of response networks working in the most affected parts of New Jersey and New York.

Although the actual disaster has now passed, rendering the processing of the available information less urgent for actual disaster response, it remains necessary to identify and test methods of data extraction, thus preparing for future disasters and crises. The mass of textual information and documents produced during the response to Superstorm Sandy will serve to provide a good platform for understanding and testing the RNA’s efficacy.

Data and methodology
Sources of data
The news articles and other documents used for this study were downloaded from the news archive LexisNexis Academic (www.lexisnexis.com/hottopics/lnacademic/), using the search query (“Hurricane Sandy” or Superstorm Sandy or Sandy Hurricane) and (New York or New Jersey) on the platform LexisNexis Smartindexing. The period for the data includes the first three months following the incident, from October 24, 2012 to January 31, 2013, which covers the preparation period, landfall and post-disaster response and recovery in the state of New York and New Jersey. A total of 1,000 articles were exported from LexisNexis Academic. After redundant and duplicate articles were eliminated, we analyzed textual information from 541 distinct articles published by 223 distinct written news sources in English.

RNA coding
A researcher and a programmer conducted the RNA of the 541 articles. Utilizing the LexisNexis Smartindexing platform, which provides a set of pre-defined textual data,
the researcher obtained a first overview list of agents, including companies, organizations and individuals participating in disaster response. Agents and the articles where these agents were identified were coded into the two-mode network data set.

The researcher identified inconsistencies in the titles of the identified agents, because different references to or labels for the same agents could lead to the creation of false distinct nodes in the network analysis, thus distorting the results of analysis (Pfeffer and Carley, 2012). To resolve errors stemming from labeling inconsistency, a thesaurus was created to convert different references to the same agents into a standardized format (Pfeffer and Carley, 2012).

Using Java, the programmer also created text-mining software to identify missing information that had not been pre-indexed by the Smartindexing system. Using the rule of the English language that proper nouns must begin with a capital letter, and adding a comment to pass the titles initially identified by Smartindexing, the Java text-mining software extracted a supplemental list of agents from the 541 articles. The additional data were again cleaned and combined with the previous two-mode network data set. The RNA work overall, from text mining to data cleaning, and to the building of the two-mode network data set, took 60 hours (12 hours/day \times 5 days), from February 8 to February 12, 2013.

**MCA coding process**

Three researchers conducted MCA on the same 1,000 articles originally exported from the Smartindexing system. The articles were divided into two groups, based on the similarity of articles: 470 unique articles, and 530 articles with some duplicates. One researcher analyzed the unique set of 470 articles, and the other two researchers analyzed the set of 530 articles. Following standard coding choices for MCA, each researcher read each article thoroughly, identified nodes and ties, and hand-coded the information on a spreadsheet. For example, researchers identified actors based on their actual involvement in the response and relationships between actors based on explicit transactions occurring during the practice of disaster response. For each document, the researchers iteratively identified and coded dyadic relationships between two organizations, an (inter)action-initiating organization and a corresponding organization. Then, the researchers combined their coding sheets. In this way, a full roster of participating organizations and ties among the organizations that participated in Superstorm Sandy response was identified. At Last, the combined data set was reviewed and cleaned by all three researchers. Overall, MCA took nearly 480 hours (12 hours/day \times 40 days), from May 1 to June 30, 2013, to complete.

**Data cleaning process**

To create a valid comparison of the two data-coding methods, each data set from each coding method was cleaned and set into comparably equivalent formats (Grimmer and Stewart, 2013; Lewis et al., 2013). During the cleaning process, first, variations in the codes for the titles of some organizational agents were identified. This discrepancy in the data was cleaned through matching and recoding the titles the affected organizational agents. Second, discrepancies in codes representing individual agents were treated. The codes for RNA separated organizational from individual agents, according to the Smartindexing categorization of the identified agents. For example, if the US Congress was represented in one article and a specific congressional representative or senator was represented in another, Congress was coded as a unique organizational agent, and the name of the representative or senator was also coded as a unique individual agent. Using the code list generated using MCA, an individual who belonged to an organization was coded as an organizational agent. Reaching a consensus on the scaling down of the data units, it was decided that codes for individual agents in the RNA data set would be scaled up and that the titles of individual agents would be recoded with the titles of the organizations if those
agents belonged to any of the identified organizations or represented certain groups. However, titles were retained for individual agents that could not be categorized into any organizations or groups.

Social network analysis
Once the data sets had been authenticated through multiple reviews by several researchers regarding the data attributes and contents, social network analysis was conducted to compare the results of analysis from each data set that was collected using the different data-collection methods. ORA (Carley, 2001–2011), a software package for network analysis developed at the Center for Computational Analysis of Social and Organizational Systems, Carnegie Mellon University, was used to analyze the two data sets. The findings from social network analysis were reviewed to identify similarities and differences in the structural and compositional characteristics of each network, as identified by the respective data-coding methods. From these results, the distinctive value of each method was identified.

Results
Network-level comparison
Network analysis was conducted to examine the general patterns of structures and the characteristics of the networks that were identified using the different coding methods. First, descriptive analyses were calculated for the two static networks. In the RNA network, 639 agents were identified that responded to the disaster, of which 46 (7 percent) were isolates, i.e., independent actors that did not interact with any other agents during the disaster response; further, 2,884 unique dyads and monads were identified. In the MCA network, 617 agents were identified that responded to the disaster, of which 177 (29 percent) were identified as isolated agents; further, 556 distinctive links were constructed by the agents identified (Table I).

The density[2], distance[3] and fragmentation[4] of the RNA and the MCA networks were then measured to determine in each network relationship patterns among agents (Table II).

As indicated by the results of descriptive analysis, the RNA network had greater density and connectedness owing to its higher number of links that that found in the MCA network. Furthermore, because the RNA network has well-interconnected agents, the average distance among them was much shorter than the average distance between agents in the MCA network. Because it had more isolated agents and found fewer links, the MCA network had much greater fragmentation and network-level values.

Each data set was graphically modeled to provide greater insight into the results of the above analysis. Figure 3 shows maps of the RNA and MCA static networks, including all agents.

<table>
<thead>
<tr>
<th>Table I.</th>
<th>Static network descriptive analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RNA</td>
</tr>
<tr>
<td>Count, node</td>
<td>639</td>
</tr>
<tr>
<td>Link count</td>
<td>2,884</td>
</tr>
<tr>
<td>Isolate count</td>
<td>46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table II.</th>
<th>Static network structure analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>RNA 0.014</td>
</tr>
<tr>
<td>Average distance</td>
<td>RNA 2.827</td>
</tr>
<tr>
<td>Fragmentation</td>
<td>RNA 0.343</td>
</tr>
<tr>
<td>Network level</td>
<td>RNA 6</td>
</tr>
</tbody>
</table>
Figure 4 gives only the core networks of the RNA and the MCA, excluding all isolated agents and groups. In both figures, the node of each agent node is colored according to its degree of centrality and its total number of direct links (where blueish = more, reddish = less), and they are sized according to their the betweenness centrality, i.e., the total number of shortest paths from all agents to all others that pass through a certain agent within the network (larger = more, smaller = fewer). Figures 3 and 4 reflect the differences between the RNA and MCA networks in terms of measures of network density, distance and fragmentation. The patterns (arrangement of node sizes and colors) for both core networks featured in Figure 4 show the structural similarities found between the networks produced by RNA and MCA.

To investigate the particular properties of the graph, centralization measures[5] were documented. The results, shown in Table III, support our claim about the overall structural
similarity between the networks produced by RNA and MCA. Even though there were actual
differences in absolute values for centralization measures, the centralization patterns in both
networks were similar. In both the RNA and MCA networks, the betweenness centralization
measures were the lowest, the degree centralization measures were slightly higher than the
betweenness centralization measures and the eigenvector centralization measures were the
highest of all. These results indicate that the structural advantages were distributed
among the agents within each network and that, overall, both networks’ agents were
connected to relatively central actors, and they shared resources or information with them
during disaster response.

Network-entity-level comparison
Analyses on the level of network entities were conducted to investigate whether the methods
of RNA and MCA congruently identified the agents participating in the Superstorm Sandy
response network. First, the percentage of agents in common in the networks, and the agents
that were distinct between them were determined. Among the agents in the RNA and MCA
networks, 371 (58 percent in the RNA network and 60 percent in the MCA network) were
found to be identical. In addition, 268 (42 percent) and 246 (40 percent) distinct, unshared
agents were identified in the RNA and MCA networks, respectively.

To determine the unique agents in each network more closely, the analysis of key
entities was conducted, using centrality measures, to examine whether the agents
had important positions in each network. The results of the analysis of key entities
indicated, also, discrepancies between the RNA and MCA in terms of the identification of
key agents with their structural roles and functions in the network of response to
Superstorm Sandy.

First, using a measure of degree centrality, the top 10 most-connected agents in each
network were identified: these were the agents that had a high degree centrality and an
immediate impact on many other agents in the disaster response they were involved in.
Despite differences in rank, seven out of ten agents were common to the RNA and MCA
networks; in addition, even though, within the top 10 most-connected agents, 6 were unique
to the top 10 of their respective system, they were all still agents that both the RNA network
and the MCA network had in common (Table IV).

<table>
<thead>
<tr>
<th>Rank</th>
<th>RNA</th>
<th>MCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>American Red Cross</td>
<td>American Red Cross</td>
</tr>
<tr>
<td>2</td>
<td>White House</td>
<td>Federal Emergency Management Agency</td>
</tr>
<tr>
<td>3</td>
<td>Federal Emergency Management Agency</td>
<td>Office of Governor of New Jersey</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>White House</td>
</tr>
<tr>
<td>5</td>
<td>Office of Governor of New Jersey</td>
<td>Congress</td>
</tr>
<tr>
<td>6</td>
<td>Office of Mayor of New York City</td>
<td>Office of Governor of New York</td>
</tr>
<tr>
<td>7</td>
<td>Congress</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Office of Governor of New York</td>
<td>Office of Mayor of New York City</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table III.</th>
<th>Network centralization measure</th>
<th>RNA</th>
<th>MCA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Network centralization, betweenness</td>
<td>0.183</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>Network centralization, eigenvector</td>
<td>0.358</td>
<td>0.642</td>
</tr>
<tr>
<td></td>
<td>Network centralization, total degree</td>
<td>0.237</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>Shared situation awareness</td>
<td>0.002</td>
<td>0.002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table IV.</th>
<th>Key entities of RNA network and MCA network by total degree centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>American Red Cross</td>
</tr>
<tr>
<td>2</td>
<td>White House</td>
</tr>
<tr>
<td>3</td>
<td>Federal Emergency Management Agency</td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Office of Governor of New Jersey</td>
</tr>
<tr>
<td>6</td>
<td>Office of Mayor of New York City</td>
</tr>
<tr>
<td>7</td>
<td>Congress</td>
</tr>
<tr>
<td>8</td>
<td>Office of Governor of New York</td>
</tr>
</tbody>
</table>
Second, using the betweenness centrality measure, the top 10 agents that were positioned to broker connections between groups were identified. These agents were the ones that could control or diffuse information among separate groups or individual agents. The results were similar to those for the analysis of the total degree centrality, regardless of the changes between the key agent lists. Seven out of ten agents were identified in common as the top gatekeeping agents. In addition, even though six agents within the top 10 gatekeeping agents for each of RNA and MCA, they were still identified as common agents in both the RNA network and the MCA networks (Table V).

Third, eigenvector centrality analysis was conducted to identify the key agents that had powerful neighbors. Through connections with powerful neighbors, agents that were eigenvector central could influence other nodes within the network. The results of analysis indicate that six common key agents enjoyed powerful neighbors; furthermore, four other agents in each network that were distinctive in terms of eigenvector central were found within the list of agents common to both the RNA and MCA networks (Table VI).

Finally, the top 10 emergent leaders in each network were examined. Emergent leaders are those that have many connections to other agents and are also engaged in multiple complex tasks requiring high levels of coordination (Carley, 2001–2011). The results of this analysis were identical to those for the top 10 most-connected agents: there were seven emergent leaders in common, with the other six emergent leaders that were not shared between the top 10s of the two networks being common to the RNA and MCA networks (Table VII).

**Conclusion**
This study investigated the RNA data-coding method. To test the method, the study compared the results of network analysis from two data sets, one developed using RNA and the other using MCA, using the same online textual information on the response to Superstorm Sandy.
Two levels of analysis were conducted on each data set: network-level analyses to investigate the overall compositional and structural patterns of the Superstorm Sandy network produced from each data-coding method, and network-entity-level analyses to examine whether and how far there was congruence between the lists of key agents identified by RNA and MCA.

From these analyses, similarities and differences were identified. The results of network analysis for RNA and MCA showed significant similarities in the general overview of the response network of Superstorm Sandy in terms of the number of agent nodes (around 60 percent), global structural patterns among those agents (patterns among the centralization measures), and the majority of key agents (centrality measures). The similarities in the results of the analyses of the two methods, combined with the efficiency of RNA, indicate that it is a more efficient method than MCA for the rapid provision of information during the initial stages of disaster response. The RNA network provided a timely overview of rapidly emerging networks of disaster response and the locations of key agents that could have aided managers in actual emergency response operations in search of information that was practically organized.

Further, differences were found between the link counts of the two networks that affected the all values of network measurement. The disparities between the link counts produced in the two networks mainly emerge from the core coding choices, including assumptions of relationship identification. The assumption of the co-occurrence of RNA methods, which automatically endows all the agents identified in the same article with relationships among one another, may overestimate the number of relationships. Overrepresentations of the link counts would result in a continual overestimation of the values of the various network measures used to examine the overall health, effectiveness and efficiency of the RNA networks. Thus, network researchers may wish to consider developing and applying new methods or assumptions for RNA to overcome its tendency to overestimate nodes and ties, as well as the consequent errors in the results of analysis that this implies.

Given the relative advantages of RNA over MCA, it is impossible to definitively conclude that either method is superior to the other for conducting evaluative analyses of disaster response networks. However, we do recommend that flexible utilization of each method, or a mixture of methods, would be desirable, in accordance with the situation. For example, during the immediate phases of disaster relief and response, the key is collecting and distributing emerging information and key resources to where it is needed rather than identifying precise information on each individual actor within an arena. The convenience and rapidity of RNA may serve the situation better than the feature of MCA. RNA is an efficient method for grasping the overall structure of a system and key actors within a short period. Therefore, RNA may be efficient for the quick review of participating agents and resource flows through their relationships during initial disaster relief and response. On the other hand, for long-term preparation for future disasters, conducted during the recovery, mitigation and prevention phases, emergency managers may need data that have a greater

<table>
<thead>
<tr>
<th>Rank</th>
<th>RNA</th>
<th>MCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>American Red Cross</td>
<td>American Red Cross</td>
</tr>
<tr>
<td>2</td>
<td>White House</td>
<td>Federal Emergency Management Agency</td>
</tr>
<tr>
<td>3</td>
<td>Federal Emergency Management Agency</td>
<td>Office of Governor of New Jersey</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>White House</td>
</tr>
<tr>
<td>5</td>
<td>Office of Governor of New Jersey</td>
<td>Congress</td>
</tr>
<tr>
<td>6</td>
<td>Office of Mayor of New York City</td>
<td>Office of Governor of New York</td>
</tr>
<tr>
<td>7</td>
<td>Congress</td>
<td>Office of Mayor of New York City</td>
</tr>
<tr>
<td>8</td>
<td>Office of Governor of New York</td>
<td></td>
</tr>
</tbody>
</table>

Table VII. Key emergent leaders in the RNA and MCA networks.
level of precision and accuracy to diagnose emergency management systems as a whole. In this context, rapidity of data accumulation is less relevant, but acquiring comprehensive and accurate data sets is the major concern. Therefore, researchers may strategically utilize either RNA or MCA, or they may take both approaches, to obtain information responding to the priorities of situations as they emerge and to assist the decision making of practitioners in the field of emergency and crisis management.

Notes
1. 4 hours/person × 3 persons.
2. Density = actual connections (links between nodes)/potential connections, potential connections = n(n − 1)/2, n = number of existing nodes.
3. Distance = 1/potential connections × (sum of shortest distance among any two nodes within a network, if one node cannot be reached by another node, the value of shortest distance between the two nodes is recorded as 0).
4. “Fragmentation = proportion of nodes in a network that is disconnected (Carley, 2001–2011).”
5. Centralization measures describe how tightly an overall connection of a network is organized around a particular focal point or how evenly nodes are sharing connections with other nodes within the network. Degree centralization is calculated based on the proportion of direct connections incident upon a central node within a network, and betweenness (i.e. the network is connected through certain nodes) and eigenvector centralization (i.e. the connectivity of the nodes depends on their neighboring or adjacent nodes) measures the proportion of connections of central nodes relative to the other nodes within the network.

References


**Further reading**


**Corresponding author**

Kyujin Jung can be contacted at: kjung1@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 quality of health web documents using their characteristics

Melinda Oroszlányová, Carla Teixeira Lopes, Sérgio Nunes and Cristina Ribeiro

Department of Informatics Engineering, University of Porto, Porto, Portugal

Abstract

Purpose – The quality of consumer-oriented health information on the web has been defined and evaluated in several studies. Usually it is based on evaluation criteria identified by the researchers and, so far, there is no agreed standard for the quality indicators to use. Based on such indicators, tools have been developed to evaluate the quality of web information. The HONcode is one of such tools. The purpose of this paper is to investigate the influence of web document features on their quality, using HONcode as ground truth, with the aim of finding whether it is possible to predict the quality of a document using its characteristics.

Design/methodology/approach – The present work uses a set of health documents and analyzes how their characteristics (e.g. web domain, last update, type, mention of places of treatment and prevention strategies) are associated with their quality. Based on these features, statistical models are built which predict whether health-related web documents have certification-level quality. Multivariate analysis is performed, using classification to estimate the probability of a document having quality given its characteristics. This approach tells us which predictors are important. Three types of full and reduced logistic regression models are built and evaluated. The first one includes every feature, without any exclusion, the second one disregards the Utilization Review Accreditation Commission variable, due to it being a quality indicator, and the third one excludes the variables related to the HONcode principles, which might also be indicators of quality. The reduced models were built with the aim to see whether they reach similar results with a smaller number of features.

Findings – The prediction models have high accuracy, even without including the characteristics of Health on the Net code principles in the models. The most informative prediction model considers characteristics that can be assessed automatically (e.g. split content, type, process of revision and place of treatment). It has an accuracy of 89 percent.

Originality/value – This paper proposes models that automatically predict whether a document has quality or not. Some of the used features (e.g. prevention, prognosis or treatment) have not yet been explicitly considered in this context. The findings of the present study may be used by search engines to promote high-quality documents. This will improve health information retrieval and may contribute to reduce the problems caused by inaccurate information.

Keywords Credibility, Prediction models, Online health information, Healh information quality

Paper type Research paper

Introduction

In the past decades, a huge amount of health-related information became available on the web. This is one of the reasons people prefer the internet as a source when seeking for information (Kim, 2009; Savolainen, 2008; Zhang et al., 2014). This implied an increased number of people being affected by online health information. Users' access to the internet and their interest in health information are both influencing aspects in health search surveys. The findings of a national survey from 2010 (Fox, 2011) about how internet users in the USA search for health information on the web report that, on a daily basis, millions of American adults have been using online resources for their health concerns. Among all adults in the USA, 74 percent went online, and 59 percent looked online for health information in 2010. These values were showed to be influenced by the health status of the...
user as well. Users' findings may have an impact on their decision making, according to the retrieved information.

Although the web is the largest source of health information available to the users, it has always been unregulated due to its distributed nature, entailing the attention to the quality of health-related information. The problem is also the increased possibility to access consumer-oriented health information, due to the rising popularity of the internet and the advance in crowd-edited websites (e.g. Wikipedia), but with potential incoherence in its quality. As summarized by Eysenbach (2002a), researchers conclude that the quality of health information varies significantly between sources.

People usually do not check the quality of health information on the internet (e.g. related to some specific medical conditions). Therefore, the goal of current research is to provide quality indicators with the perspective of helping users find trustworthy information (Eysenbach, 2002b). The quality assessment is done by introducing quality indicators and building reliable quality rating tools that can be used to improve search engine rankings. At the end of the 1990s there were already more than 45 quality-rating instruments identified (Jadad and Gagliardi, 1998), which used seals of approval for qualifying the websites. Within a few years, this number raised above 250 instruments with the focus on tools that could be used by the consumers (Bernstam, 2005). The criteria and instruments used to evaluate and rate the health-related websites can be easily accessed through the open web. There is no agreed standard of quality indicators for web-information yet, nor are the quality evaluation tools reliable in predicting high quality information (Fahy et al., 2014).

Based on the quality indicators, researchers established scoring systems for quality evaluation (e.g. HONcode, Utilization Review Accreditation Commission (URAC)), that can help searchers to encounter more reliable information. In the present work, using a previously annotated data set, composed by a set of annotated web pages and specific characteristics of web documents, we analyze the impact of several document characteristics on their quality. Our broader aim is to see whether it is possible to infer the quality of health information on the internet automatically, besides using the already reported characteristics in the literature, and without using the characteristics of HONcode and URAC criteria (Health on the Net Foundation, 2015; Utilization Review Accreditation Commission, 2015).

The extensive use of search engines (Fahy et al., 2014) makes their ranking criteria important indicators of the information reached by users. A manifold approach is needed in order to improve the quality of health information that reaches the seekers on the internet, along with better informing the searchers about the online health resources.

**Assessing the quality of health information on the web**

In the 1990s, Pallen (1995) introduced a guide to the internet that became a central concept for healthcare providers of that time, urging them to share information on health topics with the public. Later on, researchers started to focus on the information about specific medical topics available on the web, and they saw that for the users, when searching for health-related information, it might be difficult to determine the reliability of the web pages. Several types of studies that assess the quality of web content started to appear. There are studies focused on the evaluation of the quality of certain websites, studies proposing guidelines for manual evaluation of websites and studies presenting tools to automatically do this assessment. In the following subsections, we will describe some of the main initiatives in manual and automatic methods.

**Quality criteria**

Several studies have evaluated the quality of the content, and several guidelines have been published by international health organizations.
An important study in exploring the quality of health information was conducted by Impicciatore (1997), where they assessed the accuracy and completeness of website information relating to the medical topic. Accuracy is defined as the degree of concordance of the provided information with the best available evidence, and completeness means the proportion of a priori-defined elements covered by a website. Several criteria have been used to assess the quality of health-related information in the literature, termed indicators of quality. As summarized by Eysenbach (2002a), the list of the most frequently applied quality criteria contained accuracy, completeness, readability, design, disclosures and references provided. Readability means whether or not the content of a site is understandable for general consumers without medical background, and design is defined as the visual aspect of the site or layout. Disclosures refer to criteria related to how the information was presented, and references are sources of the published information. Burkell (2004) identified the following core indicators: accuracy, completeness and currency of information as well as proxy indicators: readability, design and disclosures. The author distinguishes core indicators from proxy ones. Core indicators are direct measurements of information quality on the web. On the other hand, proxy indicators predict the quality of health information indirectly. Both have their weaknesses: direct measurements require expert input, and proxy indicators might be inaccurate. In a recent review, Zhang et al. (2015) distinguished quality indicators and criteria, following a decision-making perspective based on information acquisition and processing. They grouped the criteria used in the literature into three main categories, called substance, formality and design, regarding the content. Among their findings they pointed out that criteria related to the design of the content, such as interactivity, privacy, and social and cultural appropriateness were frequently used in the studied literature. The quality indicators suggested by other studies are summarized by Zhang et al. (2015).

In Table I, we summarize the quality indicators reported by the studies we analyze. As it can be seen, there is a good accordance on the criteria that can be used to assess the quality of health contents.

### Tools for manual assessment

Based on quality indicators, researchers developed evaluation tools which derive a quality score of the given website. There are over 200 scoring systems, as reported by Bernstam (2005). The oldest, and most widely used, quality evaluation tool was created by the Health on the Net Foundation in 1995, referred to as HONcode (Boyer et al., 2011). HON recommends a code of conduct, based on eight principles summarized in Table II.

It appears in a form of an emblem on the approved websites (“Health on the Net Foundation,” 2015). Boyer et al. (2011) reported more than 7000 such certified sites. These seals of approval, used by many of the quality evaluation tools, are mostly based on the proxy indicators, as studied by Burkell (2004) and Rao et al. (2012).

Another widely used quality evaluation tool was presented by Silberg et al. (1997), in the *Journal of the American Medical Association* (JAMA) benchmarks, consisting of seven

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Completeness</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Currency</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Readability</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Design</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Disclosures</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>References</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

**Table I.**

<table>
<thead>
<tr>
<th>Quality indicators reported in our literature review</th>
</tr>
</thead>
</table>


objective assessment criteria: display of authorship, source, date of update, disclosure of ownership, sponsorship, advertising policies and conflict of interest.

The Division of Public Health and Primary Care at Oxford University, London also created a quality evaluation tool in 1998, called DISCERN, which is a questionnaire for the users consisting of 16 points (Rao et al., 2012). The URAC’s Health Web Site Accreditation program is also considered very successful in addressing the problem of health information quality, evaluating websites upon 48 quality standards (“Utilization Review Accreditation Commission,” 2015). As stated by Fahy et al. (2014), it is still unclear whether these key quality evaluation tools are accurate, due to the ever-changing character of the web. Therefore, the focus is now on considering the practical habits of users who search for health-related information on the web.

In Table III, we list the reported evaluation tools, along with their creation date and number of criteria.

As reported by Zhang et al. (2015), the quality of online consumer-oriented health information has been evaluated by researchers against predefined criteria (e.g. accuracy, credibility and readability) and using pre-existing instruments (e.g. HONcode), and it varies across domains and websites.

**Automatic assessment of health quality**

As the number of medical websites is constantly growing, automatic systems are needed to help perceiving whether a given website complies with the required principles of quality. Gaudinat et al. (2007a) proposed an automatic tool for the categorization of webpages based on document content and HONcode principles. Authors use a machine learning approach that, considering documents as vectors within a vector space, where the size of each vector is determined by the frequency of a given unit (e.g. words) in a given document, focuses on

### Table II.
The eight principles of the HONcode of Conduct for medical and health websites, and their indicators

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Authoritative</td>
<td>Indication of authorship and qualification(s) of the information provider(s)</td>
<td>Name and attainment of the author, editor or the responsible of the site</td>
</tr>
<tr>
<td>2. Complementarity</td>
<td>Declaration that the information does not replace the advice of a health professional</td>
<td>Mission, purpose and intended audience of the website, and of the organization behind the website</td>
</tr>
<tr>
<td>3. Privacy</td>
<td>Description of the privacy policy regarding the confidentiality of personal data</td>
<td>Privacy policy, copyrights</td>
</tr>
<tr>
<td>4. Attribution</td>
<td>Documentation of the sources and dates of creation and last modification</td>
<td>References, links, date of creation, last update</td>
</tr>
<tr>
<td>5. Justifiability</td>
<td>Back up all the claims with scientific evidence</td>
<td>References</td>
</tr>
<tr>
<td>6. Transparency</td>
<td>Providing accessible and clear information, and accurate contact addresses</td>
<td>Contact of the author, editor or the responsible of the site</td>
</tr>
<tr>
<td>7. Financial disclosure</td>
<td>Identification of funding sources</td>
<td>Funding sources</td>
</tr>
<tr>
<td>8. Advertising policy</td>
<td>Distinguish between editorial and advertising content</td>
<td>Advertisements, commercial intent</td>
</tr>
</tbody>
</table>

### Table III.
Evaluation tools reported in our literature review

<table>
<thead>
<tr>
<th>Evaluation tools</th>
<th>Founded</th>
<th>Number of criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>URAC</td>
<td>1990</td>
<td>48 quality standards</td>
</tr>
<tr>
<td>HONcode</td>
<td>1995</td>
<td>8 principles</td>
</tr>
<tr>
<td>JAMA benchmarks</td>
<td>1997</td>
<td>7 assessment criteria</td>
</tr>
<tr>
<td>DISCERN</td>
<td>1998</td>
<td>16 points questionnaire</td>
</tr>
</tbody>
</table>
categorizing the sentences of a document and not documents. However, it is not sensitive enough to detect the type of the context (positive vs negative) where the statement occurs. The authors have also proposed a machine learning algorithm to identify the reliability of online documents through the URL names of web pages Gaudinat et al. (2007b). Since URLs can be composed with keywords related to the HONcode principles (e.g. URL names registered for the privacy principle within the HONcode accredited database: anatome.ncl.ac.uk/tutorials/privacy.html), they were a good basis for the categorization of web pages.

Wang and Richard (2007) used a regular expression based heuristic approach for measuring information quality. The latter rule-based method of detecting technical criteria for automatic quality assessment combines structural and expression patterns, focusing on improving the detection accuracy by analyzing the structure and the content of web pages. For each criterion (e.g. author’s name, author’s credentials, author’s affiliation, reference provided, copyright notice, date of creation, date of last update, disclosure of editorial review process, disclosure of advertising policy and statement of purpose) a corresponding measurable indicator is defined with its value and expected location. After matching the indicator value with the content of a web page, the approach extracts candidate lines that may contain indicators. The overall accuracy of the rule-based method was higher than the accuracy of the direct detection method.

An adaptive learning algorithm called recursive trust labeling (RTL) was proposed by Abbasi et al. (2012) to decrease the prevalence of fake medical websites, by exploiting the complementary information utilized by graph and content-based classifiers in a dynamic manner. It was shown to be able to significantly improve fake medical website detection performance, compared to other 19 methods, by overcoming the issue of misclassification. RTL addressed this problem by adding test instances having the strongest prediction agreement across the two classifiers, to the training data. The training data set was reset during each iteration and all testing instances were reclassified in order to allow error correction.

The supervised binary classification approach of Sondhi et al. (2012) to automatically predict reliability of web pages in the medical domain used the standard HON criteria. After determining the reliability of individual web pages by a Support Vector Machine classifier, it computes the reliability of a website (not binary, but a real value) as the fraction of reliable web pages on the website. It was shown to be useful in reliability-based re-ranking and automatic website accreditation, especially when websites are a mixture of reliable and unreliable pages (e.g.: commercial websites with information about diseases and advertising products; sites where both experts and lays can edit; sites with proper or poor references).

Boyer and Dolamic (2015) discussed the usefulness of HONcode and the compliance of its principles, to judge whether the health or medical information found on the web is trustable. They remind that obtaining a HONcode certification is still based on a voluntary process, therefore some websites might be reliable, but have not requested HONcode certification. However, HONcode remains the most widely used healthcare website Code of Conduct. The authors propose further development of automated systems with criterion-matching abilities that identify good quality, trustworthy health information on the web (Boyer and Dolamic, 2015). In the present study, we discuss such criteria. Boyer et al. (2015) further studied the automated detection of health websites’ conformity to HONcode principles, focusing on the relation between the language of the document and the user’s native language (language-independent approach). In the present work, the language of the document is also one of the features we discuss in relation to do quality. HONcode certification is multilingual, therefore it allowed us to work with documents (and users) of various languages. Genova and Bender (2016) as well used the language and other concepts such as readability, layout, typography, appearance, content, risk communication, usefulness and scientific value to evaluate the quality of health information on HON approved health websites. The study of quality of health-related information on the internet of McBride et al. (2017) listed reading grade level,
authorship, references, health-care information quality certification and breadth of topic discussion as criteria that determine readability, credibility and quality. Risk factors, symptoms, prevention, treatment, prognosis, author and reading level were the most indicative features used in the detection of the quality of online health information conducted by De Groot et al. (2017). As Boyer et al. (2017) stated, automated systems for the detection of HONcode principles can identify structured elements, but they cannot understand the meaning of the content. Therefore, further improvements of the automated detection of the reliability of health websites remain important.

Methodology
The major search engines influence the users’ choice of websites by ordering their appearance in the search results. As reported by Eysenbach (2002b), most of the users visit only the top 10 websites of these ranked lists, what gives rise to consider this ordering when studying the quality of health information on the internet. The following subsections introduce the detailed description of the data set, clarifying the origin of the sample, and of the key statistical concepts used in the present analyses.

Data set
This study is based on an existing data set composed by an annotated sample of 732 health web documents, all of them being used in the current analyses. This set of documents was initially collected for a user study (Lopes and Ribeiro, 2013), where the participants performed eight web search tasks, associated with different health information seeking situations, based on questions submitted to the health category of the Yahoo! Answers service. From the list of open questions of this category, starting with the most popular one, eight questions about treatments to a symptom or disease were selected. For each question, four different search queries were defined, two in English and two in the participants’ native language. In each language, the two queries were formulated by using lay and medico-scientific terminology, respectively. Queries were built concatenating the eight symptoms or diseases with the word “treatment” with different medical terminology (lay and medico-scientific). The Google search engine was used with two collections: Google’s entire collection and Google’s indexed webpages with HONcode certification. The latter collection was filtered through Google custom search. To reduce the risk of Google learning from the previous submitted queries, it was ensured that returned links were never clicked. Further, to prevent changes in the search engine or in the HON collection, all queries were submitted within a very short time span.

For each query and collection, the top-30 results were collected. Although the top-ranked documents of Google are obviously biased towards relevance, this does not mean that the quality of the retrieved health information is good enough, as reported by Pérez-López and Pérez Roncero (2006). For the collected documents, a metadata scheme was defined, as presented in Table IV, and was later annotated using manual and automatic approaches (Sousa, 2011). The documents were assessed by a researcher and 10 percent of them were also assessed by an external health professional. The agreement rate between both assessments was measured through Cohen’s κ, where 38 indicators had concordance values greater than 0.8, 3 indicators had concordance values between 0.6 and 0.8, and 1 indicator had between 0.4 and 0.6.

Thus, the way the characteristics were annotated was, in general, well defined. The metadata schema that was used to annotate the data set is detailed in Table IV. The first column lists the document characteristics, which were categorized as related to the content, web document, entity responsible for the website and to the website. Whether the characteristic was manually (M) or automatically (A) annotated is noted in the second column. The type of variable associated with each characteristic is reported in
### Table IV.
Document characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Annotation</th>
<th>Type</th>
<th>Scale</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ari</td>
<td>A</td>
<td>C</td>
<td></td>
<td>Readability indicators</td>
</tr>
<tr>
<td>Colemanliau</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fleschkincaid</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fleschreading</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gunningfog</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smog</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smogindex</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Character of the information</td>
<td>M</td>
<td>O</td>
<td>0 – Negative</td>
<td>Their possible impact on the user (e.g. the use of “positive” or “negative” expressions)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Neutral</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 – Positive</td>
<td></td>
</tr>
<tr>
<td>Clinical cases</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td>Existence of “real” cases given by specialists</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Split content</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td>Whether the content is divided into several pages in case of html formats</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>M</td>
<td>N</td>
<td></td>
<td>Of the content (annotated according to ISO 639-1 (e.g.: en, pt))</td>
</tr>
<tr>
<td>Testimonies</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td>Of the users</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>HONcode URAC</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td>Accreditation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Last update</td>
<td>M</td>
<td>N</td>
<td></td>
<td>Date (annotated according to ISO 8,601 (YYYY-MM) and “0” if it did not exist)</td>
</tr>
<tr>
<td>References</td>
<td>M</td>
<td>C</td>
<td>0 – Not present</td>
<td>Indication of sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Commercial intent</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td>Parallel interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Advertisements</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Specific vocabulary</td>
<td>M</td>
<td>O</td>
<td>1 – Little understandable</td>
<td>Terminology</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 – Understandable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 – Completely understandable</td>
<td></td>
</tr>
<tr>
<td>Audio</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td>Type of the content</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Image</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Text</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Video</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Format</td>
<td>M</td>
<td>N</td>
<td></td>
<td>Electronic format of the document (e.g.: html, pdf)</td>
</tr>
<tr>
<td>Number of pages</td>
<td>M</td>
<td>C</td>
<td>0 – Not present</td>
<td>Of the document</td>
</tr>
<tr>
<td>Scientific Publication</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td>Documents from a publication of scientific character (e.g.: scientific papers)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Epidemiologic data</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td>Type of medical information contained in the document</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Pathologic definition</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 – Present</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Annotation</th>
<th>Type</th>
<th>Scale</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indication of health professionals</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present / 1 – Present</td>
<td></td>
</tr>
<tr>
<td>Place of treatment</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present / 1 – Present</td>
<td></td>
</tr>
<tr>
<td>Prevention</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present / 1 – Present</td>
<td></td>
</tr>
<tr>
<td>Prognosis</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present / 1 – Present</td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>M</td>
<td>ND</td>
<td>0 – Not present / 1 – Present</td>
<td></td>
</tr>
</tbody>
</table>

**Web document characteristics**

| Main type of the content               | M          | N    |  | Article, informative, message, questionnaire, comment, academic work         |

**Responsible entity characteristics**

| Contacts of the author                 | M          | ND   | 0 – Not present / 1 – Present | Author                                                                 |
| Attainment of the author               | M          | O    | 0 – No attainment was mentioned / 1 – Attainment in the health domain / 2 – Attainment in another area |                                                                 |
| Name of the author                     | M          | ND   | 0 – Not present / 1 – Present | Responsible for website                                                                 |
| Contacts of the webmaster              | M          | ND   | 0 – Not present / 1 – Present |                                                                 |
| Name of the webmaster                  | M          | ND   | 0 – Not present / 1 – Present | Reputiation                                                                 |
| Scientific nature                      | M          | ND   | 0 – Not present / 1 – Present |                                                                 |
| Governmental nature                    | M          | ND   | 0 – Not present / 1 – Present |                                                                 |

**Website characteristics**

| Objective                              | M          | ND   | 0 – Not present / 1 – Present | Mission                                                                 |
| Domain                                 | M          | N    |                           | Name (e.g.: .com, .gov, .edu)                                                                 |
| Type                                   | M          | N    |                           | Collaborative, personal, institutional-scientific, institutional-not scientific, electronic commerce |

| Copyrights                             | M          | ND   | 0 – Not present / 1 – Present | Disclosure                                                                 |
| Privacy policy                         | M          | ND   | 0 – Not present / 1 – Present |                                                                 |
| Team of revision                       | M          | ND   | 0 – Not present / 1 – Present | Editorial review                                                                 |
| Process of revision                    | M          | ND   | 0 – Not present / 1 – Present |                                                                 |

**Note:** The ones marked with italic are related to HONCode principles.

Table IV.

---

the third column with a C for continuous variables, an O for ordinal variables, an N for nominal (factor) variables and a ND for nominal-dichotomous. The fourth column contains the description of the used scales, and the last column lists the description of the characteristics.
In this study, HON certification was used as the ground truth judgment of quality, i.e., if a document had the HONcode seal it was considered to have good quality (Price and Hersch, 1999; Berland et al., 2001; Lopes and Ribeiro, 2010). All documents were assessed and labeled “1” (present) or “0” (not present).

Some of the document characteristics marked with italic in Table IV are related to HONCode principles (see Table II). The details of these relations are indicated in Table V, including the types of the relations in the last two columns.

In a previous work of the authors (Lopes and Ribeiro, 2010), several document characteristics have been analyzed, and shown to be important for determining the relation with information quality. The most relevant characteristics are listed in Table IV, and they form the basis of the present study.

In a previous study (Oroszlányová et al., 2015) we observed features of web documents other than the ones used in the HONcode principles, and found that some were associated with the quality of the document (such as split content, video, image, advertisement, English language, indication of place of treatment, prevention, prognosis, pathologic definition and clinical cases).

### Statistical analysis

In this work, the HON certification is used as the ground truth judgment of quality. The goal of HONcode certification is to assure that the health or medical information found on the web is trustable. Thus, if a document has the HONcode seal, it is considered to have good quality (Price and Hersch, 1999; Berland et al., 2001; Lopes and Ribeiro, 2010). Although it would be preferable to have health professionals assessing every document in the data set, the high number of documents and the difficulties in recruiting medical doctors or other health professionals for this task, led us to opt for using HONcode certification as ground truth.

In Section "Multivariate analysis", we analyze the variables from our data set (considering more than two variables), and we build a forecasting model. We want to predict whether a document has quality or not, based on its characteristics. For this purpose, we select the variables that build up a model with the best fit to our data. The least absolute shrinkage and selection operator (lasso) selects the best subset of predictors by shrinking the regression coefficients toward zero, and estimates their coefficients (James et al., 2013). It uses logistic regression, to model the probability of having a quality document given its characteristics. Alternative classification approaches, like linear discriminant analysis are

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>HON principle</th>
<th>Necessary?</th>
<th>Sufficient?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Content characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Last update</td>
<td>Attribution</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>References</td>
<td>Attribution</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Responsible entity characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contacts of the author</td>
<td>Transparency</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Attainment of the author</td>
<td>Authoritative</td>
<td>Almost</td>
<td></td>
</tr>
<tr>
<td>Name of the author</td>
<td>Authoritative</td>
<td></td>
<td>Almost</td>
</tr>
<tr>
<td>Contacts of the webmaster</td>
<td>Transparency</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Name of the webmaster</td>
<td>Authoritative</td>
<td></td>
<td>Almost</td>
</tr>
<tr>
<td><strong>Website characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective</td>
<td>Complementarity</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Privacy policy</td>
<td>Privacy</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Team of revision</td>
<td>Authoritative</td>
<td></td>
<td>Almost</td>
</tr>
</tbody>
</table>

Table V. Document characteristics and their relation with HON principles
Predicting the quality of health web documents

less direct to estimating this probability, and they do not tell us which predictors are important, like it is the case with $K$-nearest neighbors (James et al., 2013). The probability of having a quality document given its characteristics can be written as $\Pr(\text{HON.code} = \text{yes} | \text{characteristics})$, where the probability values range between 0 and 1: $p(\text{characteristics}) \in [0, 1]$. The above relation can be expressed in a general form as $p(X) = \Pr(Y=1|X)$. Since our model has a binomial distribution with $Y \in \{0,1\}$, and has multiple predictors, $X = (X_1, \ldots, X_p)$, we use the logistic function:

$$\log \left( \frac{p(X)}{1-p(X)} \right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p. \tag{1}$$

By definition, this can be also written as:

$$p(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p}}. \tag{2}$$

called log-odds (or logit) transformation. The lasso regression coefficient estimates are then the values that minimize the quantity composed by the squared residuals sum and the shrinkage penalty as follows:

$$\sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j|. \tag{3}$$

Selecting a good value of the tuning parameter $\lambda \geq 0$ is crucial, because it controls the relative impact of the above two terms on the regression coefficient estimates. In the shrinkage penalty, the $\ell_1$ norm of a coefficient vector $\beta$ is given by $||\beta||_1 = \sum |\beta_j|$. The lasso performs variable selection according to the $\ell_1$ penalty forcing some of the coefficient estimates to be exactly equal to zero when $\lambda$ is sufficiently large. For choosing the best $\lambda$, we use ten-fold cross-validation (CV) (James et al., 2013). For this effect of the $\ell_1$ penalty, we use lasso over other regularization methods, like the well-known ridge regression what only shrinks the coefficient estimates toward zero, but cannot exclude any predictor variable from the model (James et al., 2013). Thus, lasso can yield a more interpretable model.

The lasso fit does not carry information on statistical significance, but the resulting coefficients can be ranked by their values, which are indicators of effect size. Positive coefficients are equivalent to positive logs of the odds ratio and indicate that the variables are associated with higher risk of an event (in our context, documents having quality), and vice versa for negative coefficients. However, the importance of the effects also depends on what the variables stand for and on subjective knowledge. After the variable selection done by lasso, we include the chosen characteristics in the multiple logistic regression model, and estimate the classification and predictive accuracy (James et al., 2013). The classification accuracy measures how well the classification test predicts the two categories: documents having quality, and documents not having quality (James et al., 2013). The accuracy of the prediction model (predictive accuracy) is assessed using leave-one-out cross-validation (LOOCV) (James et al., 2013). The LOOCV error rate in our classification setting is estimated by averaging the $n$ misclassified observations, and it takes the form:

$$CV_n = \frac{1}{n} \sum_{i=1}^{n} I(y_i \neq \hat{y}_i). \tag{4}$$

The LOOCV approach uses a single observation in the validation set, using the remaining observations to form the training set, and the prediction $\hat{y}_i$ is made for the single observation.
We build a second model, which contains only the significant variables from the full model, and compare the LOOCV estimates of prediction (or test) errors for the two models.

**Multivariate analysis**

In this section, we analyze if and how the characteristics of documents are useful to predict their quality, applying multiple logistic regression. In the three following subsections, we build and evaluate three types of full and reduced logistic regression models. The first one includes every feature, without any exclusion, the second one disregards the URAC variable, due to it being a quality indicator, and the third one excludes the variables related to the HONcode principles, which might also be indicators of quality. The reduced models were built with the aim to see whether they reach similar results with a smaller number of features.

**Full model without exclusions**

Our first model considers all variables, including the URAC certification, and variables related to the HONcode principles. We start by fitting a lasso model on the training set for an automatically selected range of \(\lambda\) values in order to see the path of the variables coefficients and visualize it by plotting the coefficients against the deviance explained (section “Coefficient plot”). Then we choose the “best” tuning parameter \(\lambda\) using CV (section “Choosing the tuning parameter \(\lambda\)”), and use it to fit the lasso model on the full dataset (section “Lasso model”). With the variables selected by the lasso model, we fit a multiple logistic regression model with the variables selected by the lasso model (section “Logistic regression model”), and evaluate the results (section “Evaluation”).

**Coefficient plot**

After fitting a lasso model on the training set for an automatically selected range of \(\lambda\) values, we plot the coefficients in order to see their path (Figure 1). This coefficient plot displays the path of the variables coefficients against the deviance explained. The upper axis indicates the number of non-zero coefficients (i.e. degrees of freedom) at the current percent deviance explained, which is expected to change sufficiently from one \(\lambda\) to the next in the lasso.
Each curve on Figure 1 corresponds to a variable, and we can see that some of the coefficients will be equal to zero, due to the forcing effect of the $\ell_1$ penalty when the tuning parameter $\lambda$ is sufficiently large. The curves are annotated by numbers, which refer to the position of the variables in lasso. For example, the (first) curve annotated with 15 refers to URAC certification, and the (last) curve annotated with 24 refers to the variable that indicates whether a document is a scientific publication.

We observe that, at first, lasso results in a model containing only the variable URAC. Then the other predictor variables enter the model as we move from left to right in the figure. We see in the plot that for models above 17 degrees of freedom, many curves begin to separate from zero, as well as the intensive increases in coefficients with only modest increases in percent deviance explained. The lasso can produce a model involving any number of variables depending on the value of $\lambda$, e.g., if $\lambda = 0.163$, the model would include 1 variable, or if $\lambda = 0.040$ it would contain 11 variables. Figure 1 helps to determine the strongest predictors, showing that variables that enter the model early are the most predictive and variables that enter the model later are less important.

Choosing the tuning parameter $\lambda$

Second, to choose a value for the tuning parameter $\lambda$, we applied ten-fold CV based on the misclassification error as visualized on Figure 2. The dotted line indicates the CV-curve, and error bars indicate the standard deviation along the $\lambda$ sequence. The vertical dotted lines indicate the optimal $\lambda$'s. The first gives the value of $\lambda$ at which the minimal mean cross-validated error is achieved ($\lambda_{\text{min}} = 0.025$), and the second the value of the largest $\lambda$ at which the error is within one standard error of the minimum $\lambda_{\text{1se}} = 0.033$), i.e. the most regularized model.

To apply the ten-fold CV, we first chose a grid of $\lambda$ values, and computed the CV error for each value. Then, the tuning parameter was selected as the value with the smallest CV error $\lambda_{\text{min}} = 0.025$.

Lasso model

Using the above information, and the list of 81 potential predictor variables, we built a model predicting the quality of web documents. The lasso, with the minimal $\lambda$ chosen by CV...
(\lambda_{\text{min}} = 0.025), yielded a prediction model containing 17 variables (the description of the variables is in Table IV). The resulting coefficient estimates are summarized in Table VI. The first column contains the category of the feature, where C, D, R and S refer to content, web document, responsible entity and website, respectively. The second column contains the names of the variables and the third column the corresponding coefficients estimated by the lasso. Since 17 variables are included in this model, 64 of the 81 coefficient estimates were zero.

We observed that the variables URAC and last update less than one year before the documents were retrieved ("Last update (< 1 yr ago") have the largest coefficient estimates. The negative coefficients indicate that documents with these characteristics are less likely to have quality than the documents without these, for fixed values of the remaining variables. The negative coefficient found in the scientific publication value surprised us because we expected a direct relationship between quality and this type of publication.

Logistic regression model

After using lasso, variables from Table VI were added to the multiple logistic regression model summarized in Table VII. The first column contains the category of the feature, where C, D, R and S refer to content, web document, responsible entity and website, respectively. The second column contains the names of the variables and the third column their corresponding estimated coefficients. Variables with positive coefficients have positive effect on the likelihood of documents HON certification, and it is less likely that certified documents possess the characteristics having negative coefficients. The fourth column lists the standard error when assessing the accuracy of the coefficient estimates. The fifth column contains the z-statistic where its large (absolute) value indicates evidence against the null hypothesis of the coefficients being equal to zero. The last column lists the corresponding p-values. Our regression model was further verified by LOOCV, and its results are reported in the last row of Table VII.

The p-values associated with the variables URAC, Lastupdate (< 1 yr ago), Scientific publication and Scientific nature, marked with italic in Table VII, are statistically significant at \( \alpha = 0.05 \) in the full regression model.

<table>
<thead>
<tr>
<th>Cat.</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>URAC</td>
<td>2.504</td>
</tr>
<tr>
<td>C</td>
<td>Last update (&lt; 1 yr ago)</td>
<td>1.072</td>
</tr>
<tr>
<td>C</td>
<td>Split content</td>
<td>0.693</td>
</tr>
<tr>
<td>S</td>
<td>Privacy policy</td>
<td>0.629</td>
</tr>
<tr>
<td>C</td>
<td>Video</td>
<td>0.557</td>
</tr>
<tr>
<td>S</td>
<td>Process of revision</td>
<td>0.516</td>
</tr>
<tr>
<td>C</td>
<td>Last update (3–4 yrs ago)</td>
<td>0.439</td>
</tr>
<tr>
<td>S</td>
<td>Copyrights</td>
<td>0.266</td>
</tr>
<tr>
<td>R</td>
<td>Scientific nature</td>
<td>0.231</td>
</tr>
<tr>
<td>S</td>
<td>Team of revision</td>
<td>0.196</td>
</tr>
<tr>
<td>C</td>
<td>Advertisements</td>
<td>0.169</td>
</tr>
<tr>
<td>C</td>
<td>Place of treatment</td>
<td>0.152</td>
</tr>
<tr>
<td>C</td>
<td>Last update (2–3 yrs ago)</td>
<td>0.151</td>
</tr>
<tr>
<td>C</td>
<td>Colemanlau</td>
<td>0.083</td>
</tr>
<tr>
<td>S</td>
<td>Type (institutional-not scientific)</td>
<td>0.026</td>
</tr>
<tr>
<td>C</td>
<td>Scientific publication</td>
<td>−0.393</td>
</tr>
<tr>
<td>S</td>
<td>Type (collaborative)</td>
<td>−0.647</td>
</tr>
</tbody>
</table>

Table VI.
The 17 variables of the lasso model, and their coefficient estimates
Evaluation

We fitted a logistic regression model in order to predict documents’ quality using the above 17 variables. To better assess the accuracy of this model, we fitted it using part of the data, and then examined how well it predicts the held-out data (test data set). For this purpose, the observations were randomly split in half. We computed the probabilities of the document having quality, using our test data set. Given these predictions, we determined how many observations were correctly or incorrectly classified, based on the confusion matrix (James et al., 2013). Our logistic regression model has an accuracy of 89.0 percent, a specificity (true negative rate) of 92.1 percent and sensitivity (true positive rate) of 69.8 percent. The LOOCV estimate of prediction error from Table VII is very low (7.4 percent), meaning that the regression model is of high accuracy. That is, the predictors in our model explain more than 90 percent of the variance in quality.

According to the model in Table VII, using Equation (2) from Section “Statistical analysis”, the probability of having a quality document with almost all the above characteristics and a Colemanliau of 1.5 is given by:

$$P(Y=1) = \frac{e^{-5.637 + 3.980 \times 1 + 2.099 \times 1 - 2.822 \times 1 + 1.296 \times 1 + 1.685 \times 0 + 1.067 \times 1 + 0.967 \times 1 + 0.944 \times 1 + 0.966 \times 1 + 0.142 \times 1.5 + 0.766 \times 0 + 0.887 \times 0 + 0.221 \times 0 - 0.436 \times 1 + 0.049 \times 1 - 16.883 \times 0}{1 + e^{-5.637 + 3.980 \times 1 + 2.099 \times 1 - 2.822 \times 1 + 1.296 \times 1 + 1.685 \times 0 + 1.067 \times 1 + 0.967 \times 1 + 0.944 \times 1 + 0.966 \times 1 + 0.142 \times 1.5 + 0.766 \times 0 + 0.887 \times 0 + 0.221 \times 0 - 0.436 \times 1 + 0.049 \times 1 - 16.883 \times 0}} = 0.9897.$$  

where the value of Colemanliau (1.5) is the average value amongst the documents, and it corresponds to the reading level grade of a person (e.g. a score of 8 means that a US student in the eighth grade (14-year-old) can understand the document). Thus, the documents are in average easily readable, as indicated by the low mean of 1.5.
Similarly, the probability of having a quality document without the above characteristics and a Colemanliau of 1.5 is:

\[
Pr(\text{HON.code} = \text{yes}, \text{URAC} = \text{no}, \text{Last update (< 1 yr ago)} = \text{no}, \text{Scientific publication} = \text{no}, \text{Scientific nature} = \text{no}, \text{Place of treatment} = \text{no}, \text{Last update (2–3 yrs ago)} = \text{no}, \text{Privacy policy} = \text{no}, \text{Split content} = \text{no}, \text{Video} = \text{no}, \text{Collomanliau} = 1.5, \\
\text{Copyrights} = \text{no}, \text{Last update (3–4 yrs ago)} = \text{no}, \text{Process of revision} = \text{no}, \text{Advertisements} = \text{no}, \text{Team of revision} = \text{no}, \text{Type (institutional-not scientific)} = \text{no}, \text{Type (collaborative)} = \text{no}) = \\
\frac{e^{-(5.637 + 3.980 \times 0 + 2.099 \times 0 - 2.822 \times 0 + 1.286 \times 0 + 1.685 \times 0 + 1.087 \times 0 + 0.967 \times 0 + 0.944 \times 0 + 0.966 \times 0 + 0.142 \times 1.5 +}}{1 + e^{-(5.637 + 3.980 \times 0 + 2.099 \times 0 - 2.822 \times 0 + 1.286 \times 0 + 1.685 \times 0 + 1.087 \times 0 + 0.967 \times 0 + 0.944 \times 0 + 0.966 \times 0 + 0.142 \times 1.5 +}}
\]

\[
= \frac{0.766 \times 0 + 0.608 \times 0 + 0.087 \times 0 + 0.221 \times 0 - 0.436 \times 0 + 0.049 \times 0 - 16.883 \times 0}{0.766 \times 0 + 0.608 \times 0 + 0.087 \times 0 + 0.221 \times 0 - 0.436 \times 0 + 0.049 \times 0 - 16.883 \times 0} = 0.0044.
\]

The values obtained from this example indicate that the probability of having a quality document with the above characteristics is very high, and the probability of having a quality document without them is very low.

Reduced model without exclusions

We built a second model, including only the statistically significant variables from Table VII. Table VIII shows the coefficient estimates for this logistic regression model. In this second model, all the variables remained significant. The \(p\)-values associated with the above variables indicate that each of these characteristics is associated with quality. Compared to what happened in the full model, the coefficient for the variable scientific publication is negative, indicating that scientific publications are less likely to have quality than non-scientific publications, for fixed values of the other variables in the model. This probably happens because it is not common to have scientific publications that are HONcode certified, the criterion we used as ground truth. In fact, the proportion of scientific publications in the data set was 11.1 percent (i.e. 81 documents), what is quite low, and out of these 81 documents only 6 were HONcode certified (0.82 percent of all documents).

Our logistic regression has an accuracy of 88.4 percent, a specificity of 90.6 percent and sensitivity of 71.7 percent. The LOOCV estimate of prediction (or test) error for this model is slightly higher than the error estimate for the full regression model in Table VII.

Full model without URAC

In the previous models, we have considered the URAC variable, that is, if a document has or not the URAC certification. Since URAC is, itself, an indicator of quality, we decided to build a model without this variable. It is important to have a model that does not depend on other type of medical certification that by itself assures the quality of the document.

| Cat       | Variable                             | Estimate | SE  | z-score | Pr( > |z|)   |
|-----------|--------------------------------------|----------|-----|---------|-------|
|           | (Intercept)                          | -3.741   | 0.432 | -8.656  | 4.91E-18 |
| C         | URAC                                 | 5.570    | 0.897 | 6.212   | 5.22E-10 |
| C         | Last update (< 1 yr ago)             | 2.098    | 0.456 | 4.604   | 4.14E-06 |
| R         | Scientific nature                    | 2.076    | 0.466 | 4.455   | 8.41E-06 |
| C         | Scientific publication               | -2.623   | 1.055 | -2.486  | 1.29E-02 |

**Note:** \(p\)-values with italics are statistically significant.
Fitting the lasso model, we found the minimal $\lambda$ (0.036) and the following 15 variables summarized in Table IX. The first column contains the category of the feature, where C, D, R and S refer to content, web document, responsible entity and website, respectively. The second column contains the name of the variables and the third column the corresponding coefficients estimated by the lasso.

We observed that the variables team of revision and last update less than one year before the documents were retrieved have the largest coefficient estimates. This is natural, since the presence of team of revision and of the last update date are HONcode principles. Once again, we observed a negative coefficient for the scientific publication variable.

**Logistic regression model**

After adding the above variables to the multiple regression model (summarized in Table X), six of them were significant at $\alpha = 0.05$, including the last update less than one year and...
between two and three years before the date of the document retrieval, scientific publication, place of the treatment, domain “.edu”, and split content.

**Evaluation**

Our logistic regression has an accuracy of 86.8 percent, a specificity of 92.7 percent and sensitivity of 56.9 percent. The prediction error of 0.084 means that 8.4 percent of the cases would be misclassified by using our model.

According to the model in Table X, using Equation (2) from section “Statistical analysis”, the probability of having a quality document with all the above characteristics on and a Colemanliau of 1.5 is given by:

\[ \Pr(\text{HON.code} = \text{yes} | \text{Type (collaborative)} = \text{no}, \text{Last update (2–3 yrs.ago)} = \text{no}, \text{Last update (3–4 yrs. ago)} = \text{no}, \text{Colemanliau} = 1.5, \text{all.the.other.variables.from.Table X} = \text{yes}) = 0.9899. \]

Similarly, \[ \Pr(\text{HON.code} = \text{yes} | \text{Colemanliau} = 1.5, \text{all.the.other.variables.from.Table X} = \text{no}) = 0.0047. \]

As expected, the error rates in a model without URAC are higher.

**Reduced model without URAC**

The reduced model, containing only the significant variables from the full model, is summarized in Table XI. All six variables remained significant.

Our logistic regression has an accuracy of 85.2 percent, a specificity of 87.9 percent and sensitivity of 58.6 percent. The LOOCV estimate of prediction error for this model is slightly higher than the error estimate for the full regression model in Table X.

**Full model without HON characteristics**

Some of the variables included in the previous models specify characteristics that documents must have to be HON certified, that is, are characteristics included in HONcode principles (described previously in Table II). Since these variables are, at least partially, indicators of quality, we wanted to analyze how well we can predict quality without these characteristics.

**Lasso model**

Fitting the lasso model, we found the minimal \( \lambda \) (0.013), and the 22 variables listed in Table XII.

**Logistic regression model**

We built a multiple regression model including these variables, summarized in Table XIII, where five of them were statistically significant at \( \alpha = 0.05 \), namely, the process of revision, split content, scientific publication, place of the treatment, and domain “.edu”.

| Cat. | Variable                        | Estimate | SE  | \( z \)-score | \( \Pr(>|z|) \) |
|------|---------------------------------|----------|-----|---------------|----------------|
|      | (Intercept)                     | −3.007   | 0.316 | −9.513        | 1.85E-21       |
| C    | Split content                   | 0.606    | 0.505 | 4.564         | 5.01E-06       |
| C    | Last update (2–3 yrs.ago)      | 2.239    | 0.427 | 3.835         | 0.002          |
| C    | Last update (3–4 yrs.ago)      | 1.672    | 0.606 | 0.699         | 3.20E-04       |
| S    | Domain (.edu)                   | 3.252    | 0.505 | 3.074         | 0.002          |
| C    | Scientific publication          | 2.239    | 0.932 | −2.402        | 0.016          |

LOOCV estimate of prediction error: \( 0.107 \)

**Note:** \( p \)-values with italics are statistically significant.
<table>
<thead>
<tr>
<th>Cat.</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Intercept)</td>
<td>−5.221</td>
</tr>
<tr>
<td>S</td>
<td>Process of revision</td>
<td>1.885</td>
</tr>
<tr>
<td></td>
<td>Domain (.edu)</td>
<td>1.954</td>
</tr>
<tr>
<td>C</td>
<td>Scientific publication</td>
<td>−2.733</td>
</tr>
<tr>
<td>C</td>
<td>Split content</td>
<td>1.677</td>
</tr>
<tr>
<td>S</td>
<td>Domain (.edu)</td>
<td>2.823</td>
</tr>
<tr>
<td>R</td>
<td>Place of treatment</td>
<td>1.770</td>
</tr>
<tr>
<td>C</td>
<td>Colemanliau</td>
<td>0.022</td>
</tr>
<tr>
<td>C</td>
<td>Commercial intent</td>
<td>1.222</td>
</tr>
<tr>
<td>S</td>
<td>Governmental nature</td>
<td>0.255</td>
</tr>
<tr>
<td>C</td>
<td>Ari</td>
<td>0.610</td>
</tr>
<tr>
<td>C</td>
<td>Testimonies</td>
<td>−0.614</td>
</tr>
<tr>
<td>S</td>
<td>Copyrights</td>
<td>1.222</td>
</tr>
<tr>
<td>C</td>
<td>Commercial intent</td>
<td>0.610</td>
</tr>
<tr>
<td>R</td>
<td>Scientific nature</td>
<td>0.483</td>
</tr>
<tr>
<td>S</td>
<td>Type (institutional-not scientific)</td>
<td>0.416</td>
</tr>
<tr>
<td>S</td>
<td>Domain (.gov)</td>
<td>−0.700</td>
</tr>
<tr>
<td>C</td>
<td>Video</td>
<td>−0.180</td>
</tr>
<tr>
<td>C</td>
<td>Prevention</td>
<td>0.089</td>
</tr>
<tr>
<td>S</td>
<td>Domain (.uk)</td>
<td>−0.242</td>
</tr>
<tr>
<td>C</td>
<td>Prevention</td>
<td>−0.070</td>
</tr>
<tr>
<td>C</td>
<td>Advertisements</td>
<td>0.085</td>
</tr>
<tr>
<td>S</td>
<td>Type (collaborative)</td>
<td>−16.231</td>
</tr>
<tr>
<td>S</td>
<td>Type (electronic commerce)</td>
<td>−16.870</td>
</tr>
<tr>
<td>S</td>
<td>Domain (.sa)</td>
<td>−18.893</td>
</tr>
</tbody>
</table>

Table XII. The 22 variables of the lasso model, and their coefficient estimates

| Cat. | Variable                                   | Estimate | SE  | z-score | Pr( > |z|) |
|------|-------------------------------------------|----------|-----|---------|------|
|      | (Intercept)                               | −5.221   | 1.076| −4.854  | 1.21E-06 |
| S    | Process of revision                       | 1.885    | 0.586| 3.215   | 0.001 |
|      | Domain (.edu)                             | 1.954    | 0.661| 2.955   | 0.002 |
| C    | Scientific publication                    | −2.733   | 0.931| −2.934  | 0.003 |
| C    | Split content                             | 1.677    | 0.641| 2.615   | 0.009 |
| S    | Domain (.edu)                             | 2.823    | 1.242| 2.274   | 0.023 |
| R    | Place of treatment                        | 1.770    | 0.922| 1.919   | 0.055 |
| C    | Colemanliau                               | 0.022    | 0.013| 1.675   | 0.094 |
| C    | Copyrights                                | 1.222    | 0.846| 1.444   | 0.149 |
| C    | Testimonies                               | −0.614   | 0.595| −1.032  | 0.302 |
| C    | Commercial intent                         | 0.610    | 0.594| 1.027   | 0.304 |
| R    | Scientific nature                         | 0.483    | 0.544| 0.889   | 0.374 |
| S    | Type (institutional-not scientific)       | 0.416    | 0.586| 0.710   | 0.478 |
| S    | Domain (.gov)                             | −0.700   | 1.408| −0.497  | 0.619 |
| C    | Video                                     | −0.180   | 0.953| −0.189  | 0.850 |
| C    | Prevention                                | 0.089    | 0.495| 0.180   | 0.857 |
| S    | Domain (.uk)                              | −0.242   | 1.571| −0.154  | 0.878 |
| C    | Prevention                                | −0.070   | 0.526| −0.133  | 0.894 |
| C    | Advertisements                            | 0.085    | 0.652| 0.130   | 0.897 |
| S    | Type (collaborative)                      | −16.231  | 1,527,315| −0.011 | 0.992 |
| S    | Type (electronic commerce)                | −16.870  | 2,896,216| −0.006 | 0.995 |
| S    | Domain (.sa)                              | −18.893  | 10,754,010| −0.002 | 0.999 |

LOOCV estimate of prediction error 0.091

Note: p-values with italics are statistically significant
Evaluation

Our logistic regression has an accuracy of 86.1 percent, a specificity of 90.0 percent and sensitivity of 60.0 percent. The prediction error of 0.091 means that 9.1 percent of the cases would be misclassified by using our model. The probability of having a quality document with all the above characteristics on and a Colemanliau of 1.5 and Ari of 15 (what was the average value in our observations) is:

\[
\Pr(\text{HON.code} = \text{yes}, \text{all.the.other.variables.from.Table XIII} = \text{yes}, \text{Type (collaborative)} = \text{no}, \text{Type (electronic commerce)} = \text{no}, \text{Domain (.gov)} = \text{no}, \text{Domain (.uk)} = \text{no}, \text{Domain (.sa)} = \text{no}, \text{Colemanliau} = 1.5, \text{Ari} = 15) = 0.9927.
\]

Where the high average value of Ari (automated readability index) means that the document is more difficult to read, i.e., it has lower readability.

Similarly,

\[
\Pr(\text{HON.code} = \text{yes}, \text{all.the.other.variables.from.Table XIII} = \text{no}, \text{Colemanliau} = 1.5, \text{Ari} = 15) = 0.0109.
\]

The three full models suggest that the probability of having a quality document is slightly lower when considering the characteristics included in the third model. That is, as expected, the predictive power of the model is worse.

Reduced model without HON characteristics

The reduced model, containing only the significant variables from the above full model, is summarized in Table XIV. All five variables remained significant.

We assessed the model's accuracy using LOOCV, with estimated prediction error of 0.0983. This means that the predictors in our model explain more than 90 percent of the variance in quality. Our logistic regression has an accuracy of 87.1 percent, a specificity of 89.3 percent and sensitivity of 67.8 percent.

Discussion

In the section “Models’ evaluation and comparison”, we summarize the variables in the multiple logistic regression models, and compare the models in terms of number of variables and evaluation rates. In the section “Characteristics’ pertinence for assessing the quality of health content”, we discuss the most important variables that contribute positively or negatively to the prediction of quality, and how can they be automatically assessed.

Models’ evaluation and comparison

As expected, our significant findings in the multivariate analysis suggest that the best model to predict documents’ quality is the one that contains every characteristic. However, we found that models without the URAC certification, which is itself a medical
certification, and without the characteristics related to HONcode principles, are also very good to forecast quality. Table XV summarizes the three types of full and reduced logistic regression models, where the first one contains every feature, the second one disregards the URAC variable and the third one excludes the variables related to the eight HONcode principles as well. The variables are sorted in alphabetical order. The third column indicates which characteristics we predict to be easy (E), medium (M) or hard (H) to automatically identify.

In the last five rows of Table XIV we summarize the evaluation metrics of the three full and reduced logistic regression models. The first row contains the number of variables included in each model, indicated in the third columns. In the second row, we can see that the first full model has the lowest prediction error estimate (LOOCV error). This model has the highest accuracy and precision (true positives rate) as well, indicated in the third and fourth row of the first column. These values suggest that the full model with all features is more appropriate to predict quality. However, its utility is lower because it contains variables, which already determine quality (URAC, and the HON characteristics).

<table>
<thead>
<tr>
<th>Cat. Variable</th>
<th>Automatic ident.</th>
<th>Full model w/o URAC exclusions</th>
<th>Reduced model w/o URAC exclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>C Advertisements</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Ari</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Colemanliau</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Commercial intent</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Copyrights</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Domain (.edu)</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Domain (.gov)</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Domain (.sa)</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Domain (.uk)</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>R Governmental nature</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Image</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Last update (&lt; 1 yr ago)</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Last update (2–3 yrs ago)</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Last update (3–4 yrs ago)</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Place of treatment</td>
<td>H</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Prevention</td>
<td>H</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Privacy policy</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Process of revision</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>R Scientific nature</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Scientific publication</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Split content</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Team of revision</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Testimonies</td>
<td>H</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Type (collaborative)</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Type (electronic commerce)</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>S Type (institutional-not scientific)</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C URAC</td>
<td>E</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>C Video</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>No. of variables</td>
<td>17</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>LOOCV error (%)</td>
<td>7.4</td>
<td>8.4</td>
<td>9.1</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>86.0</td>
<td>86.8</td>
<td>86.1</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>92.1</td>
<td>92.7</td>
<td>90.0</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>69.8</td>
<td>56.9</td>
<td>60.0</td>
</tr>
</tbody>
</table>

Table XV. Summary of the variables in the multiple logistic regression models, and comparison of the models in terms of number of variables and evaluation rates

Note: p-values with italics are statistically significant

Predicting the quality of health web documents
Characteristics’ pertinence for assessing the quality of health content

The studied models are also useful to understand which characteristics are more relevant to estimate the quality of web health documents. The ones that significantly contribute to the prediction of quality might be important for this, either positively or negatively. For example, search engines might use this information to improve their ranking algorithms, providing the user high-quality contents.

Generally, we found that documents recently updated (Last update (< 1 yr ago)), documents with content split by several webpages, documents under the responsibility of a scientific entity, documents from the educational domain (.edu), documents mentioning treatment locations and documents undergoing a process of revision are associated with quality. We predict to be easy to automatically identify the last update date, the domain and whether the document’s content is split, as noted in Table XV. On the other hand, the location of the treatment might be hard to assess automatically. The levels of automatic assessment of the nature of the responsible entity and of the process of revision are medium (Table XV). Testimonies, team of revision and the type of website are characteristics which contribute negatively to quality. Surprisingly, we found that scientific publications are also inversely related with quality. We think this is a consequence of using HONcode as our ground truth.

We must note that not all websites display the HONcode logo to indicate their voluntary adherence to these principles, although they might be managing the contents of their websites with high quality (Boyer et al., 2015). This might limit the information provided, and so explain why scientific publications can exhibit an effect opposite to the expected one.

Regarding the variables group, we noticed that, among the 16 content characteristics, 9 were found to influence the prediction of quality (readability, split content, presence of testimonies, URAC accreditation, last update date, commercial intent and advertisements, presence of images and videos, scientific publications, and inclusion of place of treatment and information regarding prevention).

Regarding the responsible entity characteristics, the ones related to its reputation were found to be the most important.

About the characteristics of the website, we found that the primary domains (.edu, .gov, .sa and .uk), disclosure information (copyrights and privacy policy), presence of information about the editorial review and the type of website (collaborative, electronic commerce and institutional-not scientific) are important for quality prediction.

Among the indicators that the present work considers more informative, several were listed in previous studies (Genova and Bender, 2016; McBride et al., 2017; De Groot et al., 2017) as indicative features used in the detection of the quality of online health information (e.g.: readability, scientific value, authorship, references and prevention or treatment). Previous works (Boyer and Dolamic, 2015; Boyer et al., 2017) proposed further development of automated systems for the detection of the quality of health web documents. The process of automatic assessment could be for instance based on machine learning methods, what could help in web structure mining, e.g., in extracting metadata from health-related web pages such as the last update date. The prediction of split content, for example, could be made by detecting the existence of previous-next buttons.

Conclusions

The goal of our research was to analyze the influence of web document features on their quality, using HONcode as ground truth, with the aim of finding whether it is possible to predict the quality of a document using its characteristics. In the multivariate analysis, we analyzed whether the characteristics of documents are useful to predict their quality and how.

Our main contributes are made with the models which can predict whether a document has quality or not. We built three types of models: the first one considers all characteristics,
the second one disregards the URAC variable, and the third one also excludes the variables related to the eight HONcode principles. The best values for the evaluation metrics we have obtained were: the LOOCV estimate of prediction error for the first model which considered every feature (7.4 percent); accuracy, also for the first model considering all characteristics (89.0 percent); specificity, for the second model disregarding the URAC variable (92.7 percent), and sensitivity for the first model including all variables (69.8 percent). We have also identified characteristics that can be used by search engines to improve their ranking. We note that using HON certification as ground truth may be associated with a higher rate of false negatives.

For predicting the quality of online health information, the most indicative features of the present analyses were the following: readability, split content, presence of testimonies, URAC accreditation, last update date, commercial intent and advertisements, presence of images and videos, scientific publications, inclusion of place of treatment and information regarding prevention, documents under the responsibility of a scientific entity; primary domains, disclosure information, presence of information about the editorial review and the type of website.

Although our models consider characteristics that might be difficult to automatically identify (e.g. prevention, prognosis or treatment), some of them will be easy to identify automatically (like the ones we already automatically assess) and, to several of them, we envision ways to automatically detect them (e.g. copyrights, images, video and type). This takes us closer to a future research direction what focuses on automated processes to infer the quality of health information on the web. Thus, in future work, we aim to use these findings as a way to improve information retrieval in the health domain.

We consider exploring technologies that might help in analyzing the content of web pages. Web structure mining or extracting metadata could be done by using machine-learning methods. The visualization of the quality information within a search engine might be as well a future goal. We will also aim to study the importance of using a single variable vs a group of variables, potentially, by search engines.

References


Further reading


Corresponding author

Melinda Oroszlányová can be contacted at: melondos@gmail.com
A study of antecedents influencing eWOM for online lecture website

Personal interactivity as moderator

Mi sook Lee
Department of Business Administration, Kyungpook National University, Daegu, The Republic of Korea, and
Hongbok An
Department of Accounting, Keimyung University, Daegu, The Republic of Korea

Abstract

Purpose – The purpose of this paper is to identify the antecedent factors – perceived usefulness (PU), perceived switching cost (PSC) and perceived web security (PWS) – affecting learners’ attitude toward online lecture website (ATW), which, in turn, affects electronic word of mouth (eWOM) and finds the factor that online lecture business should focus on the most to make learners have positive attitude.

Design/methodology/approach – This paper investigates the functional relationship among those five constructs; and examines the moderating role of personal interactivity. Data were collected from learners who had taken online lectures and were using social network sites, and a research model was analyzed using structural equation modeling.

Findings – The results show that PU and PSC positively influence ATW but PWS has no significant influence on ATW; PU is the most influential factor to ATW; ATW positively influences eWOM; personal interactivity has a moderating effect on some paths; and path coefficients are higher in the high-interactivity group than the low-interactivity group for all the links except the link from PU to ATW.

Originality/value – This paper contributes to online lecture business by understanding learners’ perception and behavior to the websites. Unlike many previous studies, this study designates eWOM as dependent variable and personal interactivity as moderation variable. This study shows interesting results occurred between low- and high-interactivity groups.

Keywords Perceived usefulness, Electronic word of mouth, Attitude toward website, Perceived switching cost, Perceived web security, Personal interactivity

Paper type Research paper

Introduction

Online lectures, including massive open online courses, are rapidly gaining popularity in South Korea due to the fierce competition for university entrance exams, certificate acquisition, reeducation programs for workers and lifelong education. Thus, online lecture institutions should become forerunners in the use of information technology in order to rejuvenate and find new and more sustainable directions (Jacobs, 2006). Online learners spread positive or negative word of mouth (WOM) on social network sites (SNSs) about the lectures they have taken. New learners can decide which online lectures are useful for them after reading electronic word of mouth (eWOM), saving time and money. SNS users express their opinions and share information with other users. Thus, SNSs could create valuable opportunities for eWOM and eWOM posted on SNSs can affect new learners’ decision.

Although many studies have analyzed the antecedents and consequences of eWOM, most have explored the impact of eWOM and regarded eWOM as an antecedent of attitude or behavioral intention (BI) (Matute et al., 2016; Ye et al., 2011; Zarrad and Debabi, 2015).
In particular, few studies have explored the antecedents of eWOM in the context of online lecture business. Unlike the previous studies, this study seeks to determine the important factors influencing attitude toward online lecture website (ATW) that is connected to eWOM. For the antecedents of ATW, we selected perception factors such as perceived usefulness (PU), perceived switching cost (PSC) and perceived web security (PWS) because we judged that they are the most influential factors on learners’ attitude in online lecture website with the following reasons. Above all, usefulness of online lecture website will be very important in that learners seek knowledge and information that is useful to them. Second, some learners are thinking about switching to the other online lecture website because they would rather not like the website they are currently using, on the other hand they are not doing so because of the switching cost (SC). Third, it was revealed that security has an impact on the consumers’ attitude, in many online-related studies.

Moreover, research works dealing with the moderating effect of personal interactivity are very scarce, while several studies have just examined the interactivity of online learning. Thus, this study differentiates from the previous studies by investigating the moderation role of personal interactivity.

With those research situations, this study poses the following main questions:

RQ1. What factor is the most influential antecedent of the ATW in online lecture website?

RQ2. Does personal interactivity play a moderating role in this study?

The rest of this paper is organized as follows. The second section briefly reviews the online lecture businesses and the literatures pertaining to eWOM on SNSs. The third section provides the research hypotheses and model and the constructs for this study. The fourth and fifth sections present the study’s methodology and results, respectively. Finally, this paper concludes with a discussion of the key findings and provides an outlook on future research.

Online lecture business and eWOM
The key element of online lecture is the use of the internet. With the expansion of internet as a tool for exchanging information, online lecture businesses include a virtual space in their websites to share information among learners (Matute et al., 2016). Traditional WOM has evolved into eWOM and internet has made WOM more rapid, convenient and pervasive, without face-to-face human communication and pressure (Lee et al., 2013; Phelps et al., 2004; Wang et al., 2016; Yeh and Choi, 2011). Several online lecture businesses advertised their website through interactive activities, but new learners tend not to trust the advertising (Brecht, 2012; Wieling and Hofman, 2010); they are much more trusting the experience of and information from other existing learners. Online reviews are perceived as helpful tools in obtaining information of products or services and reducing risks and uncertainty of buying (Teng et al., 2014).

Shin et al. (2014) regarded eWOM as “the spreading of online reviews, arguments, and recommendations that pertain to personal experiences with specific products or service providers with a view to generating persuasive effects on the targeted consumers.” While people share WOM with their friends, family and colleagues, they also increasingly share eWOM with strangers through various internet channels. People enjoy talking and posting online about products and services they have consumed. Nielsen.com (2012) demonstrated that 92 percent of 28,000 internet users in 56 countries rely on the recommendations of friends and family, and about 70 percent of them rely on online reviewers/eWOM. Thus, eWOM is exercising strong influence over customer decision making (Lee and Tussyadiah, 2011). For instance, Ye et al. (2011) showed that traveler reviews and eWOM have a significant impact on online sales and highlighted the importance of online user-generated reviews and eWOM for business performance in tourism.
SNSs act as an effective communication channel for eWOM among learners, serving as an important source of online lecture-related information and opinions (Kudeshia and Kumar, 2017). According to the report of Statista.com (2017), Facebook, WhatsApp, YouTube, Facebook Messenger and WeChat are the most well-known SNSs, ranked by the number of active users. And the leading social networks in South Korea are YouTube, Facebook, Twitter, Instagram, Kakaotalk and GooglePlus. These SNSs have become the most prevalent eWOM channels for online learners to express their opinions and communicate with the others (Cheung and Lee, 2012). Moreover, recent studies have shown that SNS itself also provides educational opportunities such as the Khan Academy, Crash Course, YouTube Edu and iTunes U (Isik, 2013; Lu et al. 2017).

Albayrak and Yildirim (2015) investigated student involvement in Facebook, examining the educational uses of SNSs, and found that Facebook has the potential to increase student involvement in discussions and out-of-class communication among instructors and students.

Research hypotheses
This study establishes its hypotheses based on the conceptual background and valid research results provided by other researchers, who have found that PU, PSC and PWS have a positive relationship with ATW and that ATW has a positive relationship with eWOM. In addition, this study investigates the moderating effect of personal interactivity.

PU and ATW
PU, an important determinant in many IS/IT studies, is defined as “the degree to which a person believes that using a particular system would enhance his or her job performance” (Davis et al., 1989).

Yang and Yoo (2004) indicated that PU affects both cognitive and affective attitudes in the usage of information systems. Wang et al. (2009) included PU and PWS as the predictors of attitude in their research model and concluded that PU and PWS significantly influence cognitive and affective attitudes. Altawallben et al. (2015) stated that PU positively influences attitude toward the learning framework of e-learning. Tarhini et al. (2017) revealed that PU is a significant determinant of students’ BI toward e-learning. Mohamud et al. (2017) found that PU has a significant influence on user’s attitude toward m-learning; however, it did not influence the BI to use m-learning. Many studies have also suggested that the more satisfied the customer is with a website, the more positive the customer’s attitude toward it (Manuel et al., 2016; O’Cass and Grace, 2004). Thus, we propose the following hypothesis:

H1(+). PU will have a positive influence on ATW ($\gamma_1 > 0$): the higher the PU score, the higher the ATW score.

PSC and ATW
Recently, SC plays a critical role as a success strategy in e-commerce but consumers sensitively react to it (Lee and Romaniuk, 2009; Park et al., 2014; Ram and Wu, 2016). SC is the total cost involved in changing from one service provider to another (Aydin and Ozer, 2005; Burnham et al., 2003). SCs include costs in monetary terms, the time and psychological effort required to deal with a new service provider, the uncertainty involved in the process and the knowledge invested in understanding the products, services or relationships (Dick and Basu, 1994; Keith et al., 2010), all of which customers need not pay if they do not switch service provider. Higher SCs are likely to make customers stay with their current service provider, since they anticipate extra burdens if they switch (Ram and Wu, 2016). In the
context of online lecture, learners might think that switching to a new online lecture program will result in some unexpected hassle or that it might not be as helpful as expected. So, they would stay with the current online lecture website.

Lee and Murphy (2005) stated that SC is the main determinant in retaining customers, compared with other factors such as price and service quality. While securing new customers is important, a strategy of increasing SC can be more important for retaining existing customers. Thus, SC can provide a buffer zone for businesses to compensate for the adverse effects of factors such as high prices, low customer trust and short-term fluctuations in service quality (Yin and Shen, 2017).

Previous studies of the relationship between SC (or PSC) and ATW are scarce but there are some studies of the relationship with customer loyalty (CL), WOM, eWOM or whatever. Chou and Lu (2009) and Kotler (2000) found that SC has a positive effect on CL. Stan et al. (2013) indicated that PSCs have the strongest, positive and direct impact on CL in comparison to the other antecedents included in the model. Park et al. (2014) highlighted that SC positively influences continuous intention to use, which, in turn, helps achieve CL. Yin and Shen (2017) found that SCs exert significant effects on perceived values and brand loyalty. However, Ram and Wu (2016) showed that SC does not have a significant influence on CL. Nusair et al. (2017) found that SCs are positively related to eWOM communication. On the other hand, learners will have negative feelings after knowing that they have to pay the time and penalty to quit the existing service in use. Jones et al. (2007) and Lee and Romanuk (2009) stated that SC may prevent customers from leaving but generate harmful WOM. So, PSC may positively, negatively or not influence ATW. Considering the characteristic of online lecture website, we would like to select the following hypothesis because most literatures demonstrate the positive effect of PSC on ATW:

\[ H2(+) : \text{PSC will have a positive influence on ATW} (\gamma_{12} > 0) \text{; the higher the PSC score, the higher the ATW score.} \]

PWS and ATW

Users are exposed to a certain security risk when using websites. Perceived security is defined as “the level of security that users feel while they are doing something on a website” (Yenissey et al., 2005). Bauer et al. (2005) emphasized that new media service users tend to have concerns about data manipulation, unauthorized data access and unwanted tracking of usage patterns. So, web security is a critical factor influencing attitude toward the web, which was proved in many studies.

Shin (2009) showed that consumers’ attitudes and intentions are influenced by perceived security in the study of the consumer acceptance of mobile wallet. Cheng et al. (2006) and Wang et al. (2009) also proved that web security has a positive influence on attitudes in the internet banking and the internet ticketing, respectively. However, Nusair et al. (2017) examined the direct influence of PWS on eWOM without ATW and found that perceived security negatively influences eWOM, in contrast to existing literature results. We propose the following hypothesis on the basis of common results:

\[ H3(+) : \text{PWS will have a positive influence on ATW} (\gamma_{13} > 0) \text{; the higher the PWS score, the higher the ATW score.} \]

ATW and eWOM

Attitude is a critical construct for predicting user behavior. Fishbein and Ajzen (1975) described attitude as a sequence of cognitive constructs: beliefs, attitudes, BI's and behaviors in the theory of reasoned action (TRA). Thus, attitude could mediate the relationship between personal perceptions (beliefs) and the behaviors.
In the most studies of the relationship between ATW and WOM (or eWOM), it was proposed that WOM (or eWOM) is an influential factor on ATW (Baber et al., 2016; Kudesia and Kumar, 2017; Ladhari and Michaud, 2015; Teng et al., 2017). Contrary to this, it was rarely proposed that ATW positively or negatively influences WOM (or eWOM) in some studies (Durukan and Bozaci, 2012; Hennig-Thurau et al., 2004; Swan and Oliver, 1989).

Durukan and Bozaci (2012) found that attitude toward social media are meaningfully related to both positive and negative WOM. Hennig-Thurau et al. (2004) conceptualized attitude in terms of two dimensions, cognitive and affective attitude, and showed that affective (rather than cognitive) attitudes significantly determine eWOM intention. On the other hand, Swan and Oliver (1989) found that consumers with positive attitude toward a product do not always recommend the product to friends and relatives. Zhao et al. (2016) found that user attitudes influence eWOM not directly but indirectly through user engagement in the study of mobile sensor computing applications. Although the order in the relationship of ATW and eWOM is suggested otherwise in the various studies, we would like to propose the following hypothesis because we think eWOM is a behavior and accepts a sequence of cognitive constructs of TRA:

\[ H4(+) \]: ATW will have a positive influence on eWOM \( (\beta_{21} > 0) \): the higher the ATW score, the higher the eWOM score.

Moderating effect of personal interactivity

A broad range of studies (Huang et al., 2014, 2017; Lin, 2011; Tarhini et al., 2017) has examined moderating effects of perceived flexibility advantages, individual cultural values and experience in online learning, but research on the moderating effect of personal interactivity is scarce.

Learners’ interactivity in online lecture means learners’ active participation in the learning process, such as bookmarking, tagging, commenting, communicating via short messages, status updating and so on. Knowledge and information sharing occurs through the interactive relationships on websites (Shipps and Phillips, 2013). Online lecture websites and social networks provide an opportunity to interact, establish relationships and engage with others through information sharing. Thus, websites that promote interactivity can be perceived as helpful and responsive in addressing its learners’ unique needs. Jo and Kim (2009) found that interactivity plays an important role in building customer relationships on the web. Interactivity is also a crucial element in improving the quality of online learning (Siemens et al., 2015). Interactive learners might acquire information or maintain social relationships with others even better than non-interactive learners. Furthermore, empirical studies have demonstrated the effectiveness of interactivity in extending learners’ attention spans, enhancing their achievements, and improving the quality of their learning (Cherrett et al., 2009; Dror et al., 2011). Conversely, however, interactivity can interrupt learners’ concentration: learners may be distracted, answer messages on SNSs or focus on external requests for their limited attention resources. Thus, we suggest that learner’s interactivity positively moderates each relationship of the research model. The hypothesis and sub-hypotheses below are established to test the moderating effect of personal interactivity:

\[ H5 \]: Personal interactivity has a moderating effect on each path: there is a significant difference on each path between the low- and high-interactive groups[1].

Learners who perceive the usefulness, SC and security of the web more strongly will have more positive ATW if they have high interactivity. The same applies to the relation between ATW and eWOM:

\[ H5a(+) \]: Personal interactivity moderates the relationship between PU and ATW: \( \gamma_{11,1}^{h} < \gamma_{11,1}^{l} \).
$H5b(+)$. Personal interactivity moderates the relationship between PSC and ATW: $\gamma_{12,1} < \gamma_{12,j}$.  

$H5c(+)$. Personal interactivity moderates the relationship between PWS and ATW: $\gamma_{13,1} < \gamma_{13,j}$.  

$H5d(+)$. Personal interactivity moderates the relationship between ATW and eWOM: $\beta_{21,1} < \beta_{21,j}$.  

The research model and hypotheses of the study are depicted in Figure 1.

**Research constructs and items**  
As shown in Table I, the study’s constructs are appropriately tailored to online lecture website by drawing from the constructs developed by many researchers.

The items shown in Table II have been developed to measure these constructs. A seven-point Likert scale, with values ranging from “(1) strongly disagree” to “(7) strongly agree,” was employed.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Definition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived usefulness (PU)</td>
<td>Degree to which a learner believes the online lecture website would enhance his/her performance</td>
<td>Amin et al. (2014), Ajzen (1991), Davis et al. (1989)</td>
</tr>
<tr>
<td>Perceived switching cost (PSC)</td>
<td>Total cost involved in changing from one online lecture website to another</td>
<td>Ram and Wu (2016), Burnham et al. (2003)</td>
</tr>
<tr>
<td>Perceived website security (PWS)</td>
<td>Belief that the online lecture website is secure for transmitting sensitive information</td>
<td>Wang et al. (2009), Nusair et al. (2017)</td>
</tr>
<tr>
<td>Attitude toward web (ATW)</td>
<td>An expression of the learner’s evolution, an action tendency and an emotional feeling toward the online lecture website</td>
<td>Chen and Wells (1999)</td>
</tr>
<tr>
<td>Electronic word of mouth (eWOM)</td>
<td>The spread of information on online lecture websites about personal experiences with a view to generating persuasive effects on future learners</td>
<td>Shin et al. (2014)</td>
</tr>
</tbody>
</table>

Table I. Constructs and definitions
Methodology
To test the research model, a questionnaire-based survey was developed. A preliminary survey was conducted on graduate students in two universities in D city of South Korea. After the questionnaire was a little revised, online survey was conducted aimed at adults countrywide with the help of researchers at a university research institution in the fall of 2016. The respondents were learners who were taking an online lecture for certificate or license acquisition and who used SNS such as Facebook and Kakaotalk. After missing values and ambiguous responses were either removed or properly treated, data from 247 learners were used to test the research model.

The analysis was conducted in three phases. The first phase was to analyze the respondents' demographic information and to examine the statistics of the measurement items using SPSS23.0; this mainly involved an analysis of the measurement model to examine its reliability and validity. The second phase was to analyze the structural model via AMOS 22.0 and to test hypotheses. Finally, an invariance analysis was conducted between the respondent subgroups based on their personal interactivity.

Results
Descriptive analysis
The demographic information for the sample is presented in Table III by gender, age, residential area, education level and employment status. Of the total, 38.5 percent of respondents were male and 61.5 percent were female. The vast majority of respondents were between 25 and 44 years old (52.2 percent). Despite the educational level of the respondents was high, they are still taking online lectures. Almost two-thirds (66.4 percent) were employed part-time or unemployed.

Measurement model
Once the items for measuring the constructs were established, their validity and reliability were assessed using confirmatory factor analysis. Tables IV and V display the results for
reliability, convergent and discriminant validity (Fornell and Larcker, 1981; Hair et al., 2010).

Regarding the scales’ reliability, Cronbach’s $\alpha$ and composite reliability (CR) were calculated. As shown in Table IV, the Cronbach’s $\alpha$ values for each construct exceed the cutoff value of 0.7, and the CR values exceed the recommended threshold of 0.7, demonstrating adequate internal consistency.
Average variance extracted (AVE) was used to assess convergent validity. The AVE values ranged from 0.709 to 0.821, exceeding the recommended cutoff value of 0.5. When the square root value of AVE of each construct is greater than its correlation value with all other constructs, the construct is considered to exhibit adequate discriminant validity. The results show acceptable levels of discriminant validity since the square root value of AVE is greater than the correlation with other constructs (Table V). These results indicate that the properties of all the constructs satisfy the established criteria for reliability as well as both convergent validity and discriminant validity.

It could be assumed that the learners are somewhat positive about the online lecture website they are using because each mean of ATW and eWOM is 5.247 and 4.687, respectively, as shown in Table V.

The fit indices suggest that the model represents a good fit to the data (Table VI). The ratio of the $\chi^2$ value to the degree of freedom ($\chi^2$/df) was less than the cutoff point of 5 ($\chi^2 = 180.087$, df = 79, $\chi^2$/df = 2.28). Besides $\chi^2$/df, the most commonly reported fit indices are the CFI, GFI, NFI and the RMSEA. The relative indices (CFI, NFI and GFI) and the absolute indicator of fit (RMSEA) indicate that the proposed model is a reasonable explanation of observed covariances among the study constructs (Chou and Lu, 2009). Recently, there is a consensus that the cut-off value of these indices needs to be a little stringent (NFI $\geq 0.95$, CFI $\geq 0.95$, GFI $\geq 0.95$ and RMSEA $\leq 0.06$) as suggested by Hu and Bentler (1999). In this study, NFI (0.931), CFI (0.96), GFI (0.912) and RMSEA (0.072) are a little acceptable even though they do not satisfy the strict cutoffs.

**Structural model and path analysis**

After an analysis of the measurement model, data analysis of the relationship among the constructs for the total sample was conducted. We used criteria similar to the measurement model in order to measure the goodness-of-fit for the structural model. As shown in Figure 2, the results of the structural model show a good fit to the data ($\chi^2 = 209.807$, df = 81; $\chi^2$/df = 2.59; GFI = 0.90; NFI = 0.92; CFI = 0.95; RMSEA = 0.08).

The results of the path analysis for the sample are as follows: first, PU and PSC positively influence ATW; the path coefficients are 0.400*** and 0.171***, respectively. Second, PWS has no statistically significant influence on ATW; the path coefficient is 0.036. Finally, ATW positively influences eWOM; the path coefficient is 0.543***. Thus, both enhancing PU and increasing PSC could be seen as an important strategy that

<table>
<thead>
<tr>
<th>Mean SD PU</th>
<th>PSC</th>
<th>PWS</th>
<th>ATW</th>
<th>eWOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>PU</td>
<td>4.015</td>
<td>1.274</td>
<td>0.899</td>
<td>0.304***</td>
</tr>
<tr>
<td>PSC</td>
<td>4.623</td>
<td>1.109</td>
<td>0.865</td>
<td>0.006</td>
</tr>
<tr>
<td>PWS</td>
<td>4.171</td>
<td>1.264</td>
<td>0.904</td>
<td>0.142**</td>
</tr>
<tr>
<td>ATW</td>
<td>5.247</td>
<td>1.151</td>
<td>0.906</td>
<td>0.415***</td>
</tr>
<tr>
<td>eWOM</td>
<td>4.687</td>
<td>1.080</td>
<td>0.842</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** The italic numbers in the diagonal are the square roots of the AVE. **p < 0.05; ***p < 0.01

<table>
<thead>
<tr>
<th>$\chi^2$ df $\chi^2$/df NFI</th>
<th>CFI</th>
<th>GFI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>180.087 79 2.28 0.931 0.96</td>
<td>0.912</td>
<td>0.072</td>
<td></td>
</tr>
</tbody>
</table>

**Recommended limit** $< 5 \geq 0.95 \geq 0.95 \geq 0.95 \leq 0.06$
spread eWOM. It seems that the learners are not sensitive to the web security because they believe it is somewhat safe.

Path coefficients and the mean of each construct will differ according to the degree of personal interactivity. Therefore, it is very important to understand whether greater personal interactivity will lead to a greater path coefficient in this study. We thus test the moderating effect of personal interactivity to determine if the result is positive or negative. To measure the personal interactivity of learners, we included three items in the questionnaire, as shown in Table VII.

The total sample was divided into two groups, low and high levels, based on the mean of the three personal interactivity items. To test for the moderating effect, we used a multi-group analysis. The standardized coefficients of the construct relationships for the two groups are presented in Table VIII. The influence of PU on ATW ($\gamma_{11}$) is marginally higher in the low level group (0.638*** than in the high level group (0.219***). The influence of PWS on ATW ($\gamma_{13}$) is negative in the low level group (-0.107) and positive in the high level group (0.167**). Finally, the influence of ATW on eWOM ($\beta_{21}$) is positive in both the low and high interactive groups (0.487*** and 0.603***, respectively).

To determine which path displays a statistically significant difference between the two groups, a constraint in which the coefficients of each group are equal (e.g. $\gamma_{11l} = \gamma_{11h}$) is

### Table VII.

<table>
<thead>
<tr>
<th>Items for checking personal interactivity level</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I actively give feedback to the tutor</td>
<td>3.15</td>
<td>1.257</td>
</tr>
<tr>
<td>I discuss some issues or problems with the other learners</td>
<td>3.35</td>
<td>1.331</td>
</tr>
<tr>
<td>I share knowledge and information with the other learners</td>
<td>3.48</td>
<td>1.379</td>
</tr>
</tbody>
</table>

### Table VIII.

<table>
<thead>
<tr>
<th></th>
<th>Low level group</th>
<th></th>
<th>High level group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SC</td>
<td>t-value</td>
<td>p-value</td>
<td>SC</td>
</tr>
<tr>
<td>PU → ATW ($\gamma_{11}$)</td>
<td>0.638</td>
<td>5.11</td>
<td>***</td>
<td>0.219</td>
</tr>
<tr>
<td>PSC → ATW ($\gamma_{12}$)</td>
<td>-0.138</td>
<td>-1.40</td>
<td>*</td>
<td>0.439</td>
</tr>
<tr>
<td>PWS → ATW ($\gamma_{13}$)</td>
<td>-0.107</td>
<td>-0.251</td>
<td>0.33</td>
<td>0.167</td>
</tr>
<tr>
<td>ATW → eWOM ($\beta_{21}$)</td>
<td>0.487</td>
<td>4.418</td>
<td>***</td>
<td>0.603</td>
</tr>
</tbody>
</table>

**Notes:** SC, standardized coefficient. *p < 0.1; **p < 0.05; ***p < 0.01
imposed on each path, step-by-step, in the structural model (Steenkamp and Baumgartner, 1998). The result of the path inequality test between the two groups is shown in Table IX. For the test to be significant, the difference in the \( \chi^2 \) value (\( \Delta \chi^2 \)) between the constrained and free (unconstrained) model must be higher than the value of the \( \chi^2 \) with 1 df, which is 3.84. The difference in the \( \chi^2 \) value in the comparison with constrained model 1 and the free model is 3.529 (369.670–359.141) < 3.84, while the difference in df is 165–164 = 1. This suggests that there is no moderating effect of personal interactivity in the path \( \gamma_{11} \), even though the path coefficients differ considerably, as shown in Table VIII. The difference in \( \chi^2 \) value in the comparison with constrained model 2 and the free model is 20.986 (380.127–359.141) > 3.84, suggesting that there is a moderating effect of personal interactivity in the path \( \gamma_{12} \). This is also true in the case of \( \gamma_{13} \) and \( \beta_{21} \) (except \( \gamma_{11} \)) are supported. 

\( \gamma_{12} \), \( \gamma_{13} \) and \( \beta_{21} \) (except \( \gamma_{11} \)) are higher in the high level group than in the low level group, as shown in Figure 3, indicating that personal interactivity is an important moderator. Moreover, complete moderation is considered to occur if the hypothesis for the main effect is not significant, while partial moderation is considered to occur if the hypothesis for the main effect is still significant (Awang, 2012). Since \( \gamma_{11} \) is not significant, we can conclude that complete moderation occurs on the path; as \( \gamma_{11} \), \( \gamma_{12} \) and \( \beta_{21} \) are significant, we can conclude that partial moderation occurs on the paths.

Conclusions and implications

The education paradigm is continuously evolving based on edu-Tech products with advanced ICT technologies all over the world. Beyond online lecture, simulation-based, game-based and mobile-based learning are expected to be used for the next generation.

### Table IX.
Comparison between the free model and the constrained model (\( \chi^2 \) of free model = 359.141; df = 164)

<table>
<thead>
<tr>
<th>Constraint</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>( \Delta \chi^2 )</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained model 1</td>
<td>( \gamma_{11,l} = \gamma_{11,h} )</td>
<td>362.670</td>
<td>165</td>
<td>3.529</td>
</tr>
<tr>
<td>Constrained model 2</td>
<td>( \gamma_{12,l} = \gamma_{12,h} )</td>
<td>380.127</td>
<td>165</td>
<td>20.986</td>
</tr>
<tr>
<td>Constrained model 3</td>
<td>( \gamma_{13,l} = \gamma_{13,h} )</td>
<td>363.170</td>
<td>165</td>
<td>4.029</td>
</tr>
<tr>
<td>Constrained model 4</td>
<td>( \beta_{21,l} = \beta_{21,h} )</td>
<td>365.094</td>
<td>165</td>
<td>5.953</td>
</tr>
</tbody>
</table>

**Notes:** Subscript \( l \) = low level group, subscript \( h \) = high level group. \( \Delta \chi^2 = \chi^2 \) of constrained model – \( \chi^2 \) of free model

**Figure 3.** Comparison of standardized path coefficient for two different groups

**Perceived Usefulness (PU)**

**Perceived Switching Cost (PSC)**

**Attitude Toward Web (ATW)**

**Electronic Word of Mouth (WOM)**

**Notes:** Path coefficient of low level group vs (path coefficient of high level group). \( \chi^2 = 359.14; df=164; \chi^2/df=2.19; GFI=0.85; NFI=0.88; CFI=0.93; RMSEA=0.07. \)

\*\( p<0.1; \**\( p<0.05; \***\( p<0.01 \)
In this evolving education situation, this study aims to make theoretical proposals for the follow-up study and practical suggestions for online lecture service providers, by examining the learners’ response to online lecture website. The primary goals of this study are to develop a research model representing learners’ perceptions of online lecture website and their eWOM behavior, to empirically test it and to examine the moderating role of personal interactivity.

The results of this study provide insights into the interrelationships between PU, PSC, PWS, ATW and eWOM about online lecture website. The hypothesis test results for total sample are as follows:

1. The respondents are highly positive about an online lecture website in that the mean score of each construct is over 4, and the mean score of ATW is especially 5.247.

2. PU among the antecedents of ATW most positively influences ATW; PWS has no statistically significant influence on ATW; ATW positively influences eWOM.

In the study of Wang et al. (2009), only PU and PWS among the antecedents (PEOU, PU and PWS) of cognitive and affective attitudes have significant impact on the attitudes and the influence of PU is the greatest on both cognitive and affective attitudes. In the studies of Altawallben et al. (2015) and Yang and Yoo (2004), it was concluded that the influence of PU is the greatest on both cognitive and affective attitudes. In the study of Tarhini et al. (2017), however, it was revealed that PU influences students’ BI toward e-learning but it is not the most significant factor. When considering the results of these preceding studies, it can be concluded that PU is important on the attitude of website. In various website studies (Cheng et al., 2006; Wang et al., 2009), web security is considered to be an important factor on the attitude toward the website. However, the results of this study are clearly different from those of the studies. Overall, online lecture businesses should make students aware of the usefulness of online lectures and take a positive attitude toward websites regardless of the SC.

To identify which links cause statistically significant differences according to personal interactivity, the structural invariance for each hypothesized structural path was tested. The sample was divided into two groups, low and high levels, based on the personal interactivity of the learners. The analysis of a moderating effect of personal interactivity indicates the followings:

1. The influence of PU on ATW is considerably higher in the low level group than in the high level group, but there is no moderating effect of personal interactivity in the path in the \( \chi^2 \) test.

2. The influence of PSC on ATW is negative in the low level group and positive in the high level group. This implies that PSC has a negative influence on ATW among low interactive learners even though PSC positively influences ATW in many studies.

3. The influence of PWS on ATW is negative in the low level group and positive in the high level group. This result is similar to (2).

4. The influence of ATW on eWOM is positive in both low and high level groups.

5. Through the invariance analysis, personal interactivity is found to be an important moderator except on the path of PU to ATW.

The path coefficients are higher in the high level group than the low level group for all the links except the link from PU to ATW, interestingly implying that learners who are less interactive have more positive attitude toward the web if they perceive more usefulness.

This study contributes to academic and practical research in an environment, where online lecture businesses are becoming increasingly competitive. First, this study reviews eWOM and online lecture websites and summarizes the relevant literature. Second, the study reveals that, to generate eWOM for an online lecture website, the manager and/or
staff of online lecture businesses should strive to maintain a positive attitude toward the web among existing learners by focusing on the website's usefulness and SC. In other words, as PU among the antecedents of ATW is the most important factor and the relationship between ATW and eWOM is very strong, the executives should put more effort into improving the usefulness perceived by learners and consequentially making learners' attitude toward its online lecture website positive. Judging from these findings, online lecture providers should strive to be positive for ATW by focusing on developing programs and contents that allow learners to participate more interactively. In addition, they should consider the learners' interactivity when using SC as a strategic tool because the influence of PSC on ATW is negative for the low-interactive learners.

There are some thresholds of research. The results of the study may vary depending on the teaching style and type of the online lecture but this study did not focus on a particular course of a particular institution. However, we just would like to examine the general responses to online lectures without limitation on a particular online lecture. Another limitation is that we used a convenience sample. The sample consisted mainly of learners taking online lectures taken by graduate students who had participated in the preliminary survey. The lectures are mainly for certificate or license acquisition and the education institutions vary as well. In the future, a more detailed sample design would have to be made for comparative analysis of online lecture in various fields.

This study has some suggestions that should be noted for the future research. First, further research should investigate the elements that would most likely increase usefulness of online lecture website and focus on them because PU is proved as the most influential factor on ATW. Second, future research should find the way to make low interactive learners not have negative attitudes when PSC is strategically high because the path coefficient from PSC to ATW is shown to be negative in the low level group. Finally, although eWOM can be negative as well as positive, this study focuses on only positive eWOM. So, more comprehensive aspect should be considered as a dependent variable.

We hope that this study will help stimulate further research while contributing to the development of online lecture businesses, riding high on the information technology.

Notes
1. This study considers the "low level group" as the group of learners whose personal interactivity is low and the "high level group" as the group of learners whose personal interactivity is high (subscript \(_l\) = low level group; subscript \(_h\) = high level group).
2. There were some ambiguous responses in education level and employment. We properly treated them by logical judgment.

References


1063


**About the authors**

Mi sook Lee was Research Professor in Graduate School of Management, Kyungpook National University. She majored in both applied statistics and business administration. Her area of interest includes user’s behavior or adoption toward a new information technology. Mi sook Lee is the corresponding author and can be contacted at: leems@knu.ac.kr

Hongbok An is currently Professor in the College of Business Administration, Keimyung University. His area of interest includes corporate governance, firm innovation and business performance.

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
To get cited or get tweeted: a study of psychological academic articles

Yingxin Estella Ye
NUS Libraries, National University of Singapore, Singapore, and
Jin-Cheon Na
Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore

Abstract
Purpose – By analyzing journal articles with high citation counts but low Twitter mentions and vice versa, the purpose of this paper is to provide an overall picture of differences between citation counts and Twitter mentions of academic articles.
Design/methodology/approach – Citation counts from the Web of Science and Twitter mentions of psychological articles under the Social Science Citation Index collection were collected for data analysis. An approach combining both statistical and simple content analysis was adopted to examine important factors contributing to citation counts and Twitter mentions, as well as the patterns of tweets mentioning academic articles.
Findings – Compared to citation counts, Twitter mentions have stronger affiliations with readability and accessibility of academic papers. Readability here was defined as the content size of articles and the usage of jargon and scientific expressions. In addition, Twitter activities, such as the use of hashtags and user mentions, could better facilitate the sharing of articles. Even though discussions of articles or related social phenomena were spotted in the contents of tweets, simple counts of Twitter mentions may not be reliable enough for research evaluations due to issues such as Twitter bots and a deficient understanding of Twitter users' motivations for mentioning academic articles on Twitter.
Originality/value – This study has elaborated on the differences between Twitter-inclined and citation-inclined articles. It provides useful information for interested parties who would like to adopt social web metrics such as Twitter mentions as traces of broader engagement with academic literature and potential suggestions to increase the reliability of Twitter metrics. In addition, it gives specific tips for researchers to increase research visibility and get attention from the general public on Twitter.

Keywords Twitter; Altmetrics; Citation; Social web; Scholarly communication

Paper type Research paper

Introduction
Assessing and evaluating the impacts of scholarly outputs is an essential process in scientific progress and the knowledge advancement of society. Derived from peer-review publishing, citation analysis is the dominant form of scholarly communication and was developed as the primary mechanism for evaluating research (Haustein, Sugimoto and Larivière, 2015).

The development of the internet has cast new light on traditional scholarly communication. (Gu and Widén-Wulff, 2011). Web 2.0 tools, such as reference managing websites (e.g. Mendeley and CiteULike), wikis, blogs, and social networking sites (e.g. Twitter and Facebook), have sparked faster and less formal methods for scholarly communication and have opened up the boundaries across academia and non-academic communities (Shema, Bar-Ilan and Thelwall, 2012). An increasing number of researchers would like to use the social web to maintain awareness about latest research news and information, disseminate information, interact with diverse audiences, connect with peers in professional networks (Manca and Ranieri, 2016; Veletsianos and Kimmons, 2016; Van Noorden, 2014). Correspondingly, these tools and platforms have yielded broader and...
timelier impact assessments of scholarly outputs using altmetrics based on social web activities, such as readership, discussions, sharing, and recommendations for a variety of audiences (Fenner, 2014; Das and Mishra, 2014).

Utilizing psychological academic articles as a case study, we examine the differences between citation counts and Twitter mentions in research evaluations. We studied the Twitter mentions of academic articles for the following reasons: Twitter is a primary social media platform used by scientists and researchers (Peoples et al., 2016); and the diversity of the Twitter community allows us to draw a comprehensive picture of scholarly communication on the social web, and also the meaning of social web metrics (Barthel et al., 2015). The field of psychology was selected because it is one of the most popular subjects on social media platforms (Mohammadi and Thelwall, 2014).

The main goal of this study is to better comprehend the role that Twitter plays in scholarly communication, and the value of Twitter mentions in research assessment. It aims to provide useful information to parties including funding agencies, university administrators, and policymakers who would like to adopt social web metrics such as Twitter mentions as traces of broader engagement with academic literature. These traces may be of interest to these parties as the use of academic literature extends beyond the academic community. We also try to encourage and offer tips for researchers to share their research outputs with a larger audience on the social web for higher research visibility and impact.

In this study, starting with a review of existing scholarship examining citation counts, altmetrics, and their determinants, we compared highly cited and highly tweeted articles to identify distinct differences between citation counts and Twitter mentions. We first explored how selected attributes including document characteristics, sources, accessibility, wording and research topics influence differences between citation counts and Twitter mentions. We then investigated the patterns of Tweets citing these academic articles with a simple content analysis approach in which we covered various variables including hashtags, user mentions, and retweets.

Literature review

Citation counts and altmetrics

The emergence of altmetrics has attracted wide attention from scholars. Altmetrics are regarded as alternative metrics of research evaluations because they make up for the deficiencies of traditional impact assessments; for example, they can handle the most recent publications (Brigham, 2014), they expand the targets of measurements to scholarly works in various formats not limited to written works (Piwowar, 2013), and aggregated altmetrics are able to reflect the achievements of researchers (Mounce, 2013). However, there are some debates regarding the validity of altmetrics. For example, due to the lack of rigorous regulations, cyber-based metrics may not be as reliable as peer-reviewed citations. They can be easily manipulated: for example, “buying” or “selling” posts engagements such as “like” and “share” can lead to a dramatic increase of social web mentions. Moreover, without robust filtering mechanisms or standards, the results of research assessment can be exaggerated due to “spam” posts (Barthel et al., 2015). Altmetrics, in use as a type of social web metrics, may also be influenced by heuristics such as authority and bandwagon cues through which the contents posted by authorities and users with many followers are more easily mentioned and disseminated. It is also more likely for social media users to cite their peers’ posts (Lee and Sundar, 2013; Lin et al., 2016). Stemming from a deficient understanding of the actual meaning of altmetrics, traditional bibliometrics are still the only assessment mechanism of research impacts used by majority of academic institutions (Peoples et al., 2016).

Considerable research has been undertaken regarding correlations between citation counts and social web mentions. A majority of these studies have concentrated on the following questions: whether altmetrics can predict traditional bibliometrics, whether
Altmetrics can be applied as an impact assessment tool by interested parties, and whether intensive scholarly communication on the social web can lead to higher citation rates. Based on the analysis of Facebook mentions and citation counts of articles from a variety of disciplines, Ringelhan et al. (2015) suggested Facebook mentions as an early indicator of the future impact of scientific works according to their positive association. However, they admitted that the validity of the early prediction may differ between disciplines; for instance, the correlation was found to be more significant for psychology articles compared to non-psychology fields such as business and life sciences. Similar results were found in Eysenbach’s (2011) study, which showed that tweet mentions can predict highly cited articles within the first three days of the publication of scholarly works. He argued that with capacities to measure the social impacts of articles, Twitter metric should be proposed as a supplement to traditional bibliometrics.

These studies have provided inspiring insights into the relationship between citation counts and altmetrics. However, concentrating on correlation testing, some studies may have ignored the fact that these measurements can sometimes be at odds with each other. In addition, the statistical analysis of correlations, either related or not related, is not able to tell the stories hidden under these numerical metrics.

Citations counts, social web mentions and determinant factors

Citation counts may sometimes fail to reflect the quality or intellectuality of scholarly publications (Gargouri et al., 2010). They were found to be positively related to the accessibility of articles. Taking measures such as making articles open access and the individual or institutional archiving of academic papers increases the number of citation counts (Xia et al., 2011; Norris et al., 2008). As mentioned by Didegah and Thelwall (2013), the journal impact factor and the citation impact of referenced articles could be the most effective determinants of citation counts. Besides, an article’s internationality and number of authors may determine its citation rate as well. For instance, Gazni and Didegah (2011), through examining the Harvard University affiliated publications between 2000 and 2009, demonstrated that articles involving a larger number of authors and institutions had higher chances of attracting more citations. Taking published research articles from Finland as a case study, Puuska et al. (2014) concluded that on top of the impact of the number of authors, international co-publications usually receive more citation counts than domestic collaborations. Similarly, Sin (2011) pointed out that library and information science articles involving international collaborations and authors from high-income nations are more likely to be cited.

Based on 1.3m Web of Science (WoS) indexed articles published in 2012, Haustein, Costas and Larivière (2015) discovered that document characteristics, such as document type, the number of pages, the length of articles, and collaboration patterns, can affect the number of citation counts and social media metrics. For instance, it was found that the number of pages is positively correlated with citation counts but negatively associated with an article’s popularity on Twitter, especially in the disciplines of life and earth sciences, math and computer science, social science and humanities. A longer article title may attract more citation counts but relatively less social media mentions. However, differences existed across disciplines; for example, a longer title led to fewer citations for articles in disciplines such as biomedical and health sciences, natural sciences, and engineering, but brought more citations for articles in social sciences and humanities.

Inspired by the scholarship mentioned above, four main factors, including document characteristics, sources, accessibility, and research topics, were examined in this study in an attempt to attain consistent findings in the field of psychology. We selected articles with incongruent rates of citations and Twitter mentions to clearly visualize the affiliation
between relevant factors and the metrics, and also to better understand the differences between citation counts and Twitter mentions. The comparisons were conducted between articles with incongruent rates of citation counts and Twitter mentions.

**Twitter analysis**
While previous studies have provided compelling investigations into altmetrics, including Twitter metric, the actual contents of tweets were seldom dissected. To enrich the understanding of Twitter counts, a basic analysis of tweets was utilized in this study to evaluate the meaning of Twitter mentions.

Existing studies related to Twitter analysis have mainly focused on Twitter activities such as the use of hashtags, user mentions (@username), replies, retweets, favorites, including URLs, and so forth. For example, Small (2011) found hashtags in tweets worth exploring because they are keywords assigned to the information, and they provide a summarized description of the contents. According to Small, hashtags can indicate the topics or events embedded in tweets. In a study by Börgmann et al. (2016), the authors examined Twitter discussions on urologic oncology using frequency analysis of words and hashtags in tweet contents, while Naveed et al. (2011) took the use of elements such as URLs, exclamations and question marks, and emoticons into account in their Twitter analysis of interestingness. Bruns and Stieglitz (2013) made efforts to develop a set of standardized metrics for Twitter analysis with particular focus on hashtagged exchanges, for instance, counts of tweets sent, counts of original tweets, counts of user mentions sent, counts of genuine replies, retweeted statuses (unedited or edited), unique users, and so on.

We understand that approaches such as sentiment analysis and network analysis are also very popular in Twitter analysis, but due to the limited time and resources, only a basic analysis of Twitter activities such as hashtags, user mentions, and retweets was conducted to discover the patterns of tweets citing academic psychology articles.

**Methodology**

*Extraction of psychological articles and related posts*
Psychological articles published between 2012 and 2016 were extracted from WoS on November 18, 2017. The extraction was based on the International Standard Serial Number (ISSN) of journals under Social Science Citation Index (SSCI) 2015, covering ten disciplines including applied psychology, biological psychology, clinical psychology, developmental psychology, educational psychology, experimental psychology, mathematical psychology, multidisciplinary psychology, psychoanalysis and social psychology.

The number of related journals was 587. Articles from WoS were extracted from the Request API for WoS Data based on queries combining the above-mentioned ISSNs and publication dates between “January 1, 2012” and “December 31, 2016” under the SSCI collection. The publication date indexed in WoS refers to “the date on which the records were entered in the product database” (“Web of Science Core Collection Help,” 2017). A total number of 217,768 articles (identified by the accession number “UID” in WoS) were retrieved with their titles, authors, abstracts (if available), publishers, sources, publication dates, page counts and citation counts. At the same time, we extracted Twitter mention data from Altmetric.com through the Altmetric Explorer by entering the same list of ISSNs and limited the publication period between “January 1, 2012” and “December 31, 2016” in the advanced search. In total, 568 out of 587 journals and 124,974 articles (identified by item ID) were retrieved. Next, Digital Object Identifiers (DOI) was used as the linkage between WoS records and records from Altmetric.com. Before this, we removed 157 WoS records with duplicate DOIs due to incorrect data entries. Records of 92,435 unique articles (identified by DOIs) were matched. Table I shows the number of articles retrieved in each discipline. It is
possible for one article or one journal to be classified into more than one category. We used R Studio to process data in this study.

We recognized limitations of the data collection in this study: first, there are inconsistencies between WoS-indexed articles and those captured by Altmetric.com. For example, not all articles are cross-indexed by both databases. Only successfully matched articles were taken into account in this study. Second, Altmetric.com tracks social web mentions based on identifiers like DOIs and URLs ("How it works – Altmetric," 2017). This may lead to a low coverage rate.

Clustering articles based on citations counts and twitter mentions

To better demonstrate the distinct differences between citation counts and Twitter mentions, we make comparisons between highly cited and highly tweeted articles. We targeted articles in the following two clusters: citation-inclined: articles with high citation counts captured by WoS but low Twitter mentions; and Twitter-inclined: articles with high Twitter mentions but low citation counts.

As Twitter mentions and citation counts were found to be heavily affected by recency (date of publication) and research disciplines (Zahedi et al., 2014; Hammarfelt, 2014), in order to investigate articles in different disciplines and those that were published in different years, normalization was applied to Twitter mentions and citations separately in each discipline and publication year. The formula \((i - 0.5)/n \times 100\) used in the study of Bornmann and Haunschild (2016) was adopted for the calculation of percentiles; \(i\) is the integer value of rank position (ascending) of an article in the set of articles within the same discipline and published in the same year, while \(n\) refers to the total number of articles in the set. If an article is classified into \(m\) number of disciplines, we divided the sum of percentiles by \(m\) to get an average percentile. We used Avg Tp to represent average Twitter mention percentile and Avg Cp to represent average citation count percentile.

According to the 80/20 scientometric data quality rule (Strotmann and Zhao, 2015), bibliometric data needs to have at least 80 percent coverage in the database to allow a reliable field-specific study. Therefore, before data normalization, we excluded sets of publications in each discipline and publication year in which over 20 percent of articles received zero mention on Twitter (see Table II). The number of remaining unique articles was 49,150.

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Number of articles indexed in Altmetric.com</th>
<th>Number of articles indexed in WoS</th>
<th>Number of articles matched</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychology, Biological</td>
<td>4,271</td>
<td>10,586</td>
<td>3,374</td>
</tr>
<tr>
<td>Psychology, Clinical</td>
<td>26,307</td>
<td>44,267</td>
<td>19,872</td>
</tr>
<tr>
<td>Psychology, Educational</td>
<td>5,549</td>
<td>11,793</td>
<td>4,950</td>
</tr>
<tr>
<td>Psychology, Developmental</td>
<td>17,662</td>
<td>26,279</td>
<td>13,516</td>
</tr>
<tr>
<td>Psychology, Applied</td>
<td>10,282</td>
<td>20,777</td>
<td>8,700</td>
</tr>
<tr>
<td>Psychology, Multidisciplinary</td>
<td>25,809</td>
<td>57,559</td>
<td>20,923</td>
</tr>
<tr>
<td>Psychology, Psychoanalysis</td>
<td>870</td>
<td>3,814</td>
<td>597</td>
</tr>
<tr>
<td>Psychology, Mathematical</td>
<td>2,022</td>
<td>3,132</td>
<td>1,687</td>
</tr>
<tr>
<td>Psychology, Experimental</td>
<td>21,731</td>
<td>42,908</td>
<td>17,771</td>
</tr>
<tr>
<td>Psychology, Social</td>
<td>11,647</td>
<td>18,388</td>
<td>10,583</td>
</tr>
</tbody>
</table>

Table I. Number of articles retrieved from WoS and altmetric.com
We classified an article with an Avg Cp above the 70th percentile but an Avg Tp below the 30th percentile as citation-inclined, while an article with an Avg Tp above the 70th percentile but an Avg Cp below the 30th percentile was classified as Twitter-inclined. We selected the 70th and 30th percentiles as thresholds for the following reasons: the difference between the Avg Cp and Avg Tp makes the articles in citation-inclined and Twitter-inclined clusters distinctive; and this allowed us to keep a relatively large and balanced set of articles that provided a higher level of accuracy in data analysis. In this way, 3,536 articles (7.19 percent) and 3,026 articles (6.16 percent) were, respectively, classified as Twitter-inclined and citation-inclined articles (see Figure 1).

As shown in Table III, articles in the Twitter-inclined cluster had been mentioned at least three times on Twitter, while the most-tweeted articles had been mentioned 695 times on Twitter. Regarding citation counts of Twitter-inclined articles, the range was between 0 and 4. The citation counts among citation-inclined articles fell between 3 and 114, while the Twitter mentions ranged from 0 to 2.

<table>
<thead>
<tr>
<th>Discipline</th>
<th>2012%</th>
<th>2013%</th>
<th>2014%</th>
<th>2015%</th>
<th>2016%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychology, Applied</td>
<td>56.03</td>
<td>71.91</td>
<td>77.66</td>
<td>79.08</td>
<td>88.04</td>
</tr>
<tr>
<td>Psychology, Biological</td>
<td>60.28</td>
<td>69.37</td>
<td>76.12</td>
<td>86.52</td>
<td>94.99</td>
</tr>
<tr>
<td>Psychology, Clinical</td>
<td>69.12</td>
<td>77.7</td>
<td>79.76</td>
<td>83.53</td>
<td>91.02</td>
</tr>
<tr>
<td>Psychology, Developmental</td>
<td>71.57</td>
<td>79.05</td>
<td>82.72</td>
<td>89.17</td>
<td>92.25</td>
</tr>
<tr>
<td>Psychology, Educational</td>
<td>58.78</td>
<td>70.53</td>
<td>76.49</td>
<td>83.3</td>
<td>89.64</td>
</tr>
<tr>
<td>Psychology, Experimental</td>
<td>54.77</td>
<td>67.71</td>
<td>74.82</td>
<td>79.62</td>
<td>90.14</td>
</tr>
<tr>
<td>Psychology, Mathematical</td>
<td>44.65</td>
<td>62.88</td>
<td>67.35</td>
<td>74.21</td>
<td>86.29</td>
</tr>
<tr>
<td>Psychology, Multidisciplinary</td>
<td>68.87</td>
<td>78.84</td>
<td>82.05</td>
<td>86.67</td>
<td>89.97</td>
</tr>
<tr>
<td>Psychology, Psychoanalysis</td>
<td>85.19</td>
<td>75.76</td>
<td>83.56</td>
<td>71.62</td>
<td>67.46</td>
</tr>
<tr>
<td>Psychology, Social</td>
<td>72.29</td>
<td>81.25</td>
<td>83.64</td>
<td>84.57</td>
<td>90.60</td>
</tr>
</tbody>
</table>

Table II. Twitter coverage of publications by discipline and publication year

Figure 1. Clustering articles based on the 70th and 30th percentiles of citation counts and twitter mentions
In order to better comprehend Twitter mentions, we extracted tweets mentioning Twitter-inclined articles from Altmetric.com. Information retrieved included the username of the tweet author, publication data and time, tweet ID, and tweet contents. The number of tweets collected was 47,146. We further utilized the Twitter API for Python to gather additional details of the tweets. Matching with tweet IDs, Twitter API provides information such as tweet contents, users involved, interactions (i.e. retweet, reply, and quote), geolocations, etc. In this study, retweeted status (retweeted_status), quoted status (quoted_status), user mentions (user_mentions), and hashtags (hashtags) were retrieved. Since some of the tweets had been deleted by their authors, only 32,394 tweet records were successfully retrieved and used in our data analysis.

**Data analysis**

*Article attributes affiliated with citation counts and Twitter mentions.* Four main attributes, including document characteristics, source, accessibility, and the topics and wording of article titles and abstracts, were used for analysis (see Table IV). Statistical analysis was utilized to test the correlations between the above-mentioned variables and the clustering of articles. A $\chi^2$ test was used for categorical variables and an independent sample $t$-test was used for numerical variables.

*Patterns of Twitter mentions.* In order to better comprehend Twitter mentions, we conducted basic analysis of tweet contents that mentioned selected articles. Table V shows how we defined relevant attributes in this study.

By presenting sample tweets, we made further interpretations on users’ Twitter activities where psychological academic articles were mentioned, e.g. the top hashtags based on their occurrences in Tweets, the way hashtags were used in tweets, the patterns of the most popular tweets based on the number of retweets, and so forth. Through this, we tried to demonstrate the diversity of user patterns in communication and interactions. Lastly, we added a section to share explorative findings during our studies as well to provide more ideas about tweets and Twitter metric.

**Findings and discussion**

*Article attributes affiliated with citation counts and Twitter mentions*

In this section, we compare articles in Twitter-inclined and citation-inclined clusters to identify potential attributes that are affiliated with the traditional counts and Twitter mentions of academic articles.

*Document characteristics.* First, the content size of articles is found to be an important factor contributing to the Twitter inclination of articles. Articles with less pages and shorter titles and abstracts seem to have higher chances of becoming highly tweeted. In contrast, articles with longer titles and abstracts and more pages are more appealing to citation counts. These mean differences (see Table VI) are all significant at 0.01 level. For example, the average length of abstracts of Twitter-inclined articles is 150.80, while, for the

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Mean</th>
<th>Median</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twitter-inclined (3,536 articles)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twitter mention</td>
<td>15.67</td>
<td>10</td>
<td>695</td>
<td>3</td>
</tr>
<tr>
<td>Citation count</td>
<td>0.5</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Citation-inclined (3,026 articles)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twitter mention</td>
<td>0.81</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Citation count</td>
<td>10.13</td>
<td>7</td>
<td>114</td>
<td>3</td>
</tr>
</tbody>
</table>

Table III. Descriptive statistics (citation counts and twitter mentions) of clustered articles
citation-inclined articles, the mean is 172.02. Twitter-inclined articles also tend to have shorter titles and contents. Regarding the number of authors, citation-inclined articles are more likely to have more authors (4.12 authors on average) than Twitter-inclined articles (3.47 authors on average). This mean difference is significant at 0.01 level.

**Impacts of sources.** The impacts of journals were reported to have a strong correlation with citation counts (Didegah and Thelwall, 2013). This conclusion can also be applied to Twitter mentions, according to our study. Figure 3 shows that around 85 percent of citation-inclined articles and over 74 percent of Twitter-inclined articles were published in journals in Quartile 1 and Quartile 2 (average JIF percentile between 50 and 100).
However, generally, the impacts of journals may have stronger affiliation with citation counts. The mean of average JIF percentiles of citation-inclined articles is 71.65 percent, while for Twitter-inclined articles, the mean of average JIF percentiles is 64.65 percent. The mean difference is significant at 0.01 level in t-test ($t = 13.29, p < 0.01$).

**Accessibility.** Compared to citation counts, accessibility drives more Twitter mentions. In our study, 12.46 percent of Twitter-inclined articles are open access articles. This percentage is higher than that of citation-inclined articles (3.48 percent) and higher than the percentage of open access articles among all selected articles (7.62 percent). This difference is significant at 0.01 level in $\chi^2$ test ($\chi^2 = 160.24, p < 0.01$).

Accessibility is a very important precondition for article dissemination and discussion on social web platforms such as Twitter. Twitter users outside of academia may have very limited access to non-open access articles, let alone the ability to share them with others. However, this may not be a critical problem for scientists affiliated with universities or research institutes, which should provide them accessible resources for research including subscription-based articles.

**Research topic and wording.** The feature words extracted indicate a similar finding: articles with jargon and scientific expressions are less likely to be Twitter-inclined. For example, as shown in Figure 2, highly tweeted articles have less jargon compared to highly cited articles, e.g. “sct” (sluggish cognitive temp), “nssi” (non-suicidal self-injury), “dsm” (Diagnostic and Statistical Manual of Mental Disorders), “comorbidity,” and so forth. In addition, theory or methodology-related terms are more likely to appear in citation-inclined articles, e.g. “bifactor,” “discriminant,” “validity,” “confirmatory,” “conceptualization,” “narcissism,” etc.

Twitter mentions seem to have a close affiliation with topics related to sex and gender (represented by terms such as “infidelity,” “heterosexual,” “mating,” and “sexuality”), developmental disorders (e.g. “autistic,” “autism”), crime and violence (e.g. “crime,” “victim,” “violent”), and so on (Figure 3).

A potential explanation of these differences could be the diversity of Twitter users. Twitter users from the general public may not have the strong knowledge backgrounds of researchers and scientists that allow them to understand academic theories or methodologies in articles.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hashtags</td>
<td>Binary variable: whether the article was mentioned with hashtags</td>
</tr>
<tr>
<td>User mentions</td>
<td>Binary variable: whether the article was shared with user mentions</td>
</tr>
<tr>
<td>Retweet</td>
<td>Binary variable: whether the tweet is a retweet (either retweet or quote tweet, which refers to a retweet with comments). We did not analyze quote tweets separately because this feature was introduced in 2015 (Griffin, 2015) and, hence, the status of tweets before this launch may not be properly captured</td>
</tr>
<tr>
<td></td>
<td>Numeric variable: number of tweets retweeting the tweet</td>
</tr>
</tbody>
</table>

Table V. Identifying patterns of tweets

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Mean in citation-inclined cluster</th>
<th>Mean in Twitter-inclined cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of title (word counts) $t = 5.63, p &lt; 0.01$</td>
<td>14.03</td>
<td>13.56</td>
</tr>
<tr>
<td>Length of abstract (word counts) $t = 14.07, p &lt; 0.01$</td>
<td>172.02</td>
<td>150.80</td>
</tr>
<tr>
<td>Number of pages $t = 10.05, p &lt; 0.01$</td>
<td>12.78</td>
<td>11.33</td>
</tr>
<tr>
<td>Number of authors $t = 10.64, p &lt; 0.01$</td>
<td>4.12</td>
<td>3.47</td>
</tr>
</tbody>
</table>

Table VI. Document characteristics: mean differences between Twitter-inclined and citation-inclined articles
Hence, they tend to selectively absorb and share information or knowledge which is easier to understand on the internet. This can be regarded as a weakness of Twitter mentions because they may be biased against articles with academic or scientific expressions, such as academic concepts and experimental methods. Moreover, they have strong preferences toward particular topics. This could also be a reason why certain highly cited articles within academia have failed to arouse wide discussions on the social web.

**Patterns of twitter mentions**

In this section, we examined tweets mentioning Twitter-inclined articles with the purpose of understanding patterns hidden behind the numeric counts.

**Twitter activities**. Hashtags and user mentions were found to have positive correlations with the number of Twitter mentions. Among Twitter-inclined articles, articles assigned with at least one hashtag when mentioned on Twitter received 18.32 mentions on average, while those mentioned without hashtags received 11.55 mentions on average ($t = 4.09, p < 0.01$). In terms of user mentions, articles shared without any Twitter handle or username received 7.15 Twitter mentions, while those that mentioned other users had higher Twitter mentions with 16.23 on average ($t = 4.13, p < 0.01$). Both mean differences are significant at 0.01 level.

Figure 4 shows hashtags that occurred at least 10 times in the tweets mentioning Twitter-inclined articles. The size of the circles reflects the number of tweets. Hashtags that occurred together are represented by links, while the thickness of the links refers to the number of times the hashtag occurred. Most of these hashtags represent the topics of articles, and they also show that topics such as mental health (e.g. #mentalhealth, #depression, #PTSD, and #anxiety), developmental disorders (e.g. #autism), family violence (e.g. #childabuse), and sex and gender (e.g. #LGBT, #sexual and #HIV) are
relatively popular among Twitter users. In addition, Twitter users may assign generic tags such as “#psychology” and “science” together with the topics of articles.

Another interesting pattern among these hashtags is the mention of open access status (e.g. #openaccess, #Free, and #OpenAccess). These Twitter users could be researchers, publishers, universities, research institutes, and others. By sharing the academic articles with others, these Twitter users intended to announce the release of their publication and to attract the audience on Twitter. Below are three example tweets; the first is a tweet shared by the department of psychology of a university, the second tweet was posted by a scientific publisher, and the third tweet was shared by a research fellow working in a research institute of a university:

Our department has published a new paper on maternal cognitions and principles #openaccess http://t.co/xxxxxx

#openaccess in our journal; research article on #autism #PDA http://t.co/xxxxxx

Our #review paper on in #cats! #OpenAccess #communication https://t.co/xxxxxx

In addition, publishers (e.g. #springerlink and #IJED) are highlighted in the form of hashtags. It is interesting that not only publishers added themselves as hashtags in their tweets in order to announce new publications, researchers and practitioners also tagged publishers when they share academic articles on Twitter. For example, an associate
A professor in the field of speech language pathology wrote: “I’m reading this on https://t.co/xxxxx #springerlink.” Another example is from a Juris doctoral student, sharing his first journal article with the following tweet: “My first journal article has just been published! https://t.co/xxxxx #springerlink.” These tweets were probably modified based on tweet templates from journal websites. For instance, the tweet template for Springer journals is “the title of article + URL of the article + #springerlink.”

It is also common among Twitter users to mention other users when sharing articles on Twitter. For example, some users would like to mention the authors in their post, such as “A preregistered replication study of a classic psychological theory [...] http://t.co/xxxxxxxxx from @username,” “An interesting replication attempt by @username and @username. http://t.co/xxxxxxxxx” and “Congratulations to our lecturer @username on her latest publication http://t.co/xxxxxxxx.” Twitter mentions were also used to alert others to the article, for instance, “[...] This is a possible explanation for gender paradox [...]@username” and “[@username This is an interesting paper! http://t.co/xxxxxxx.” Twitter users may use user mentions to relate the article to other relevant users as well. For instance, a professor of psychology mentioned another professor in the same research area in his post “[...] Robot-enhanced psychotherapy works [...] http://t.co/xxxxxxxxx @username is crying tears of joy.”

Retweets. We observed that 56.78 percent of tweets mentioning Twitter-inclined articles are retweets. And due to Twitter’s character limit, contents posted by Twitter users were quite straightforward. Most of these top tweets simply highlighted the findings of articles or posted short comments about articles or related social phenomena with additional URLs. The patterns of the top ten most retweeted English-written tweets are listed in Table VII. First, these top tweets were posted by various types of users, such as a literature

<table>
<thead>
<tr>
<th>Tweet author</th>
<th>Tweet content pattern</th>
<th>Topics</th>
<th>Number of retweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature recommendation page</td>
<td>summary of the article + URL of the article + screenshot of the abstract</td>
<td>Gender stereotypes and socialization</td>
<td>229</td>
</tr>
<tr>
<td>Professor of sociology</td>
<td>highlights of research findings + URL of the article</td>
<td>Racism and internet</td>
<td>167</td>
</tr>
<tr>
<td>Postdoctoral fellow of psychology</td>
<td>comments on the article (“this article has addressed a huge gap in the scientific literature”) + URL of the article + screenshot of highlighted contents</td>
<td>Psychological characteristics of professionals</td>
<td>167</td>
</tr>
<tr>
<td>Software engineer</td>
<td>highlights of research findings + URL of the article</td>
<td>Feminism</td>
<td>165</td>
</tr>
<tr>
<td>Professor of psychology (1)</td>
<td>discussion on social phenomenon based on findings of the article + URL of the article quote from the article + URL of the article</td>
<td>Gender stereotypes and socialization</td>
<td>93</td>
</tr>
<tr>
<td>Scientific publisher</td>
<td>summary of research finding to support personal view + URL of the article</td>
<td>Education and psychology</td>
<td>92</td>
</tr>
<tr>
<td>Blogger</td>
<td>summary of research findings + URL of the article + screenshot of highlighted contents</td>
<td>Racial microaggressions and psychotherapy</td>
<td>70</td>
</tr>
<tr>
<td>Literature recommendation page (2)</td>
<td>highlights of research findings + URL of the article + screenshot of highlighted contents</td>
<td>Feminism</td>
<td>69</td>
</tr>
<tr>
<td>Professor of psychology (2)</td>
<td>comments on the articles (“amusing and plausible”) + URL of the article</td>
<td>Autism</td>
<td>68</td>
</tr>
</tbody>
</table>

Table VII. Top 10 most retweeted English-written tweets
recommendation page that mainly shares the latest studies, including academic articles, in specific research disciplines, researchers or scientists (e.g. professors and research fellows), scientific publishers, other professionals (e.g. software engineers), and members of the general public (e.g. bloggers). This finding is consistent with the diverse user community of Twitter.

Additional explorative findings

Replies to tweets. Even though the tweets mentioning psychological academic articles may not involve in-depth discussions, we noticed some interesting discussions and interactions between Twitter users, especially in replies to tweets. In order to tell a diverse story, we showed below some examples of replies to the top retweeted tweets mentioning psychology articles. We understand that without systematic sampling or text clustering analysis, the replies covered in this section are not completely representative.

The first three replies question the research methods of the articles, while the fourth example refers to the research gap. In other words, we may assume that scholarly discussions do exist among Twitter mentions, though we did not study their motivations of the sharing further. It is also evident that Twitter users link real-life phenomena or their personal views with the academic articles mentioned on Twitter (see the 5th, 6th, and 7th examples):

Would there be any value in performing the same study using men?

Their sample size is too small […] This really needs replication.

[…] it is a persuasive argument […] but it seems to be post-hoc […].

[…] still massive gap for studies on non-Italian clowns […].

[…] does that mean we should teach children with autism to read using the whole word rather phonics method […].

[…] We do not understand why you, as a professional, consider autism to be a disorder […].

[…] I find it insulting to be told I’m (as a female) weak enough to be oppressed.

Issue of Twitter bots. Consistent with previous studies (Barthel et al., 2015), we are aware of the Twitter bots among the Twitter mentions that we examined in this study. Ferrara et al. (2016) demonstrated that a significant difference between human accounts and bot accounts is that bots tend to retweet more than human accounts, while by contrast, human accounts usually receive more balanced interactive engagements, such as replies, mentions, and retweets. In accordance with this, we attempted to identify suspects by the imbalance between their activities and user engagements that they received.

For instance, we plotted the number of tweets mentioning selected psychological articles by Twitter accounts to represent the level of activity and the number of selected tweets being retweeted to reflect the engagements which they attained in Figure 5. It is apparent that the three accounts circled on the left of the chart were considerably active regarding the number of tweets. However, their tweets were seldom retweeted. In other words, these accounts may have received very limited responses from other Twitter users. Additionally, we happened to find that three of them were linked together with the account highlighted on the right, with almost all of their posts were shared at the same time with the same content. When looking into the content of the tweets, a majority of their posts attempted to share publications including academic books and articles related to the topics of HIV, sex workers and prison health from different publishers. Without concrete evidence, we cannot determine the actual purpose of these posts, but we have to admit that suspicious accounts like these may partly pollute the Twitter mentions, as well as their validity and reliability in measuring impacts.
Limitations
There are some limitations in this study. First, only psychology articles and Twitter were taken for analysis. The findings may be different when the focus is shifted to other disciplines or platforms with different demographics. Second, this study only covered articles indexed by the WoS and Altmetric.com, while articles without valid web identifiers were ignored. Additionally, this study lacks an analysis of interactions among Twitter users and of their motivations for scholarly communication on the social web.

Conclusion
Comparisons between Twitter-inclined and citation-inclined articles showed the “preferences” of citation counts and Twitter mentions. Twitter mentions were found to be closely affiliated with the readability of articles, defined as the content size (length of the article, abstract, and title) and the usage of scientific jargon, and the accessibility of articles. This may be attributed to the diverse audience on Twitter, as some Twitter users may not have the background to understand jargon or academic theories. Twitter mentions were “picky” toward specific article topics. In contrast, citation counts were friendly to articles with longer titles and abstracts, a larger number of pages, and those that placed greater emphasis on scientific theories and methodologies. In this sense, both Twitter mentions and citation counts may not perfectly reflect academic or other impacts of articles.

We noticed other weakness of Twitter mentions in this study. However, though this study did not provide a profound interpretation of the meaning of Twitter metrics, the great potentials of Twitter metric in measuring the broader engagement of scholarly outputs other than academic values in the strict sense should be recognized. One reason is that we observed a variety of users and audience involved in the article sharing on Twitter, e.g. academic researchers, research institutions, publishers, professional practitioners, members from the general public and so on. Moreover, discussions toward the articles exist though they may not be very rigorous or insightful. For example, as shown in this study, Twitter users may recommend interesting articles, question the methodology of the mentioned articles, summarize research gaps, link real-life experiences with academic articles and so forth. To enhance the reliability of Twitter metrics, we suggest the following measures as
potential solutions: first, in addition to the number of tweets, the engagement of Twitter users should be analyzed to reflect the popularity of articles on Twitter. For example, in our study, the contents of replies seem to be very informational and sometimes more comprehensive than the original tweets. Unfortunately, these interactions were not analyzed in this study. Future studies can explore the possibility of adopting social network analysis to measure the impacts of articles on Twitter. Second, it is necessary to use filtering algorithms to identify Twitter bots and spam posts because they have the power to dramatically increase Twitter mentions. For instance, the suspicious accounts spotted in this study “contributed” over thousands of Twitter mentions combined. Third, it would be useful to apply sentiment analysis to tweets, since some tweets actually criticize the articles rather than favoring them.

Twitter activity may increase the odds of Twitter mentions. Hence, we encourage researchers to engage with a larger audience including research peers, industrial practitioners, and also the general public on Twitter or other social platforms. While we are not certain about the kind of impact that can be reflected by Twitter metric, sharing scholarly output on the social web can improve knowledge exchange and also promote the transformation of academic knowledge to technological and societal advancement. Below are some practical tips for researchers who would like to share their research outputs on Twitter or other social media platforms:

(1) Increase the accessibility of articles by publishing open access. If there is not enough funding to support gold open access, an alternative way to increase the visibility of scholarly outputs is through self-archiving or depositing appropriate versions (e.g. preprints or postprints) into open repositories.

(2) Tag articles with relevant research topics, and try to adopt tags commonly used by other researchers or online communities.

(3) Link articles with related social phenomena or industrial trends that resonate with the general public.

(4) When mentioning or sharing academic articles on the social web, try to translate them into simple messages, e.g. avoiding excessive use of jargon and summarizing important contents for diverse audiences.

References


Corresponding author
Yingxin Estella Ye can be contacted at: estella.ye@nus.edu.sg

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
Mediating impact of fan-page engagement on social media connectedness and followers purchase intention

Zoha Rahman
Department of Media Studies, University of Malaya, Kuala Lumpur, Malaysia

Sedigheh Moghavvemmi
Department of Operation and Information Management System, University of Malaya, Kuala Lumpur, Malaysia

Kumaran Suberamanaian
Department of Indian Studies, University of Malaya, Kuala Lumpur, Malaysia

Hasmah Zanuddin
Department of Media Studies, University of Malaya, Kuala Lumpur, Malaysia, and

Hairul Nizam Bin Md Nasir
Department of Software Engineering, University of Malaya, Kuala Lumpur, Malaysia

Abstract

Purpose – The purpose of this paper is to identify the mediating effect of fan-page followers’ engagement activities and moderating role of followers’ demographic profile and trust level on their purchase intention.

Design/methodology/approach – This study utilised the customer engagement behaviour and consumer involvement theory as a foundation to explore the impact of variables. Structural equation modelling was utilised to test the model with the data collected from 307 Facebook fan pages’ followers of five Malaysian companies.

Findings – It was shown that following fan pages will influence fan page engagement, which in turn affects purchase intention and social media connectedness. Further analysis indicated that the impact of “follow” and “engagement” on purchase intention differs between genders, ages, level of trust and income.

Research limitations/implications – The study serves as a basic fundamental guideline for academics and researchers to interpret the concept of following fan pages and engagement actions and its effects on purchase intention and social media connectivity, as well as opening a vast area of unexplored researches on the subject of social media.

Practical implications – The research provides information for business-to-consumer companies in utilising fan page based on user categories.

Originality/value – This study proposes the application of an empirically tested framework to the fan-page follow actions. The authors argue that this framework can provide a useful foundation for future social commerce research. The results would help academics be aware of fan page and its user’s engagement actions, which will provide a new avenue of research.

Keywords Social media, Purchase intention, Fan page, Social media connectedness, Social media engagement

Paper type Research paper

This research is funded and supported by University Malaya Research Grant (UMRG) Project No. RP024-15HNE and PPF (Project No. PG182-2015A) Program of University of Malaya, Malaysia. The authors would like to give our special thanks and gratitude to the University of Malaya Research Grant Program for injecting financial support to have necessary research equipment, research-workers, research assistants associated with this research.
1. Introduction

Company fan pages are fast becoming a prominent marketing channel, due to their contribution towards increasing sales (Chetna et al., 2016). Facebook pages result in more selling opportunities for business to consumer (B2C) companies, as it increases web-traffic (Matista, 2015). However, the mechanism of a fan page and its management remains obscure (Benedikt and Werner, 2012), especially in the context of customer engagements (Benedikt and Werner, 2012). Facebook marketers attempted to determine the value of Facebook fan and determine whether joining a brand’s social network changes consumer behaviour (Leslie et al., 2017). Researchers are also interested in elucidating whether or not the positive associations between following a brand on social media and consumer positive behaviour is causal (Leslie et al., 2017). They are also interested in determining consumer behaviour towards a brand, as little is known about the relationship between consumers’ brand “likes” on a social network and their behaviour vis-à-vis a brand (Wallace et al., 2014; Chetna et al., 2016).

Customer brand engagement (CBE) is the subject of many research works. While the study on CBE in a social media-based brand community, such as Facebook brand pages, began, several research works were needed to help improve its understanding (Yap and Lee, 2014). Most studies focussed on identifying factors associated with joining online communities (Abeer and Abdelhamid, 2017), while others examined the implications of participation in online brand communities in the context of consumer behaviours, such as brand liking and the intention to purchase (Úbolvá, 2014; Panos et al., 2015; Laurence et al., 2015; Abeer and Abdelhamid, 2017; Chetna et al., 2016). Many researches were conducted on different product categories (Úbolvá, 2014; Angella and Kim, 2016; Constantinos et al., 2016) to identify the impact of social media engagement, as social media reaches consumers at different phases of the purchase funnel (Fulgoni and Lipsman, 2015). The subsequent results are important for industrial players, as it increases their understanding of online consumer’s behaviour and brand engagement. It will also create new knowledge that companies can use to improve its marketing strategy (Schultz, 2016; Lisette et al., 2017).

Facebook is the most widely used social networking site by consumers (Schultz, 2016; Hutter et al., 2013) and business owners. Of the Fortune 500 companies, a total of 348, or 70 per cent, are on Facebook (Barnes, 2014). Users of corporate fan pages are more experienced, and recent studies suggested further research on Facebook brand page community members (Wimmala et al., 2017), as the fans are the measure of its potential (Naylor et al., 2012). Participation, “involvement”, or simply, “membership” are issues that most researches investigate, due to many works suggesting inherent differences between active/passive fans (Úbolvá, 2014; Laurence et al., 2015). However, not many studies examined fan engagement and their corresponding activities (Like, Share, Comments) and its effect on the intention to purchase a brand (Abeer and Abdelhamid, 2017). Another important factor that is fast gaining traction is the impact of trust on community engagement behaviour (Li-Chun, 2017). A consumer’s trust on a retailer positively impact consumers’ attitudes toward the retailers’ physical stores (Kim and Park, 2005), and further investigation, particularly for online extensions, is strongly recommended (Yu-Teng et al., 2015).

Most companies utilise various online channels to connect to their customer, especially social media. Schultz (2016) suggested that researches should address the interactions between multiple social media across one or more levels as well. Furthermore, it is also vital for the marketers to develop a clear idea as to what extent each media platform (e.g. website, Facebook, Twitter, television) interacts with others (Richard et al., 2011), or the connectedness of social networking sites. This will create a clear picture of customer reaction to company activities and new knowledge towards improving marketing strategies.

Taking into account previous researches, this study aims to examine the effect of following fan pages on customer engagement, purchase intention and social media connectedness among followers of fan pages. The study also aims to explore the mediating impact of fan-page engagement.
effect of fan-page engagement on social media connectedness and purchase intention. The moderating effect of the users’ age, gender, income and trust on purchase intention will also be examined. The results of this study will create new knowledge that will be useful for academic and industrial practitioners. Research shows that 91 per cent of online users in Malaysia shop online, and 85 per cent of these shoppers spend ~RM500/month (Wong, 2014). Facebook is the most popular social networking in Malaysia, used by 41 per cent of the population (Wearesocial, 2016). It will therefore be useful to elucidate their online behaviour vis-à-vis the company’s Facebook fan pages and their corresponding engagement and purchase intention. The findings of this study will contribute to knowledge pertaining to customer behaviour, which will help industries improve their online presence.

2. Literature review and hypothesis development

During the past decade, information system (IS) researchers focussed on understanding social media/networks user behaviours. Several theories/models have been proposed to explain the engagement, usage, predictors and outcome of using social media/social networks by individuals and organisations. They indicated that customer engagement can be a special form of a broadly discussed concept in the service research community (Brodie et al., 2011). Consumer engagement behaviour (CEB) is related to interactions with an aspect of a brand or media property, while engagement can turn into action/behaviour, e.g. communication and/or purchase (Roderick et al., 2011). Customer engagement behaviour (CEB) goes “beyond transaction, and may be specifically defined as a customer’s behavioural manifestation toward a brand or firm, beyond purchase (e.g. including word-of-mouth (WOM) activity, recommendations, helping other customers, blogging, writing reviews) resulting from motivational drivers” (van et al., 2010).

Based on consumer involvement theory (CIT) (Sangeeta, 2017), individuals’ purchase behaviour depends on their devoted time and energy in making purchase decisions (McNamara, 2016). According to CIT, consumers can be classified as brand loyalists, information seekers, routine brand buyers, and brand switchers. Taking into account the significance of involvement for consumer researchers, James discussed three distinct concepts under “involvement”, namely communication, commitment, and response involvement in the context of the evolving body of knowledge (James and Shelby, 1984). Based on this theory, “communication involvement” suggests that individuals might react during communication; it will not begin before the commencement of the communication, and will only continue as long as communication is ongoing (Krugman, 1970). Houston and Rothschild introduced the term “response involvement”, which was defined as “the complexity of cognitive and behavioural processes characterising the overall consumer decision process”. Commitment is another concept that has been studied in the context of CIT that affect buyer behaviour, including purchase intention (Assael, 1981).

2.1 Fan-page follow and social media connectedness

Scholars classify social media based on features such as blogs, cooperative projects (e.g. Wikipedia), content communities (e.g. YouTube), virtual social worlds (e.g. Second Life), and social networking sites (SNS) (e.g. Facebook & MySpace) (Kaplan and Haenlein, 2010). Most companies use one (or more) of these channels to connect to their customer or other businesses. However, for companies using different channels, managing multiple social media accounts (e.g. Facebook, Twitter, LinkedIn) is challenging. Panos et al. (2015) reported that most companies prefer an integrated account management system that help them maintain social media connectivity across multiple platforms. The connectedness of social media is linked to users’ online ties (Riedl et al., 2013). It can also be seen as integration, as sites, resources, and people being connected via links shared by users on various social media platforms (Chen, 2014). Connectedness is a potential source of social capital, where people may realise network benefits by managing both strong and weak ties (Riedl et al., 2013).
Social media users become fans or followers of a Facebook fan page by pressing the “like” button, which specifies to their social network that they like a brand. The new content or update of this fan page is automatically posted to their personal Facebook news feed, and they can then comment on it, post on that fan page, forward offers, and contact the company and interact with other fans. Following a fan page is an action of consumer brand involvement and emotional attachment with a brand (Benedikt and Werner, 2012; Chetna et al., 2016), and all social media platforms form similar communication environments (Linjuan and Wan-Hsiu, 2013), which could lead to fan-page followers connecting with other social platforms of similar brands.

Consumer involvement affects the ways in which consumers seek, process and transmit information. As the level of consumer involvement increases, they have greater motivation to gather, comprehend, elaborate, and assimilate information (Sangeeta, 2017). Pereira argued that ~10.9 per cent fan page followers found the brand profile from a fan page. In the “communication involvement” concept of CIT, involvement with something is time-specific (i.e. the communication), rendering it situationally specific and transitory. Based on the definition of “communication involvement”, we assume that being involved in fan page activities also lead users to be involved in other online/social platforms of the same brand simultaneously (e.g. clicking on other social networking sites of the same brands from fan page links, diverting to a company discussion board, or redirecting to company blog).

It was also reported that self-oriented value of the fan page followers leads to a positive engagement (Benedikt and Werner, 2012). Thus, the study hypothesise that:

\[ H1. \] There is a significant positive relationship between the fan-page “Follow” and “social media connectedness”.

### 2.2 Fan-page follow and fan-page engagement

Fan page becomes an ideal platform for cooperation, information sharing and collaboration. Fan page delivers updated contents that were created and shared with other fans (Musser and O'Reilly, 2006). Social media platforms provide numerous opportunities for the customers to affirm their engagements with their favourite brands (Chetna et al., 2016). Benedikt defined fan-page engagement as an interactive/integrative participation in the fan-page community. He differentiated between the engagement construct from the solely follow action of a member, and signified that fan page usage leads to fan-page engagement (Benedikt and Werner, 2012). Consumer engagement on a social networking site begins by establishing a connection between oneself and a brand page (Champoux and McGlynn, 2012; Naylor et al., 2012). Engagement is what the public feel about social media content and what they do about it, including searching for, commenting on, and sharing content online (Benedikt and Werner, 2012). Úblová (2014) suggested that consuming, participating, and producing are engagement activities. Smith (2014) argued that four main engagement activities from users’ perspective are information consumption, sense of presence, interest immersion and social interaction (Smith, 2014).

Companies that publish brand posts on their pages expect that it would be shared and consumed by the fans (Muk et al., 2014). A study conducted on Brand followers pointed out that 29.8 per cent followers “read the brand post”, 20.1 per cent followers “navigate in the brand application”, 15.4 per cent “put ‘like’ on the brand post”, 1.09 per cent “navigate in the brand profile”, 6.2 per cent “comment on the brand post”, and 3.4 per cent “write on the brand wall” (He’lia and María, 2014). Evidence confirms that social networking platforms facilitate the engagement with other community members (Laurence et al., 2015). Thus, users who “like” the same corporate SNS page also often “like” and comment on each other’s posts, and the communal atmosphere encourages users to share user-generated content, such as product reviews, photos, and videos, which in turn strengthen the group dynamics and public engagement (Linjuan and Wan-Hsiu, 2013).
In the consumer involvement theory, Houston and Rothschild introduced the term “response involvement”, where high response involvement represents situations where individuals are highly active and information-processing beings that are out to gain as much information as possible, then use it to make optimal choices. Thus, according to the concept of “response involvement” of CIT, we can assume that when users are involved with brand fan pages, they are also simultaneously active in the fan page to find more information about a brand, as the contents delivered by Facebook fan pages fosters consumers’ intention to do online information search and engagement (Ojala and Tyrväinen, 2011). Thus, this study hypothesise that:

H2. There is a significant positive relationship between the fan page “follow” and Fan-page “Engagement”.

2.3 Fan-page engagement and social media connectedness

Research revealed that when people engage a social media via “like” or “retweet”, they usually do it through their common networks (Panos et al., 2015). Hirschman’s (1970) classic model identifies CEB as behaviours designed to curtail/expand their relationship with a brand (Hirschman, 1970). Brodie et al. confirmed that engaged consumers exhibit consumer empowerment, connection, and commitment (Hollebeek, 2011). Evidence shows that engagement with both the online brand community and brand is closely related, and the online brand community engagement enhances consumer brand interactions (Laurence et al., 2015). The behavioural manifestation of online brand community engagement is called endorsement (Laurence et al., 2015), which transcends community boundaries (van et al., 2010).

As per cognitive CBE, CBE generates customer’s levels of brand activity-related thought process and elaboration (Mayada, 2016). Moreover, learning represents a sub dimension of consumer’s behavioural engagement action that confirms that engagement is a committed act of looking for information (Laurence et al., 2015). Thus, we assume that if consumers engage in online brand community, they may also search for information in other brand related linked pages. However, consumer engagement could not exist along a singular consumer-brand nexus, but a complex web of interactions, where engagement is intertwined in multiple sites, including other platforms (Laurence et al., 2015). Based on CEB assumption, we assume that fan-page engagement actions may drive users or consumers to connect with other linked sites of the brand. Thus, the study hypothesise that:

H3. There is a significant positive relationship between fan-page engagement and social media connectedness.

2.4 Fan-page engagement and purchase intention

Hsu defined purchase intention as consumer willingness to buy and repurchase. Purchase intentions refer to the degree of perceptual conviction of a customer to purchase a particular product (or service) (Bamini et al., 2014). Fan-page users exhibiting high usage intensity are in regular contact with the brand, which in turn effect their relationship with the brand and increase their likelihood to repurchase (Jahn and Kunz, 2014). Studies on fashion brand reported that fans who behave actively and participate in various activities are ready to buy a product (Úblová, 2014). Active customers build brands by increasing awareness and stimulating purchases (Úblová, 2014). Previous research works suggest that consumer interactions on a social networking site lead to recommendations and purchase intentions (Cheung et al., 2012). Engaged community members are likely to purchase brand products or services endorsed by their respective online communities (Li-Chun, 2017).

Consumer engagement lead to both rational loyalty (overall satisfaction, intent to repurchase, and intent to recommend) and emotional attachment (including confidence in a
brand, belief in its integrity, and pride/passion in the brand) (Roderick et al., 2011). Community engagement enhances consumer loyalty (Algesheim et al., 2005), which reflects consumer’s willingness to purchase or use a product (Li-Chun, 2017). A fundamental aspect of online community engagement refers to how consumers are ready to continue using a brand product or repurchase (Algesheim et al., 2005). Consumer engagement on a social networking site increases consumer’s brand commitment, and consequently recommendation and purchase intention (Cheung et al., 2012). According to CEB, consumer engagement is a behavioural construct that goes beyond purchase behaviour alone, which includes retention and cross-buying sales and transaction metrics (Jenny et al., 2010). Thus, we hypothesise that:

H4. There is significant positive relationship between fan-page engagement and purchase intention.

2.5 Fan-page follow and purchase intention

The usage of fan pages by leading businesses raises its direct value on the basis of improved contacts and better turnovers, obtained via followers’ recommendations and WOM (Hopkins, 2012). Social media is key towards increasing WOM, which will in turn increase sales and return on investment due to increased brand awareness (Kumar and Mirchandani, 2012). Researches show that following a fan page will increase viewing of advertisements, which increases both brand awareness and purchase intention (Nielsen, 2011). Research conducted among multinational companies who have strong presence and fan following on Facebook brand pages confirmed the relationship between following fan page and purchase intention (Abeer and Abdelhamid, 2017; Chetna et al., 2016; Li-Chun, 2017). They found a significant relationship between fan page “Like” and purchase intention (Chetna et al., 2016). Considering the value of fans for a company, companies are now trying to increase the depth of commitment and loyalty among fans, which will affect the purchase intention of friends of the fans (He’ tia and Maria, 2014; Barnes, 2014). Fan-page follow is a state of brand involvement/attachment action (Chetna et al., 2016), and higher involvements can create positive attitude towards products and brands (Dahlen and Rosengren, 2003). Increased interaction via social network and community membership will influence purchase intention (Schultz, 2016), and change customer behaviours towards a brand (Yu-Teng et al., 2015). According to commitment “concept of CIT involvement has been hypothesised to be related to buyer behaviour including purchase intention” (Assael, 1981). Considering the above evidence, this study assumes that:

H5. There is a significant positive relationship between fan-page “Follow” and “Purchase Intention”.

2.6 Mediating impact of fan-page engagement

Having a large number of fans is inadequate for businesses, as the fans needs to be kept involved (He’Tia and Maria, 2014). Research revealed that liking a fan page have little impact on purchase intention, and the study emphasised fan page engagement activities for a positive output from users (Chetna et al., 2016). Creating a brand fan page and generating traffic data (e.g. likes, visits, page impressions, etc.) is no longer adequate. A brand fan-page strategy involves completely engaging customers and integrating themselves in the online community to increase sales (Benedikt and Werner, 2012; Úblová, 2014). Following brand page might result in active engagement (Linjuan and Wan-Hsiu, 2013), and the activities on brand pages not only affect the perception of brands, but also influences consumers’ purchase decision (Hutter et al., 2013). However, consumers who follow brands on social networks are considered as potential brand activists, and they publicly share their affection
on various social networking sites (Hollebeek, 2011), which eventually enhance social media connectedness. CBE was proved to be a mediator between social media marketing and brand equity (Mayada, 2016). However, recent research showed that social media engagement mediates social media usage and communication behaviour (Paek et al., 2013). Based on literature, we assume that there is a relationship between followers’ activity in company fan page and purchase intention and social media connectedness. Therefore this study hypothesise that:

\[ H6a. \] Fan page engagement mediates the relationship between “Follow” and “connectedness”.

\[ H6b. \] Fan page engagement mediates the relationship between “Follow” and “Purchase Intention”.

2.7 Moderating role of trust, gender, age and income

In online communities, trust has a major influence on consumer behaviour (Hoffman et al., 1999), and significantly affect purchase intention (Lil and Kim, 2007). Trust is seen as a co-existing mechanism that reduces uncertainty and complexity of transactions and relationships in electronic markets, thus increasing online shopping (Grabner, 2002). Besides, consumer’s intention to adopt a service would be significantly influenced by his/her perceived trust on the service provider (Pavlou and Gefen, 2004). Lacking in confidence on a marketer or retailer, consumers would hardly place orders, whether shopping offline or online (Gefen, 2000). When consumers believe in a retailer, it would decrease their perceived risk in related online transactions (Harridge-March, 2006). Thus, it can be argued that consumers’ trust in social media marketers will affect their perceived value and purchase intentions (Chen, 2014). Therefore, we hypothesis that:

\[ H7a(i). \] Trust will moderate the influence of fan-page “follow” towards purchase intention.

\[ H7a(ii). \] Trust will moderate the influence of fan-page engagement towards purchase intention.

Gender is regarded as a moderating variable in the social media behaviour of consumers (Duggan and Brenner, 2013) and a significant difference was found between males and females in terms of motivation in online shopping (Huang and Yang, 2010). Moreover, gender has been extensively used as a moderator variable, mainly when analysing consumer behaviour (Saad and Gill, 2001). The differences between males and females have been observed in online communication and usage behaviours (Kang, 2011). Studies showed that the effect of online consumer reviews on purchase intention is stronger for females than males (Bae and Lee, 2011). Chetna reported that the relationship between liking a fan page and intention to purchase the followed brand is significantly moderated by gender (Chetna et al., 2016). Therefore, this study posit the following hypotheses:

\[ H7b(i). \] Gender will moderate the influence of fan-page “follow” towards purchase intention.

\[ H7b(ii). \] Gender will moderate the influence of fan-page engagement towards purchase intention.

Age is an important factor influencing internet usage behaviours such as messaging, searching, interacting, and purchasing (Teo, 2001). Wang and Fesenmaier (2004) pointed out that in an online community, younger groups (i.e. up to 40 years old) appreciated functional benefits (e.g. information gathering and ease of transactions) from online participation (Wang and Fesenmaier, 2004). Ernst and Young (2014) reported that those who are 15–25 years old are the largest active social media audiences. Moreover, young men are regarded
as the typical profile of the early adopters of online shopping (Patricia et al., 2005). Thus, this study posits the following hypotheses:

- **H7c(i).** Age will moderate the influence of fan-page follow towards purchase intention.
- **H7c(ii).** Age will moderate the influence of fan-page engagement towards purchase intention.

A study on the Malaysian online consumers’ purchase intention explored the fact that income has as strong and indirect effect on online shopping intention (Narges et al., 2011). A study on Malaysian social media users found that the consumers that usually purchase goods online are young and earns ~RM2001 a month (Elisabeta and Ivona, 2014). Taking into account the results from the previous study, we hypothesise that:

- **H7d(i).** Income will moderate the influence of fan-page “follow” towards purchase intention.
- **H7d(ii).** Income will moderate the influence of fan-page engagement towards purchase intention.

Following previous research based on the classical concepts of CEB and CIT, this study explains how following a fan page might influence fan-page engagement, social media connectedness and purchase intention (Figure 1). We assume that the effect of follow and engagement towards purchase intention is different among different gender, age, and income with different level of trust. We assume that following a fan page may drive followers to be active, maintain social media connectedness and increase purchase intention. If users engage in a fan page of particular brand, they tend to purchase that product. Therefore, we developed the following model based on these assumptions.

### 3. Method

A survey on fan page followers was conducted to test the hypotheses in our research framework. We selected fan page followers as our respondents because they can provide information on fan page behavioural patterns and measure engagement metrics (Elliott, 2011). Four popular fan pages in Malaysia (Lenovo Mobile Malaysia, Samsung Malaysia, Zalora Malaysia and KFC Malaysia) that use fan page as a strong marketing tool and are active in fan pages with a large number of followers or high user growth rates were selected. We selected these firms based on judgment sampling (according to number of followers, PTA (people talking about) metrics, fan page growth rate).
A pilot study was conducted among 100 followers of the selected fan pages and the results of the reliability test shows that all the constructs have Cronbach $\alpha$ more than 0.7, which is acceptable. The final survey questionnaire, along with the research objective and motivations were made accessible through a web-link to the followers of the four sample firms. In all, 150 active followers from each fan page were identified, and 100 respondents were selected randomly for the final survey.

The weblink of the questionnaires was personally emailed to the followers. Out of 400 questionnaires, 326 completed questionnaires were returned by the respondents, out of which 307 responses were found to be valid and complete, and thus used for the study. The respondents’ demographic profile is listed in Appendix 1.

3.1 Measurement instrument
The research model contains four constructs; fan page following, fan-page engagement, social media connectedness, and purchase intention (Figure 2). For the constructs of the research model, multi-item scales were adopted from the previous study and modified for the current research. Fan page following comprises six measured variables adopted from Chetna et al. (2016); “fan page engagement” comprises 13 items adopted from Linjuan and Wan-Hsiu (2013); He’lia and Maria (2014), “social media connectivity” comprises five items adopted from He’lia and Maria (2014); Chen (2014). Consequently, “Purchase Intention” comprises six items adopted from Chetna et al. (2016); Angella and Kim (2016) (see Appendix 2). Respondents were asked to rate the questions on five-point Likert scales, from 1 = strongly disagree, and 5 = strongly agree.

4. Data analysis
4.1 Reliability measurement
Internal consistency reliability of the items evaluated through Cronbach $\alpha$ and the results show the $\alpha$ value of each construct achieving a floor criterion of 0.7. The reliability results of the measured items are shown in Appendix 2. The factor-item composition was verified by undertaking exploratory factor analysis. Principal component analysis based factor extraction, together with Varimax Rotation, reinforced the factor-item composition. All the items (except two deleted items) achieved factor loadings above the criteria of 0.5 (Guadagnoli and Velicer, 1988). This study tested the research model using structural
equation modelling (Amos 19). We used maximum likelihood estimation method to analyse the data due to the good estimation result of this method with minimum variance that best represent the population (Chen, 2014). The AMOS programme has always a feature that allows Maximum Likelihood Estimation (Ferron and Hess, 2007). To analyse the data using AMOS, it is necessary that the univariate and multivariate of normality be checked. An examination of these values shows minimal indication of serious multivariate outliers.

4.2 Confirmatory factor analysis
Confirmatory factor analysis is a multivariate statistical procedure used to test how well the measured variables represent the number of constructs and able to confirm the validated measurements developed by previous research (Solutions, 2013).

4.2.1 Construct validity. Construct validity was determined by measuring the fit indices of the measurement model: root mean square of error approximation (RMSEA), goodness of fit index (GFI), comparative fit index, $\chi^2$/degree of freedom (RMSEA < 0.08, GFI > 0.90, CFI > 0.90, and CMIN/df < 3) (Table I). Also, uni-dimensionality is achieved, as all items have factor loading of $> 0.05$ (Solutions, 2013). In this study, two items were deleted as they did not meet the requirement. The individual construct with factor loading is shown in Figure 2.

4.2.2 Convergent validity. The following criteria satisfy the convergent validity: CR $> 0.7$, CR $> AVE$, and AVE $> 0.5$ (Hair et al., 2010). The $\beta$ value for all the items were found to exceed 0.7, while the AVE of the constructs were found to be greater than 0.5 (Hair et al., 1998).

4.2.3 Discriminant validity. Discriminant validity is achieved when no redundant item in the model or the correlation between pairs of latent constructs is $\leq 0.85$ (Hair et al., 1998). The inter-construct correlation was less than 0.85 and shown in Figure 3.

To further verify the discriminant validity, maximum shared variance (MSV) and average shared variance (ASV) for each of the constructs have been measured, and the measurement model was established to be valid, as both the MSV and ASV of all four individual constructs were lower than their respective AVE statistics (Hair et al., 2010) (Table II). Therefore, these analyses proved discriminant validity. Based on the un-rotated factor statistics, the percentage variance account by the single factor was found to be 46 per cent below the mark of majority (i.e. 50 per cent). This established the absence of CMV in our study.

4.3 Analysis of the structural model
The structural model is estimated by examining the path coefficients ($\beta$ weights), which indicate the strength of the relationship between the dependent and independent variables. From the five paths proposed in the model, four were found to be statistically significant and positively impact the predicted construct. The relationship between follow and connectedness was found to be significant with a little positive impact ($\beta = 0.43$, $p < 0.05$), therefore accepting H1. The relationship between follow and engagement was found to be significant with a medium positive impact ($\beta = 0.59$, $p < 0.001$), confirming H2. The effect of engagement to connectedness ($\beta = 0.66$, $p < 0.001$) and engagement to purchase Intention ($\beta = 0.87$, $p < 0.001$) was significant, which support H3 and H4, respectively.

<table>
<thead>
<tr>
<th>Indices</th>
<th>Recommended value (Hair et al., 2010)</th>
<th>Model fit indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFI</td>
<td>$&gt; 0.9$</td>
<td>0.902</td>
</tr>
<tr>
<td>CFI</td>
<td>$&gt; 0.9$</td>
<td>0.984</td>
</tr>
<tr>
<td>Chisq/df</td>
<td>$&lt; 5.0$</td>
<td>4.609</td>
</tr>
<tr>
<td>RAMSEA</td>
<td>$&lt; 0.08$</td>
<td>0.087</td>
</tr>
</tbody>
</table>

Table I. Model fit indices
Follow was a significant positive predictor of purchase intention ($\beta = 0.18, p < 0.05$), thus supporting $H5$. The structural model is shown in Figure 4. The hypothesis summary table is shown in Table III.

The results of the squared multiple correlations ($R^2$) shows that follow and engagement are able to explain 98 per cent of the variance in purchase Intention, which proves the
importance of fan-page follow and engagement in predicting the purchase intention. In turn, follow and engagement are able to explain 96 per cent of the variance in connectedness, while Follow explains 34 per cent of the variance in the follower engagement.

<table>
<thead>
<tr>
<th>Construct</th>
<th>CR</th>
<th>AVE</th>
<th>MSV</th>
<th>ASV</th>
<th>Convergent validity</th>
<th>Discriminant validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follow</td>
<td>0.794</td>
<td>0.510</td>
<td>0.396</td>
<td>0.304</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Engagement</td>
<td>0.962</td>
<td>0.660</td>
<td>0.562</td>
<td>0.497</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Connectedness</td>
<td>0.856</td>
<td>0.673</td>
<td>0.532</td>
<td>0.392</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Purchase intention</td>
<td>0.733</td>
<td>0.686</td>
<td>0.562</td>
<td>0.462</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table II. Reliability and validity for individual construct

\[ \chi^2=502.437 \]
\[ df=109 \]
\[ p=0.000 \]
\[ RMSEA=0.087 \]
\[ GFI=0.902 \]
\[ CFI=0.984 \]
\[ Ratio \chi^2/df=4.609 \]

Figure 4. Structural model

Path proposed (hypotheses) | Estimate | p-value | Supported
---|---|---|---
Follow→Purchase intention | 0.18 | 0.03 | Yes (H5)
Follow→Engagement | 0.59 | 0.00 | Yes (H2)
Follow→Connectedness | 0.43 | 0.02 | Yes (H1)
Engagement→Connectedness | 0.66 | 0.00 | Yes (H3)
Engagement→Purchase intention | 0.87 | 0.00 | Yes (H4)

Table III. Results of the Structural model and testing hypotheses
4.4 Exploring the mediating effect of engagement

We tested the significance of the mediating effects using bootstrapping procedures. Unstandardised indirect effects were computed for 10,000 bootstrapped samples, and the 95 per cent bias-corrected confidence interval was computed. The two tailed and standardised indirect effects were found to be significant ($p < 0.05$). The results show that the relationship between follow and purchase Intention was mediated by engagement. The standardised regression coefficient between follow with engagement and engagement with purchase intention was statistically significant. The standardised indirect effect was $(0.59) (0.87) = 0.51$. Moreover, the relationship between follow and connectedness was mediated by engagement. However, the standardised regression coefficient between follow with engagement and engagement and connectedness was statistically significant. The standardised indirect effect was $(0.59) (0.66) = 0.39$. The results showed that engagement serves as a partial and positive mediator that establishes a strong relationship between the follow and purchase intention, which confirms the acceptance of $H6a$, while engagement is a partial mediator that establishes a strong relationship between follow and social media connectedness, which confirms the acceptance of $H6b$. Thus, the result led us to conclude that engagement is an important mediating factor to enhance the relationship of follow to purchase intention and social media connectedness. The direct and indirect summary of this causal model is illustrated in Table IV.

4.5 Multi-group moderation analyses

Multi-group moderation analysis was conducted to determine the trust, gender, age, and income based variances in terms of hypothesised relationships as part of the proposed structural model. The significant moderating effects are evidenced by the hypothesised wise Z-score statistic. The outcomes of the analysis can be seen in Table V.

Respondents were asked to provide their positive or negative opinion about trust on their followed brand. About 66 per cent of the respondents provided a positive response on trusting their followed brand. To test the hypothesised moderation model in the SEM, two group models (trust–yes and trust–no) were tested. From the moderation analysis, trust is found as a significant moderator between the relationship of follow and purchase intention, confirming the acceptance of $H7a(i)$. Without trust (trust–no), follow have no significant impact ($\beta = 0.07$, $p > 0.05$) on purchase intention, while with trust (trust–yes), follow have a significant impact ($\beta = 0.14$, $p < 0.05$) on purchase intention. From the moderation test results, we also infer that trust, as a variable, moderates the relationship between engagement and purchase intention, which significantly supports the acceptance of $H7a(ii)$. The without trust (trust–no) engagement metric has little impact ($\beta = 0.32$) on purchase intention, while with trust (trust–yes) engagement metric shows a large impact ($\beta = 0.89$) on purchase intention. Findings indicating that follow and engagement activities in fan pages are largely affected by consumer trust on a particular brand to generate purchase intention.

As per the moderation analysis findings, $H1$ and $H5$ significantly differ across gender groups (males, females) confirming the acceptance of $H7b(i)$ and $H7b(ii)$. Among the

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Direct effect</th>
<th>Indirect effect</th>
<th>Result</th>
<th>Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement mediating follow to purchase intention</td>
<td>$0.18^{**}$</td>
<td>$0.51^{***}$</td>
<td>Partial mediating effect</td>
<td>Yes ($H6a$)</td>
</tr>
<tr>
<td>Engagement mediating follow to connectedness</td>
<td>$0.43^{**}$</td>
<td>$0.39^{***}$</td>
<td>Partial mediating effect</td>
<td>Yes ($H6b$)</td>
</tr>
</tbody>
</table>

Table IV.
Mediating effect

Notes: $^{**}p < 0.05; ^{***}p < 0.01$
| Hypothesis                          | Trust-Yes | Trust-no | Male       | Female     | Young | Middle | Adult | Employed | Unemployed | Z-score | |---------------------------------|------------|----------|------------|-----------|-------|--------|-------|----------|-----------|--------| | Est. | p     | Est. | p     | Zscore | Est. | p     | Est. | p     | Zscore | Est. | p | Est. | p | Zscore | Est. | p | Est. | p | Zscore |
| Follow→Purchase intention          | 0.14       | 0        | 0.07      | 0.2       | 1.1*** | 0.04   | 0.15  | 0      | -1.2**    | 0.07    | 2      | 0.19  | 0      | 0.05  | 1      | -0.9** | 0.18  | 0    | 0.08  | 2    | 1.02*** |
| Engagement→Purchase intention      | 0.89       | 0        | 0.32      | 0         | 4.5*** | 0.34   | 0.79  | 0      | 2.2**      | 0.51    | 0      | 0.79  | 0      | 0.47  | 0      | -1.73  | 0.78  | 0    | 0.27  | 0    | -3.2*** |

Notes: *p-value < 0.10; **p-value < 0.05; ***p-value < 0.01
respondents, 61 per cent of followers are females, and 39 per cent are males. In the case of female followers, a small impact ($\beta = 0.15, p < 0.05$) was found between the relationship of follow and purchase intention, while male followers did not show any significant relationship between the follow and purchase intention. In the relationship between engagement and purchase intention, the males showed a small impact ($\beta = 0.34, p < 0.05$), while females showed medium impact ($\beta = 0.79, p < 0.01$). Thus the findings indicate that female engaged followers are more likely to have purchase intention compared to their male counterparts.

In the age group moderation analysis (young (below 20), middle aged (20–40), adult (above 40)), a significant moderating effect is found between follow and purchase intention, which confirms the acceptance of $H7c(i)$. Moreover, no significant moderating effect has been found between the hypothesised relationship of $H5$, confirming the rejection of $H7c(ii)$. The respondent is 36 per cent young, 43 per cent middle aged and 21 per cent adults. Among the three age groups, the relationship between engagement and purchase intention, the middle-aged participants engaged followers are likely to have the most intensive purchase intention ($\beta = 0.79, p < 0.01$), followed by the young age group ($\beta = 0.51, p < 0.05$), and adults ($\beta = 0.47, p < 0.05$). In the context of the relationship between follow and purchase intention, only middle aged followers are likely to have purchase intention ($\beta = 0.19, p < 0.05$), while the young age group ($\beta = 0.07, p > 0.05$) and adults ($\beta = 0.05, p > 0.05$) lack any significant relationship between follow and purchase intention. The findings indicated that even though the young age group and adults follow fan pages, their purchase intention is not significant. The findings also confirm that all aged engaged followers possess purchase intention.

In terms of income (employed, unemployed), two groups differ significantly between the hypothesised relationship of $H1$ and $H5$, which confirms the acceptance of $H7d(i)$ and $H7d(ii)$. In total, 46 per cent of the respondents were employed, while 54 per cent were not.

The results show that the employed engaged followers are more likely to have purchase intention ($\beta = 0.78, p < 0.01$) compared to their unemployed counterpart ($\beta = 0.27, p < 0.05$). Moreover, the employed follower is more likely to have purchase intention ($\beta = 0.16, p < 0.05$), while the unemployed follower is more likely to lack any significant impact on purchase intention ($\beta = 0.08, p > 0.05$). Thus, the findings indicate income or employment status as an important moderating variable between follow and purchase intention and engagement and purchase intention.

5. Discussion
The main purpose of this study is to analyse the effect of following a company’s fan page on follower engagements, purchase intention, and connectedness. The results revealed that fan-page engagement have a strong impact on generating social media connectedness and purchase intention. Without engaging the users of the fan pages, fans will be little encouraged of getting involved in other sites suggested by particular fan pages. Moreover, followers who are less engaged on fan pages tend to exhibit less purchase intention. These findings are consistent with a previous study indicating that social media engagement infuses a sense of empowerment over social media activities (Kang, 2014). Marketers should engage clear engagement strategies specifying the appropriate content type, media type, and posting time to increase the level of engagement over moderator posts (Úblová, 2014). The results of fan-page engagement support others research findings on social media engagement in communication research, including Bruce and Shelley (2010), who called for a focus on engagement as “an umbrella term that covers the full range of an organization’s efforts to understand and involve participants in its activities and decisions” (p. 3). The results of this study revealed that even though Facebook “Likes” or follow is important, it only tell a part of the story, and marketers with a social media presence should focus on
brand engagement and sales lift as well. Marketers should simultaneously increase the number of followers and use strategy to increase followers’ activities in brand fan pages. Also, it is no longer enough to merely incorporate social media as standalone elements of a marketing plan (Richard et al., 2011), rendering the management of multiple online platforms a crucial issue.

This study also explored the moderating impact of trust, gender, age and income of fan page followers, and active users (engaged followers) on purchase intention. Without trust on a particular brand, fan page followers are discouraged from purchasing it. Although users might be engaged in a fan page, without trust, their purchase intention is also very low, as per (Grabner, 2002; Hoffman et al., 1999; Li and Kim, 2007). To establish online trust, it was suggested that a long-lasting virtual membership be created via continuous interaction with users (Ridings, 2002). The results also indicated that Malaysian female followers tend to exhibit purchase intention consistent with previous findings suggesting marketers adopt a feminine marketing policy to encourage female participation (Chetna et al., 2016).

Although previous factors indicated age as an important moderating factor in the context of online settings (Patricia et al., 2005; Ernst and Young, 2014; Wong, 2014), this study reveals that the purchase intention of active users (engaged users) do not vary according to age, which differs from previous study arguing that age do effect purchase intention, consumer decision making style, and consumption pattern (Safiek, 2009). This study contributes new findings in the context of Malaysian users on the moderating role of age in the sector of fan page. The results show that followers’ income is a significant moderator between the relationship of following fan pages and purchase intention and engagement and purchase intention, which agrees with (Elisabeta and Ivona, 2014; Narges et al., 2011). The study confirmed that female users are the best potential target segment for e-marketers.

5.1 Research implication

5.1.1 Managerial implication. This research provides new knowledge for B2C companies for utilising fan page based on user categories. This work thus recommends that marketers improve their engagement with followers and marketers and use real time and content marketing strategies to realise this (Chetna et al., 2016). Firms should not merely use this platform to push their product, they must hear from users and “on the moment marketing” to provide marketers with plenty of opportunities to engage their audience (Jahn and Kunz, 2014; Linjuan and Wan-Hsiu, 2013). The evidence showed that the engagement actions on the fan pages not only impacts the users’ social media connectedness, it also affects purchase intention. Thus, if marketers regard fan page as an important marketing tool and is integrated with users on regular/frequent basis to keep them engaged, it will result in a substantial increase of sales for a business. Simultaneously, interactivity with users will also lead them to engage with other pages suggested by the fan pages, which might create a new opportunity for marketers to divert and attract existing consumers to another business page or new products’ page. Marketers can use fan page as a cost effective and less time consuming tool to promote new products or pages to existing consumers and utilise this platform for managing multiple social media platforms. It will help companies and brand representatives redirect certain problems to other channels or social connected platforms and subsequently satisfy customers (Schultz, 2016). Marketers should move consumers from a state of pre-attention to elaboration. This is important for consumers who are typically low on involvement due to their basic personal traits. Increasing the content of the fan page via explainer videos, slideshow, infographics, instructional slideshows, specialised creative briefs, video and sample ads, plus sample briefs, pro tips and the extended linked slideshow (connected with other related sites) increases the richness of the information (Sangeeta, 2017)
and attract new/existing followers to the connected sites. Media considerations, with highly interactive information kits with music/video and brochures will evoke high involvement/emotional appeal in consumer that will lead them to purchase the product. Moreover, one important aspect that should be highlighted is that the customer engagement dynamic is changing (Lyle, 2015). In today’s world, customers are now creating image of marketers by checking reviews from other social sites and looking up reputation from blog sites. Thus, marketers should always be present in all of the channels, take part in the conversation, find and answer every negative review, thank every positive one, and respond to every question. There are many customers in interconnected channels (e.g. Facebook, Twitter, WhatsApp, Sky TV, and Mumsnet), and whichever matters most to customer should also matter most to marketers (Lyle, 2015). In the current engagement dynamic setting, customers’ total experience of marketers is what marketers are to their social channels. Thus, marketers should update their channels in order to maintain prospects in their database and should also make sure the linked apps are easy to add/remove. In current engagement dynamic situation, customers are interactional instead of transactional. Therefore, marketers should meet customers in all linked social channels as a trusted friend instead of a salesman.

5.1.2 Theoretical contribution. Fan-page engagement is regarded as an important mediator in the relationship between follow and social media connectedness and follow and purchase intention. Finally, this study developed a theoretical framework in the area of fan page by combining CEB and CIT. The study signifies that involving in fan page have a positive impact on fan-page engagement, social media connectedness and purchase intention which support the “response involvement”, “communication involvement” and “commitment involvement”, concepts of CIT. It also improves and supplements literature of CEB in the area of social media by linking the concepts of fan pages’ engagement activities with social media connectedness and purchase intention. This study also attempts to modify and validate the scale of Fan page follow, fan-page engagement, and social media connectedness. This study proposes the application of an empirically tested framework to the fan-page follow actions. We argue that this framework can provide a useful foundation for future social commerce research. The results would help academics be aware of Facebook and its fan page, which will provide a new avenue of research.

5.1.3 Limitation and prospects for future research. Although this study explored a new area, there are limitations that created new avenues for future research. The current study shows the impact of engagement actions on purchase intention. In the future, it will be very effective if the direct impact of each engagement action upon multiple behaviours is identified. We also explored the impact of follow and engagement actions on purchase intention, and did not divide the purchase intention into online and offline categories, and as social media drive both online and offline sales (Cao et al., 2014), both can be researched in more detail. In addition, future research can measure the moderator effect of trust, gender, age and income on social media connectedness since there is a lack of research in this area.

In this study, we selected different types of fan pages (fashion, restaurant, and electronics) in order to increase selectable varieties for followers. Therefore, it is also possible to explore the impact of social media engagement and follow actions on purchase intention based on business categories. This study was conducted on fan pages, and it would be very helpful if the same empirical study be conducted on other social media platforms (e.g. twitter, YouTube), since social media reach consumers at different phases of the purchase process (Fulgoni and Lipsman, 2015). Overall, the study serves as a basic fundamental guideline for academics and researchers to interpret the concept of following fan pages and engagement actions and its effects on purchase intention, as well as opening a vast area of unexplored researches on the subject of social media.
6. Conclusion
It remains unclear how social and cognitive theories can be generalised to online networks, which differ from real-world social networks (Kang, 2015). We empirically tested the fan page following actions of actual active online participants and reported their social media connectivity actions and purchase intention. We study user behaviour by determining the direct effect of user’s social media involvement action (follow) on their engagement action and purchase intention. We also explored the mediating impact of user engagement actions on their purchase intention and connectedness behaviour. We study the actual involved users’ or follower’s behavioural actions by applying a model that combines CIT and CEB theories, which provides a valuable source of information for marketers about the real-world social network. Thus, we bridged consumer involvement and consumer engagement actions to show their individual effect on media connectedness behaviour and purchase intention, which will help marketers pinpoint the importance of each selected variable and the mediating impact of user engagement. Besides, the moderating impact of users’ demographic profile will assist e-marketers in designing their respective target market segment.

References
Chen, L. (2014), The influence of Social Media on Consumer Behavior, Aarhus University, China.


Social Media”, The Swedish School of Textile, Sweden, available at: www.uppsatser.se/ 
uppsats/0f579c2e54/

engagement behavior: theoretical foundations and research directions”, *Journal of Service 

brand love and WOM outcomes”, *Journal of Product and Brand Management*, Vol. 23 No. 1, 
pp. 33-42.


in-2016

available at: https://doi.org/10.1108/JPBM-08-2015-0956

www.ecommercemilo.com/2014/01/e-commerce-infographic-malaysia-understanding-online-
shoppers.html


channel retailing”, *Multimedia Tools Applications*, Vol. 74 No. 1, pp. 159-178, doi: 10.1007/s11042-
013-1430-z.

Further reading

recommendations for getting the most from your analysis”, available at: http://pareonline.net/

Houston, M.J. and Michael, L.R. (1968), “A paradigm for research on consumer involvement”, 
working paper, University of Wisconsin, Madison, WI.

Maurer, C. (2016), *Effectiveness of Advertising on Social Network Sites: A Case Study on Facebook*, 
ResearchGate.

(The Appendix follows overleaf.)
Appendix 1. Respondents’ demographic data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Response Categories</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td>Fan-page</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Count</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>KFC Malaysia</td>
<td>63</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Lenovo Mobile</td>
<td>31</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Samsung Malaysia</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Zalora Malaysia</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>119</td>
<td>100</td>
</tr>
<tr>
<td>Female</td>
<td>KFC Malaysia</td>
<td>51</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Lenovo Mobile</td>
<td>43</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Samsung Malaysia</td>
<td>34</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Zalora Malaysia</td>
<td>60</td>
<td>32</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>188</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th></th>
<th>Fan-page</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Count</td>
<td></td>
</tr>
<tr>
<td>Young (&lt;20)</td>
<td>KFC Malaysia</td>
<td>29</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Lenovo Mobile</td>
<td>38</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Samsung Malaysia</td>
<td>32</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Zalora Malaysia</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>110</td>
<td>100</td>
</tr>
<tr>
<td>Middle aged (20–40)</td>
<td>KFC Malaysia</td>
<td>32</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Lenovo Mobile</td>
<td>43</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Samsung Malaysia</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Zalora Malaysia</td>
<td>30</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>132</td>
<td>100</td>
</tr>
<tr>
<td>Adult (&gt;40)</td>
<td>KFC Malaysia</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Lenovo Mobile</td>
<td>17</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Samsung Malaysia</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Zalora Malaysia</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>64</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Income</th>
<th></th>
<th>Fan-page</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Count</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>KFC Malaysia</td>
<td>44</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Lenovo Mobile</td>
<td>31</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Samsung Malaysia</td>
<td>25</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Zalora Malaysia</td>
<td>41</td>
<td>29</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>141</td>
<td>100</td>
</tr>
<tr>
<td>Unemployed</td>
<td>KFC Malaysia</td>
<td>56</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Lenovo Mobile</td>
<td>27</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Samsung Malaysia</td>
<td>23</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Zalora Malaysia</td>
<td>60</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>166</td>
<td>100</td>
</tr>
</tbody>
</table>
Appendix 2. Measurement items and reliability

Trust: (I trust the commitment of those brands that I follow in fan pages)
(1) YES (2) NO

<table>
<thead>
<tr>
<th>Construct</th>
<th>α</th>
<th>EFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan Page Following/Joining (F) (Adapted)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1: I joined to have direct interaction with the company.</td>
<td>0.768</td>
<td>0.716</td>
</tr>
<tr>
<td>F2: I joined fan page to get my queries solved.</td>
<td>0.712</td>
<td></td>
</tr>
<tr>
<td>F3: I joined fan page because I found those products unique*</td>
<td>0.789</td>
<td></td>
</tr>
<tr>
<td>F4: I joined to know more about it from the contents of fan-page.</td>
<td>0.823</td>
<td></td>
</tr>
<tr>
<td>F5: Joining fan page helps me to increase my self-status**</td>
<td>0.712</td>
<td></td>
</tr>
<tr>
<td>F6: I joined fan page to interact with people like me on this fan page*</td>
<td>0.716</td>
<td></td>
</tr>
<tr>
<td>Fan page Engagement (E) (Adapted)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1: I am an integrated member of this fan-page community**</td>
<td>0.960</td>
<td>0.964</td>
</tr>
<tr>
<td>E2: I am an engaged member of this fan-page community.</td>
<td>0.861</td>
<td></td>
</tr>
<tr>
<td>E3: I am an active member of this fan-page community.</td>
<td>0.921</td>
<td></td>
</tr>
<tr>
<td>E4: I am a participating member of this fan-page community.</td>
<td>0.932</td>
<td></td>
</tr>
<tr>
<td>E5: I am an interacting member of this fan-page community.</td>
<td>0.867</td>
<td></td>
</tr>
<tr>
<td>Fan page Connectedness (C) (Self-administered) conceptualized from (Chen, 2014; He lía, 2014)</td>
<td>0.917</td>
<td>0.798</td>
</tr>
<tr>
<td>C1: I like to “follow” other social platforms (e.g. twitter, other related fan-pages) suggested by my followed fan pages.</td>
<td>0.821</td>
<td></td>
</tr>
<tr>
<td>C2: I share the contents of fan page to other social platform through links of fan-page.</td>
<td>0.756</td>
<td></td>
</tr>
<tr>
<td>C3: I like to click on Brand discussion board and/or company website from fan pages.</td>
<td>0.743</td>
<td></td>
</tr>
<tr>
<td>C4: I often use same login id to logon different social media platform**</td>
<td>0.813</td>
<td></td>
</tr>
<tr>
<td>Fan page Purchase Intention (P) (Adapted) (Angella, 2016)</td>
<td>0.918</td>
<td>0.761</td>
</tr>
<tr>
<td>P1: I purchased the same brand when required as I like it on FB Page.</td>
<td>0.754</td>
<td></td>
</tr>
<tr>
<td>P2: I buy its products directly from FB when required.</td>
<td>0.802</td>
<td></td>
</tr>
<tr>
<td>P3: I intend to purchase this product/brand in future as well.</td>
<td>0.801</td>
<td></td>
</tr>
<tr>
<td>P4: I am loyal customer of the brand I “like” on Facebook.</td>
<td>0.791</td>
<td></td>
</tr>
<tr>
<td>P5: I would consider buying the promoted products on fan page.</td>
<td>0.745</td>
<td></td>
</tr>
<tr>
<td>P6: I consider my followed fan page information before buying.</td>
<td>0.756</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *Crobach’s z; †exploratory factor analysis. *Deleted after EFA; **deleted after CFA

Corresponding author
Zoha Rahman can be contacted at: joha_du@yahoo.com

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
Challenges for an SNS-based public sphere in 2016

Kyle John Lorenzano, Miles Sari and Colin Harrell Storm

Edward R. Murrow College of Communication, Washington State University,
Pullman, Washington, USA

Samuel Rhodes

Department of Politics, Philosophy, and Public Affairs,
Washington State University College of Arts and Sciences,
Pullman, Washington, USA, and

Porismita Borah

Edward R. Murrow College of Communication, Washington State University,
Pullman, Washington, USA

Abstract

Purpose – Political polarization and incivility manifested itself online throughout the 2016 US presidential election. The purpose of this paper is to understand how features of social media platforms (e.g. reacting, sharing) impacted the online public sphere during the 2016 election.

Design/methodology/approach – After conducting in-depth interviews with politically interested young people and applying deductive coding procedures to transcripts of the interviews, Dahlberg’s (2004) six normative conditions for the public sphere were used to empirically examine this interview data.

Findings – While some participants described strategies for productive political discussion on Social Networking Sites (SNS) and a willingness to use them to discuss politics, many users’ experiences largely fall short of Dahlberg’s (2004) normative criteria for the public sphere.

Research limitations/implications – The period in which these interviews were conducted in could have contributed to a more pessimistic view of political discussion in general.

Practical implications – Scholars and the public should recognize that the affordances of SNS for political discussion are not distributed evenly between different platforms, both for the sake of empirical studies of SNS moving forward and the state of democratic deliberation.

Originality/value – Although previous research has examined online and SNS-based political discussion as it relates to the public sphere, few attempts have been made understand how specific communicative practices or platform-specific features of SNS have contributed to or detracted from a healthy public sphere.

Keywords Political communication, Social media, Public sphere, 2016 US presidential election, Qualitative research methods

Paper type Research paper

Introduction

Increasing partisanship in the USA made the 2016 presidential election one of the most polarized in recent history (Doherty et al., 2016; Soergel, 2016). While free political expression is nothing new, the platforms to voice these opinions are changing. In 2016, nearly one-third of Social Networking Site (SNS) users sometimes or frequently commented, discussed or posted about politics. Additionally, the election coincided with the first time most Americans received their news from social media (Gottfried and Shearer, 2016), which is most pronounced among young adults between the ages of 18 and 29 (Mitchell et al., 2015). While SNS are often perceived as close-minded spaces for discussion and some areas of the web provide safe haven for hate groups (Duggan and Smith, 2016; Feldman, 2015; Southern Poverty Law Center, 2016), some are optimistic about their democratic potential (Gil de Zúñiga et al., 2012). Scholars like Bennett et al. (2011) reject narrow definitions of “political participation” in favor of less formal SNS-based participation they say is the future
of political involvement. These outcomes point to the greater potential of SNS to serve as online public spheres that foster productive political discussion.

Jürgen Habermas’s (1996) idea of the public sphere—a place to share opinions and news in an open environment—has transitioned from solely having implications for traditional media and face-to-face conversations, to impacting computer-mediated communication. SNS are evolving in their functionality and information sharing capabilities, with some existing as ideological “echo chambers” (Sunstein, 2001). While social media can be used to facilitate exposure to cross-cutting viewpoints (Messing and Westwood, 2012), they can also be used as tools for spreading misinformation to impressionable audiences of likeminded people (Bessi et al., 2015). These empirical findings call into question the quality of democratic engagement in the online public sphere.

Amongst all research on SNS-based political discussion, there is a gap in the literature regarding a specific evaluation of young adults’ communication practices on SNS and their relation to Habermas’s conception of a healthy public sphere. Scholars like Weller (2016) have called for more investigation into platform-specific features of SNS like “favoriting,” hashtags, or unfollowing, which has been explored previously (Boyd, 2010; Marwick and Boyd, 2011), but not in the context of a Habermasian public sphere. Prior qualitative research has also established that young people engage in sophisticated political talk, as Jahromi (2011) found in her qualitative work on how teenagers constructed their American identities. From a normative and empirical perspective, it is crucial that both scholars and the public better understand the underlying mechanisms of online deliberation that can make or break the quality of political discussion. With Dahlberg’s (2004) six normative conditions for the public sphere serving as a theoretical framework, this study aims to evaluate the democratic quality of everyday political communication among young adults on SNS during the 2016 US presidential election. Using in-depth qualitative interviews and deductive coding procedures as the primary means of analysis (Glesne, 2011; Miles et al., 2014; Saldaña, 2009), we find that some conditional aspects of SNS allowed for practices like ideal role taking and critiquing validity claims, but that ultimately participants’ discussions on SNS did not meet Dahlberg’s (2004) criteria for a Habermasian public sphere.

Habermas’s public sphere

With respect to democratic deliberation, no researcher’s influence has been more widespread than that of Jürgen Habermas. Habermas’s idea of a “public sphere” was initially conceived as a direct democracy-style space where citizens could engage in dialogue, which would then produce a coherent “public opinion” toward an issue(s) and influence decision-making practices (Habermas, 1989a, b). As Lunt and Livingstone (2013) argue, this conception of the public sphere was rooted in aspects of German and European bourgeoisie society—a feature of Habermas’s initial conception of the public sphere that would invite criticism from other scholars who alleged that members of the public could be excluded from this space based on race, sex, nationality and citizenship (Fraser, 2007; Dahlberg, 2004). Although Dahlberg (2014) provides a defense of Habermas’s original thesis, Habermas himself also addressed these criticisms (Habermas, 1996, 2006). This new work by Habermas moved away from a bourgeoisie conception of the public sphere to something more fluid and conceded that multiple forms/sites of deliberation were legitimate (Lunt and Livingston, 2013).

Habermas’s (1996) Between Facts and Norms outlines a new clarification of the public sphere that functions with the same working definition, where it is “[…] best described as a network for communicating information and points of view […][which are then] filtered and synthesized in such a way that they coalesce into bundles of topically specified public opinions” (p. 360). Habermas further clarifies his theory with a discussion of civil society, referring to all non-governmental and non-economic associations between people that allow for underlying communicative structures within a public sphere. The importance here of
civil society is that it is deeply enmeshed with the political public sphere and that any iteration of a public sphere can only be sustained with an engaged civil society. Habermas also addressed the changing mass media environment and its contribution to: making the public sphere more inclusive by expanding its reach; and making the idea of what constitutes a public sphere more abstract. He argues that the role of the press should be to reinforce an enlightened civil society, while simultaneously being independent of external political/social pressure and receptive to civil society’s concerns.

Habermas (2006) further clarifies this relationship between the media and the public sphere, citing two critical requirements: the independence of media from their social environments and free-flowing feedback between a self-regulated press and the public. Despite Habermas’s assertion that many mediated forms of deliberation amount to a weak public, he goes on to argue that these publics still have an impact on decision-making processes at elite levels, as evidenced by the rise of so-called “issue voters.” At the same time, Habermas points to market-based motivations and a relatively uninformed civil society as threats to his decades-long model of deliberation.

The internet as a public sphere
With the emergence of the internet, Habermas’s idea of the public sphere has migrated to online forums. To help foster healthy online political spheres, Bohman (2004) states the public sphere must promote free and equal dialogue where dissenting opinions can exist. Thus, the emergence of a global, online political public sphere has led some to theorize this will either increase or decrease heterogeneous political discussion (Plant, 2004). Whereas once people were exposed to those in their close personal network, the internet has allowed for individuals to gather perspectives globally and hear other worldviews (Papacharissi, 2002). However, there are significant concerns about this globalization of ideas, namely that individuals can now seek out likeminded people to constantly reinforce their beliefs.

This fear of homogenous political discussion has been widely studied within political communication literature. Sunstein (2001) argued that specific features of computer-mediated communication can ultimately destabilize open political dialogue, as individuals can self-select into online spaces made up entirely of like-minded arguments resembling something like an echo chamber. Within these echo chambers, individuals are more likely to share societal identities (i.e. political opinions), where individuals are then more vulnerable to group influence, gender typing, and stereotyping (Postmes et al., 1998). Indirect inequalities can also be procured from an ability, or lack of ability, to access the internet, resulting in an overrepresentation of some demographics over others (Baek et al., 2011). Despite this, many still see the merits in creating a place online for dialogue, with the possibility of increased opportunities for political participation, deliberation and information gathering in a transnational context (Bohman, 2004; Jennings and Zeitner, 2003; Wang et al., 2009).

Some research has found that some specific features of the internet also foster a more open dialogue (McKenna and Bargh, 1999). While some research has criticized anonymity online, others have found that a lower sense of social risks might encourage individuals to espouse dissenting opinions (Stromer-Galley, 2003). Likewise, online forums can enhance contributions from individuals who might be disadvantaged (i.e. social anxiety, dissenting views) and reduce the impact of being judged on the basis of race or gender (McKenna and Bargh, 1999). Due to the potential for increased involvement in political discussion, there is also a possibility for increased civic participation (Kavanaugh et al., 2005). Online news plays an important role in this relationship. Looking exclusively on sharing news on SNS, Ma et al. (2014) found that users who felt their networks were made up of close, likeminded ties were more likely to share news on their profiles. Others like Wells and Dudash (2007) have found that young people online who used the internet for news and information often
found it difficult to determine the credibility of information they consumed, and that even learning more about politics through credible information discouraged them from getting more involved in politics. However, when individuals seek out traditional news media to complement their online media use, the likelihood of a heterogeneous political discussion and civic participation network increases (Brundidge, 2010; Kaufhold et al., 2010; Shah et al., 2005). Other research has supported this idea that traditional media or even word-of-mouth sources of information complement these kinds of online information seeking habits (Parmelee and Perkins, 2012).

The public sphere in a SNS context
Habermas defines the public sphere as “[…] a domain of our social life in which such a thing as public opinion can be formed” (Habermas, 1989a, b, p. 1). These spaces ideally share what Freelon (2010) describes as Habermas’s “conceptual trio”: the notion that only rational-critical arguments are used when individuals contribute to discussion; that only topics of public concern are discussed; and that discussions remain open to all members of the public (p. 1181). Therefore, public opinion can be thought of as the coalescing of citizens’ opinions and thoughts, generated from dialogue, with the goal of influencing public policy. Habermas further articulates his conception by suggesting that face-to-face deliberation is necessary for considered public opinions – opinion changes based on quality information from discussion and objective argument (Habermas, 2006). This led to the consideration of “mediated” public spheres, where there is no face-to-face interaction, a lack of reciprocity between speakers/listeners, and the outsized gatekeeping role of the mass media. Specifically, Boyd’s (2010) essay on “networked publics” spoke to this ability of SNS to restructure publics entirely but provide the same basic functions of a traditional space for discussion. Rather than being restricted by physically disseminating messages between two parties, Boyd argues that networked publics constituted of digital “bits” allow for the dissemination of messages to be quicker and more diffuse. She goes on to explain that certain structural affordances like permanently digitizing exchanges (persistence), the ease of copying messages (replicability), the democratization of spreading information (scalability), and the relative ease of cataloging/searching for information (searchability) inevitably change the structure, but not necessarily the content, of these publics. The much-discussed phenomenon context collapse, another structural feature of SNS, refers to the clash of personal and professional ties in a public digital space where the appropriateness of personal or professional expressions vary (Marwick and Boyd, 2011).

Optimistic attitudes toward the role that SNS play in public discussion often convey the redefined power relationship between media producers and consumers that has characterized the diminished power of mass media in the late twentieth century (Williams and Deli Carpini, 2004). Benkler (2006) argues that the ease of communicating via SNS allow individuals to reorient themselves from passive audiences into active participants. Others envision SNS development as a means of facilitating “counter publics” where challenges to the status quo can be articulated and enacted openly (Jackson and Welles, 2015; Lee et al., 2015; Papacharissi and Oliveira, 2012). Public opinion scholars have found some empirical support for the role of “issue publics,” where citizens are only interested in a single political topic (Krosnick, 1998, Neuman, 1998; Krosnick et al., 1994; Converse, 1964; Almond, 1956).

Other scholars have been less sanguine about the emerging role that SNS platforms have had in public opinion formation. Jackson and Valentine (2014) highlight some of the challenges Twitter must overcome related to anonymity and a lack of physical presence. Others like Iosifidis (2011) argue that the democratizing effects of the internet are overstated, citing the chaotic and highly partisan discussion that tends to drown out the
rational-critical discussion that Habermas (1989a, b) envisioned. Iosifidis’s (2011) criticism of the internet also extends to its role as a profitable hub of corporate activity, a view Fuchs (2014) echoes, who argues that a critical, materialistic understanding of these platforms should serve as the foundation for SNS analysis.

Understanding the role that exposure to cross-cutting dialogue serves in the formation of rational-critical arguments on SNS remains a major focus of communication scholars. Previous research has found that SNS use alone has had a positive relationship with cross-cutting exposure (Kim, 2011; Burgess and Green, 2009) and that SNS use contributes to an individual’s network diversity (Kim et al., 2015). This diversity has been found to extend to the offline world as well, with Barnidge (2017) documenting how social media users perceive more political disagreement than non-users, especially among those who consume news on SNS. While Barnidge (2017) relied on self-reported measures of perceived political disagreement, Colleoni et al. (2014) used machine learning to find that Twitter can resemble an “echo chamber-like structure of communication” among certain partisan actors (p. 328). Other research indicates that the greatest opportunity for individuals to be exposed to cross-cutting political dialogue may include apolitical online groups where political discussion occurs incidentally (Wojcieszak and Mutz, 2009).

Though scholars have devoted attention to the benefits and detriments of online political discussion, little research has assessed the quality of everyday political talk within an SNS-based public sphere. While Dahlberg (2004) provides a set of six normative conditions to assess the public sphere-like quality of everyday communicative practices, these criteria have yet to be examined empirically. Consequently, there is an opportunity to draw upon Dahlberg’s framework to critically assess the quality of everyday political talk in the online public sphere, particularly among young adults, considering this population predominantly uses SNS for receiving information. In addition to its popularity amongst young people, previous research has demonstrated that using these platforms for information and discussion constitute new arenas for political engagement, participation, and involvement that are consequential to the political landscape (Gil de Zúñiga et al., 2014; Pang and Goh, 2016; Valenzuela et al., 2016; Zhou and Pinkleton, 2012). Not only will this introduce an overlooked theoretical framework to the analysis of democratic deliberation and discussion, such a starting point will allow us to examine and reflect upon everyday discourse practices on SNS, as well as suggest future avenues of research for quantitative scholars interested in assessing the quality and conditions of healthy SNS-based public spheres.

Dahlberg conceptualizes six public sphere conditions for evaluating everyday communicative practices. For Dahlberg, discussion must: allow for reasoned critique of problematic validity claims (i.e., claims about the truth); demonstrate that those involved critically examine their own values or positions (i.e., reflexivity); involve individuals that genuinely seek to understand opposing positions (what Dahlberg calls “ideal role taking”); prevent forms of deception or self-deception (i.e., be sincere); ensure that all relevant positions and people are included in the discussion (what Dahlberg calls “formal and discursive equality”); and be autonomous from state or corporate power structures. Though Dahlberg concedes that such a conceptualization is indeed an idealization for evaluating the democratic value of routine political talk, he argues that this framework may still be useful for assessing how it is “approximated in practice” (p. 13). As such, the following research questions have been posed:

*RQ1.* How do young adults’ everyday communicative practices on SNS promote a healthy online public sphere?

*RQ2.* What everyday communicative practices of young adults are possibly detrimental to the maintenance of a healthy online public sphere?

*RQ3.* How do the underlying tools associated with SNS platforms contribute/detract from a healthy online public sphere?
Methods

The present study relies on qualitative interview data, collected during the 2016 US presidential election. 30–45-min semi-structured interviews were conducted with students from a large university in the Pacific Northwest. Members of the research team used convenience and snowball sampling methods to recruit college students who expressed an interest in politics (Miles et al., 2014). Given the importance of SNS-based political engagement amongst young people (Bennett et al., 2011; Gil de Zúñiga et al., 2014; Zhou and Pinkleton, 2012), it was determined that this would be an appropriate population to sample from. In total, 33 participants were recruited, which was deemed to be the saturation point for this sample of participants (Corbin and Strauss, 2008). Informed consent was established in all interviews and participants were compensated for their time with extra credit or a gift card.

The average participant was 20 years of age, which is only two years younger than the average age of undergraduates at this specific university. In total, 15 males and 18 females were included in our sample. In all, 24 participants were white, 3 identified as “mixed,” 2 were African-American, 2 were Hispanic, 1 identified as Asian-American and 1 participant preferred not to answer. 13 participants identified as liberal/Democrat, 14 identified as conservative/Republican, 4 identified as independents and 2 identified as libertarians.

Interview questions from previous literature were adapted for the present study, which included broad questions on posting about political posts or discussion on the internet/SNS (Stromer-Galley, 2003), general social media habits (Wang et al., 2009), political issues that were most important to participants (Jahromi, 2011), social media vs the internet generally as a source for political information (Parmelee and Perkins, 2012), and participants’ most trusted source of political information (Wells and Dudash, 2007). In addition, unique questions were developed for this study to understand more about online political participation and young adults’ motivations for seeking information and discussing politics on SNS during the 2016 US presidential primary elections. The protocol for this study consisted of a variety of prompts, such as feelings questions, background questions, experience questions, opinion questions and grand tour questions (Miles et al., 2014; Glesne, 2011). The interview protocol with the relevant questions for the current study is included in supplementary materials as Appendix 1.

Once all interviews were conducted, the research team transcribed the interviews and assigned pseudonyms to each individual participant. Then, researchers followed inductive and deductive coding procedures outlined by Miles et al. (2014) and Saldaña (2009) to analyze the data. First, a sample of the interviews were inductively coded in order to create a preliminary codebook organized around the central concepts of the study. Then, to capture statements involving the processes and feelings associated with how participants used SNS as a public sphere-like environment for political discussion, Elemental and Affecting coding methods were used. Once relevant statements were identified, they were further deductively coded with Dahlberg’s (2004) six normative conditions of democratic communication in the Habermasian public sphere that were previously outlined. A summary of findings from the present study as they relate to these six conditions is provided in Table I. Following the coding of the interviews, researchers generated code categories based on Dahlberg’s six-pronged criteria, and from these categories, data themes were developed using a thematic analysis-style approach. Specifically, the following themes were identified: negotiating the consequences of SNS political expression, reluctance to welcome all relevant positions, and competition over consensus.

Findings

Negotiating consequences of SNS political expression

In a majority of the interviews, participants regularly emphasized their reluctance in discussing politics publicly on social media as a means of maintaining personal and
professional relationships. Often, they felt that any political expression on SNS would invite backlash from users in their network who did not share the same beliefs or disliked seeing political information on their newsfeeds. Amelia, a 20-year-old Democrat, expresses this same sentiment and her strategies in avoiding "family friction":

Amelia: My entire family outside of, like, my mom and grandpa are all extreme conservatives and we're, like, the only liberals of the family, so it really kinda causes some family disconnect […] since I got my new Facebook I didn’t – I haven’t really told any of my family members that I have it […] especially since it’s election season, I really could care less, um, of what their beliefs are verse [sic] mine and cause any family friction that isn’t already there […].

These kinds of tensions have been explored in previous literature and are often referred to as "context collapse" by some scholars, with the idea being that a range of diffuse social groups are collapsed into one (Marwick and Boyd, 2011). However, Dahlberg’s (2004) notion of ensuring that all relevant positions are included in discussion is in direct contradiction with the quote above and most of the interviews analyzed in the present study. For those who felt like online discussion resulted in their views being mischaracterized, strong negative backlash, or relational strife, there was a clear lack of opportunity in discussing politics as a means of generating cohesive public opinions (Habermas, 1996, 2006). In turn, this general avoidance of SNS-based political discussion and a recognition of its consequences presented major obstacles for all other aspects of Dahlberg’s (2004) criteria for a healthy public sphere.

The same chilling effect on discussion and deliberation was also found amongst a small, but notable group of participants who explicitly mentioned their fear of offending current or future employers with public political discussions. Close to a fifth of all participants corroborated this point about professional or academic affiliations preventing them from engaging with politics in a public setting online. Even some participants who enjoyed provoking users with “triggering” political statements recognized this need to be cautious, as 18-year-old conservative Fernanda stated in her interview:

Fernanda: […] we have to be very careful, because later on in the future when we proceed to higher education or, like, we try to pursue like a job at, like, the White House or any other government positions […] it will be kinda thrown back at us if we post something that people might find very offensive […] we have to be careful in what we post, because then, uh, they can use it against us later on in the future […].

The above quote has some implications concerning Dahlberg’s (2004) point about power structures – namely, these kinds of concerns over employment and maintaining one’s public image speak directly to the power of state and corporate forces in impeding deliberative activities.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Summary</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoned critique of problematic validity claims</td>
<td>Allowing for problematic truth claims to be debated amongst those potentially affected by the claims</td>
<td></td>
</tr>
<tr>
<td>Reflexivity</td>
<td>Individuals involved in the discussion critically examine their own values and positions</td>
<td></td>
</tr>
<tr>
<td>Ideal role taking</td>
<td>Genuinely seeking to understand individuals with opposing positions</td>
<td></td>
</tr>
<tr>
<td>Sincerity</td>
<td>Preventing forms of deception or self-deception</td>
<td></td>
</tr>
<tr>
<td>Formal and discursive equality</td>
<td>Ensuring that all relevant positions and individuals are included in the discussion</td>
<td></td>
</tr>
<tr>
<td>Autonomy from state and corporate power</td>
<td>Discussion is wholly independent of state or corporate influence</td>
<td></td>
</tr>
</tbody>
</table>

Table I. Summary of participants’ adherence to Dahlberg’s (2004) criteria for a public sphere
Although a majority of participants avoided SNS altogether as a space for political discussion, it is worth noting that several other participants claimed to use these platforms to engage in political discussion. Out of these participants, some claimed that Facebook often had a more public, permanent quality that comprised of a diverse network including family members, which was not an ideal environment to talk about politics. Instead, these users said they favored Twitter for discussion because they could maintain relatively anonymous Twitter accounts; tailor their Twitter network to users who did invite political topics; and feel safe knowing friends, family or employers were not aware they maintained a Twitter account. Other participants, like 19-year-old Democrat Derrick, explained that they used platforms like Snapchat because it, “[…] deletes itself so I can avoid any repercussions in the long run.” Several participants also mentioned using private group text or chat features as a safe place for discussing politics with close ties who either agreed with their positions or were known to disagree with their beliefs in a civil manner. Out of those participants who did use Facebook for political purposes, some claimed that they used the platform’s “Messenger” feature or private groups that only included these same kinds of close ties. These conversations were often characterized as an opportunity to engage in genuine dialogue about politics, even with those who could “agree to disagree” if necessary. In some sense, it is encouraging that these individuals use SNS tools strategically to engage in these conversations. Although a majority of participants did not discuss politics publicly on social media platforms, at least a third of all participants relied on private messaging or invite-only group features as a way to discuss politics. At the same time, several of these participants also mentioned that they use these techniques to only discuss politics with ideologically similar peers or family, which does not bode well for the principles of ideal role taking, reflexivity, sincerity and critically examining one’s own positions (Dahlberg, 2004).

Reluctance to welcome all relevant positions

As stated in the previous section, participants often avoided SNS altogether to prevent potential backlash from users in their networks. An altogether different issue was that a significant minority of participants regularly disdained the idea of talking to users with opposing political viewpoints on SNS or those outside of their network. It is not as though participants were entirely opposed to talking to those with opposing beliefs, but these discussions were often relegated to face-to-face discussions by a majority of participants. When asked why he does not engage with political posts on social media, 21-year-old Independent Ezekiel responded by saying:

Ezekiel: Because it just gets so heated. And it can get very serious very quickly […] On social media, like, I don’t really care what these random people have to say. And I don’t think that they really care what I have to say, either. So, I kind of just stay away from the discussion. I kind of just let them talk and just do my own thing.

Similar sentiments expressed by Ian, a 19-year-old Democrat, also point to how political affiliation can often be the deciding factor for whom to engage with in discussion:

Ian: I tend to engage less with strangers because usually back-and-forth online results in nothing except wasted time and maybe I’ll get upset, so I tend not to do that. But usually it’s in agreement or building upon the ideas of people who share similar ideas to myself on social media.

Both answers are representative of similar answers throughout these interviews which often referred to public SNS-based political discussion as “a waste of time,” “heated,” or unnecessarily stress-inducing in some way, which speaks to the inability of social media to foster a sense of discursive equality (Dahlberg, 2004). In tightly controlled environments with strong ties, participants could engage in a productive dialogue. However, when discussions were based around public political posts where anyone could participate,
some avoided comment sections that included uninformed “idiots” and shouting matches. Again, these sorts of pejorative statements directed at the “other side” do little to contribute positively to an environment where ideal role taking or reflexivity could realistically take place (Dahlberg, 2004).

Almost half of all participants reported seeing friends in their network, weak ties in their network, or strangers commenting on posts made by larger outlets dismiss those with opposing political views in real time. Both Democratic and Republican-leaning students pointed to conservative or pro-Trump positions receiving a large amount of backlash. Derrick, a 19-year-old liberal, gives one example of witnessing this in an online discussion group:

Derrick: Online there was a huge, huge debate between the representative of the liberal, Bernie Sanders’ point-of-view on immigration, and the Donald Trump, conservative guy. And while this conservative guy was civil and respectful […] the Bernie guy is probably not the best person to represent my candidate because I’m not a Bernie supporter. He was very aggressive online and brought in a professor onto online discussion […] And so he brought in friends on the event page in which there was slandering [sic] and, which, that’s the time I was moderating saying, “Please abstain from personal attacks […]”

This scenario would not allow for the kind of sincere, all-inclusive debate of wide-ranging validity claims that are essential to a well-functioning public sphere (Dahlberg, 2004; Habermas, 1996). Participants with varying political affiliations repeatedly explained that unpopular views were often met with ostracism and getting blocked or banned from pages. Other experiences demonstrate a genuine desire to engage in political discussion. Matthew, a 20-year-old libertarian who typically votes Republican, expressed a fondness for achieving the kinds of democratic ideals underlying Habermas’s vision of a public sphere:

Matthew: I like that it [political discussion online] presents the opportunity to see, you know, if you have an opinion, it presents you the opportunity to see that other side, and to kind of flesh out your ideas a little bit more with somebody else. It also, you know, sometimes you feel like it – or I feel like it – serves a higher purpose […]

Matthew was one of the few participants in our sample who admitted that political discussion on social media was an opportunity to learn about other perspectives and reflect on his own positions. Matthew’s sentiments were echoed by John, a 21-year-old fellow conservative, who described political discussion on Facebook as an opportunity to find clarification and learn about how other people are thinking about politics. As Dahlberg (2004) explains, rational communication in the public sphere requires that people be able to critically examine their own “values, assumptions, and interests in light of all other relevant claims and reasons” (p. 8). Discussion in the public sphere must rely on the willingness to think reflexively about one’s own opinions and possibly accept that there may be other opinions that are better. Thus, one of the results of reflexivity is that participants may end up “taking the position of the other” (p. 8), rather than limiting themselves to one point of view. Matthew’s statement that political discussion online “serves a higher purpose” provides an example of how the democratic ideals of the Habermasian public sphere are possible within the realm of SNS.

Despite Matthew’s penchant for reflexivity, he stated throughout his interview that his support for Donald Trump in the 2016 US presidential election was met with hostility online. Later in the interview, when he was asked if he used anonymous apps to discuss politics, Matthew responded with the following:

Matthew: […] one of the websites I visit is the politically incorrect board on 4chan, I don’t know if you’ve heard of it. It’s hard to take seriously sometimes, and I don’t know if it should be taken seriously a lot of the time, but it’s one of the, what’s credited as like the birth places of the “alt-right” […] I don’t know how to define it. It’s just where all the people who became disaffected with politics, the “alt-right” movement I’m talking about just became banded together I guess. I haven’t really spent much time researching it but, um, and there is an appeal to that […]
Here, Matthew describes the freedom that anonymity affords him when expressing his political opinions and references the “alt-right,” a far-right political movement that the Southern Poverty Law Center has categorized as a hate group (Southern Poverty Law Center, 2016). He goes on to reference the website 4chan and its “politically incorrect” message board (“/pol/” for short), a known repository of Nazi and white nationalist propaganda (Feldman, 2015). Nothing in Matthew’s interview suggests an endorsement of hate groups, but these sorts of groups and websites are in direct opposition to the kind of public sphere that has been described here (Habermas, 1996, 2006). If anything, Matthew’s response suggests that the normative consequences of feeling ostracized by one’s own SNS network are potentially dire.

It is worthwhile to consider the democratic consequences of SNS users who block individuals who express dissenting opinions online. While Matthew shared his experiences with being blocked, there was no clear consensus among our participants on blocking or unfollowing someone over a political disagreement. This divide in our participants’ responses points to a larger trend of blocking and unfollowing among social media users. According to the Pew Research Center, close to 40 percent of social media users say they have blocked, unfriended, or taken steps to reduce the number of posts they see from certain users in response to their political posts (Duggan and Smith, 2016).

Most of our participants also cited examples of people in their network who openly expressed that they would block or unfriend people who held different political opinions, saw examples of this happen in real time, or were blocked themselves because of their political affiliations. Consider the following example from Rachel, a 20-year-old conservative:

Rachel: I literally just saw this post yesterday. Somebody was posting, like, “I’m unfollowing anyone who is — shows like, admiration or is following Trump at all. If I see you like something I will block you.” Like, stuff like that. Like, maybe, like, let’s just consider everyone’s viewpoints.

For me, it’s like, why are we getting so upset about this?

Considering the hostility and polarization of the 2016 election, such reactions are, to an extent, understandable. Yet, it is difficult to ignore the severe democratic consequences of simply blocking opposing perspectives, as this behavior further prevents social media users from considering all relevant positions when forming and defending their political opinions. In terms of Dahlberg’s (2004) normative criteria, the steps that social media users take to block dissenting views has a direct impact on sincerity and potentially encouraging forms of self-deception, not to mention the aforementioned implications of a more homogenous network on discursive equality and ideal role taking. Since SNS users are free to block and unfollow dissenting views, such acts may deceive users into thinking that their political opinions are dominant positions, and that dissenting views are only held by an extreme minority of people. This is not to say that blocking and unfollowing are not appropriate responses in the face of online harassment or political trolling. However, the act of blocking and unfollowing dissenting opinions in the context of rational discussion creates a deceptive environment that hinders users’ “sincere efforts to make known all relevant information, including that which relates to their intentions, interests, needs, and desires” (Dahlberg, 2004, p. 9; Habermas, 2001, p. 34).

**Competition over consensus**

For the select group of participants who willingly engaged with political opponents on SNS, there was a sense amongst a significant minority that the goal of the discussion was to advance the cause of their own side, either by demonstrating that their side had the facts to back up their arguments or by purposely frustrating the other side. Others said they explicitly avoided these discussions because they devolved into a zero-sum game, with some like 19-year-old moderate Alexia calling it a “lose-lose” scenario. Echoing some of the
statements made by Fernanda about “triggering” opponents, 22-year-old conservative Rebecca expressed a similar fondness for frustrating political opponents:

Rebecca: And then also, um, there’s just this really extreme, annoying person on Facebook that posts political things and I think it’s just hilarious to engage with her, ‘cause I think it’s funny to […] get her going […]

Some like Alexia, a 19-year-old moderate, express that they avoid political discussion out of a fear that they will “lose” publicly:

Alexia: […] I feel like if I did get into some type of Facebook altercation, my goal would to not be political per say, because I feel like when you get political you’re making arguments that aren’t winnable, and I have a problem with wanting to be right, so I feel like if I got in a political discussion on Facebook I would try to make it more factual than political just because I wouldn’t want to lose in front of all of those people […]

Such goals for engaging in political discussion online only further reveal that there is a lack of critical self-reflexivity in the SNS public sphere. Several participants used SNS strategically to avoid these public discussions where “winning” was emphasized over consensus, with certain trusted friend groups or private Facebook pages acting as places where all positions were critically examined (Dahlberg, 2004). However, competitive political discussion could also erupt within these spaces. As discussed earlier, a lack of reflexivity prevents individuals from examining their own positions and possibly transcending their original opinions. However, when the goal of political discussion online is to merely compete with others, the democratic potential of deliberation and political discussion in the SNS public sphere erodes. Rather than attempting to come to a consensus or at least understand other perspectives, it seems that “winning the argument” is the primary motivation for political discussion online.

As Alexia admitted, there is a fear of losing a political argument online, and by extension, a fear of being wrong in her opinion publicly. Such an approach to online deliberation ultimately prevents people from coming to terms with the fact that they might be wrong or that it is acceptable to make mistakes in a political discussion. This competitive framework only prevents users from engaging in reflective dialogue, and as a result, may even fortify a deceptive SNS environment where the perception is that the individual user is always correct.

At the same time, certain tools on Facebook like the “react” button allowed for opposing factions of a political argument to snipe at each other publicly, as Rebecca goes on to explain:

Rebecca: Yeah, I think the thing that gets me […] you know how you can “like” somebody’s comment? When people – so, let’s say you post something […] this happened to me a long time ago, and I commented on something […] and a person underneath said some snarky, sarcastic comment, and then if somebody likes it, that’s what really bothers me […] I feel like the liking thing on Facebook can cause a lot more controversy because you can see the people that like it […] it’s like, “Aw, that pisses me off!”

Later in the interview, Rebecca described this technique as “passive aggressive,” which was echoed by a handful of other participants. With several participants describing SNS-based discussion as “arguing” in a pejorative sense and select participants wanting to tear the other side down, rational-critical, issue-focused deliberation seems to be lacking in the social media experiences of many participants in this study (Freelon, 2010). As a whole, we find that these participants are not able to engage in Dahlberg’s (2004) conceptualization of ideal role taking or critical self-reflexivity, either because they avoid discussion entirely or actively attempt to tear the other side down as a means of entertainment.

A minority of participants said they were able to engage in some form of “reasoned critique of problematic validity claims” (Dahlberg, 2004). Specifically, out of those
participants who openly discussed politics publicly on social media, many expressed that one of their primary motivations for engaging in SNS-based political discussion was to correct falsehoods. Citing Habermas (1984), Dahlberg (2004) explains that, “engaging in argumentation requires the thematization and reciprocal rational testing of problematic validity claims” (p. 7). Furthermore, Dahlberg defines validity claims as statements that are always “explicitly or implicitly raised in communication and include appeals to the meaning of statements, the truth of propositions, the rightness of norms, and the truthfulness of expressions” (p. 5). Thus, the reasoned critique of problematic validity claims requires an identification and evaluation of any statements that include troublesome facts, definitions, and truth claims. Ultimately, Dahlberg argues that this condition is necessary for the formation of rational public opinion. While this critique relies on SNS users’ willingness to examine their own positions, it also rests on their inclination to identify and critique others’ positions and validity claims.

Though our participants’ experiences during the 2016 US presidential election did not contain much critical reflection occurring within the SNS public sphere, there was a sense that many of these participants’ motivations for discussion were fueled by a desire to identify and correct blatant falsehoods. Consider the following examples from Ken, a 19-year-old liberal, and Eric, a 22-year-old moderate conservative:

Ken: I guess the thing that would probably motivate me the most [to comment on a discussion] is if somebody is posting something that is extremely inaccurate or really in opposition to something that I believe in.

Eric: I have a deep-seated dislike of the misconstruing of facts, and especially when doing so propagates some sort of injustice. And so, if I see someone whose political opinion is being deliberately or otherwise misconstrued in a political discussion, and they are suffering under the onslaught of that, I sometimes – I think I’ve commented to come to the defense of somebody.

Both Ken and Eric’s sentiments are illustrative of our participants’ desire to correct what they perceive in their networks as factually incorrect or misconstrued information. Considering these individuals expressed an appreciation for politics, such motivations for engaging in political discussion within the SNS public sphere are arguably atypical when compared to the average millennial. However, what is most unique about Ken and Eric’s actions is that they seemed to be motivated by extreme circumstances of misinformation or misconceptions. Specifically, Eric’s motivations for defending people in a discussion whose views are being misrepresented seem to require a high threshold. While it is reassuring that these participants are willing to intervene in situations where validity claims are especially problematic, it is concerning that these responses are only triggered by extreme circumstances, and such motivations still may be fueled by a competitive desire to win an argument.

**Discussion and conclusion**

Using Dahlberg’s (2004) six normative conditions, this paper used in-depth interviews with politically motivated young adults to understand the role of SNS in fostering a healthy public sphere during the 2016 US presidential primaries. Certain features of these platforms allowed for rational-critical deliberation to take place, especially those that allowed for private or group-based discussion. In these private spaces, participants were able to engage in sincere discussion with trusted ties who often varied widely in their ideological beliefs. Relegating this kind of political discussion to private chats/groups could potentially constitute one of the many “spheres” Habermas declared as legitimate in his later works, as well as shield these discussions from corrosive power structures (Lunt and Livingston, 2013; Dahlberg, 2004). Some participants also mentioned that platforms like Facebook allowed them to gauge public reaction to issues or
be exposed to counter-attitudinal views, further contributing to thoughts resembling something like reflexivity (Dahlberg, 2004). Others also recalled coming to the defense of users in their network whose positions were mischaracterized. In some sense, these participants responding to a user’s opinions being misrepresented could be construed as a willingness to challenge problematic truth claims, and perhaps even a kind of ideal role taking that would not have been possible without these SNS-based affordances.

At the same time, many interviews revealed that participants’ use or aversion to SNS contrasted sharply with Habermas’ public sphere. These participants admitted to only discussing politics with like-minded users and blocking/unfollowing users to shield themselves from counter-attitudinal content. Along with many participants who echoed concerns over the consequences of public political statements, this reluctance to engage with a range of users in one’s own network is in direct contradiction to Dahlberg’s (2004) “formal and discursive equality” requirement. An analysis of these interviews also demonstrated that this lack of welcoming all relevant positions was partly asymmetric, as many participants commonly recognized conservative alienation throughout their newsfeeds. In one instance, this resulted in a conservative participant showing a fondness for more anonymous and openly hostile corners of the internet, which both speaks to the normatively undesirable consequences of this kind of alienation and how the relative anonymity of certain platforms over others functions as a tool that shapes the confines of SNS-based political discussion. For those who willingly engaged in SNS-based political discussion, these exchanges presented a problem for participants who did not want to lose arguments or appear ill-informed in a public setting. In turn, some participants described conversations where the goal was to save face at all costs, presenting obvious problems for Habermasian goals of reflexivity, the avoidance of self-deception, and ideal role-taking.

As with any study, this paper has its limitations. First, the present study does not seek to make any definitive causal arguments for the relationship between SNS use and a healthy public sphere. More work should be done to formally tests that kinds of associations alluded to in these finding here. There is also something to be said for the period in which these data were collected. It could be argued that the results here paint an overly pessimistic view of online political discussion by students, which could be the result of a sampling bias across the participants recruited for this study. However, given that much has been said in the news media about the divisiveness of this election (Soergel, 2016), it is equally plausible that these participants are reflective of the current mood of young people as it relates to political discussion. Is it possible that the most recent election could be a foreshadowing of future mid-term or general elections that will continue along this path of divisiveness? Future studies should attempt to learn more about how specific SNS and their affordances are being used by individuals to accelerate or mitigate this trend of divisiveness.

Despite these limitations, these results point to an overarching sense that many young people in our sample were displeased with online political discussion.Few described SNS-based deliberations as something close to all of Dahlberg’s (2004) normative conditions for the public sphere. Others who were more politically engaged or afraid of public criticism seemed to relish the opportunity to engage in political discussion online, sometimes for the sake of learning from other or defending them and other times to prove opponents wrong publicly. If anything, these interviews suggest that SNS functions neither as a panacea for or the death knell for a healthy SNS-based public sphere. Scholars should attempt to understand the myriad ways that individuals use SNS for political discussion and determine the best set of practices possible for creating a Habermasian vision of a public sphere.
References


Burgess, J. and Green, J. (2009), YouTube: Online Video and Participatory Culture, Polity, Cambridge.


Glesne, C. (2011), Becoming Qualitative Researchers: An Introduction, Pearson, Boston, MA.


Saldaña, J. (2009), The Coding Manual for Qualitative Researchers, Sage, Los Angeles, CA.


Further reading

Davis, R. (2005), Politics Online: Blogs, Chatrooms, and Discussion Groups in American Democracy, Routledge, New York, NY.


Appendix 1. Online supplementary materials

Interview protocol

1. Describe your social media use in a typical day, from the morning to the night.
2. What do you typically share on social media? Is it personal information, political information and opinions, news, photos?
3. How frequently do you post political information on social media?
4. What motivates you to discuss politics on social media, specifically share your political viewpoints?
5. With whom do you typically discuss politics with online?
When are you most likely to comment on others’ posts in your social media networks?

How likely are you to express disagreement in an online political discussion?

How does that compare to how likely you are to express disagreement in a face-to-face discussion?

How are political discussions different for you in comparison to face-to-face discussion?

How often are you likely to engage with someone in your network that holds different political opinions than you?

Corresponding author
Kyle John Lorenzano can be contacted at: kyle.lorenzano@wsu.edu
The perception of e-servicescape and its influence on perceived e-shopping value and customer loyalty

Ana Cuic Tankovic
Faculty of Tourism and Hospitality Management, University of Rijeka, Opatija, Croatia, and
Dragan Benazic
Faculty of Economics and Tourism “Dr Mijo Mirkovic”, Juraj Dobrila University of Pula, Pula, Croatia

Abstract
Purpose – The popularity of online shopping has grown in the recent years, enhancing the interest in identifying the factors that influence shoppers’ loyalty. The purpose of this paper is to investigate, through aesthetic appeal, layout and functionality, and financial security, the impact of customers’ perceptions of the e-servicescape in online shopping on perceived e-shopping value and customer loyalty defined as a two-component construct (attitudinal loyalty and behavioural loyalty), following the sequence of the S–O–R framework.

Design/methodology/approach – Data were gathered using an Internet survey in which 221 active online shoppers participated who had at least one online purchase in the past six months. The research model was analysed using the partial least squares approach to analyse structural equations (SEM).

Findings – Based upon empirical evidence of a web-based survey of online shoppers, this study shows that consumers’ interpretation of e-servicescape exerts a positive influence over perceived e-shopping value and loyalty. Specifically, the study finds that layout and functionality and financial security, as the salient attributes of the e-servicescape dimension, set the point of perceived e-shopping value which affects attitudinal loyalty.

Practical implications – The authors expose several practical implications how marketing management could use the dimensions of e-servicescape perceived e-shopping value. Managers are advised to incorporate the dimensions of the e-servicescape to build perceived e-shopping value and loyalty. Online shops need to invest in particular in layout and functionality and financial security, dimensions that had a stronger effect on perceived e-shopping value in order to achieve costumers’ loyalty in a fashion product context.

Originality/value – The transition from physical service locations to virtual service processes has increased the research interest of e-servicescape. This study analyses it through its aesthetic appeal, layout and functionality, and financial security connecting them with the constructs of perceived e-shopping value and customer loyalty in the application for fashion products. This work emphasises the connection between servicescape attributes in creating perceived value which affects attitudinal and behavioural loyalty. In this way, the e-servicescape is seen as stimuli, the perceived e-shopping value represents the organism and the loyalty the response of the sequence of the S–O–R framework. This research provides insight into the new effects of each dimension of e-servicescape on perceived e-shopping value and the two dimensions of loyalty, thus improving the existing knowledge in the field of servicescape and online shopping customer behaviour.

Keywords Customer loyalty, Internet shopping, S–O–R, E-servicescape, Perceived e-shopping value

Paper type Research paper

Introduction
Technological developments have created new channels for services and products, where the internet offers many advantages for organisations interested in promoting and selling their services. The internet is a powerful merchandising tool; services offered by websites are constantly growing, and online media marketing communication is becoming the most important and fastest-growing promotion strategy in this century (Dong et al., 2017). The network serves as a channel both to disseminate information and to cover the...
market which remains uncovered. For these reasons, hundreds of millions of dollars have been invested by the business sector in efforts to establish an electronic presence on the internet (Vilnai-Yavetz and Tifferet, 2009). In general terms, online shopping is a process which involves a user accessing the internet in order to search, select, buy and review goods and services to satisfy his or her needs. In Europe, e-commerce has rapidly increased reaching “an average annual growth rate of 22%, surpassing €200 billion in 2014 and reaching a share of 7% of total retail sales” (Marcus and Petropoulos, 2016). The fact that many business activities are carried out using computers and the internet reveals the core issue of how internet businesses can make themselves the most trusted and shopped websites of internet customers (Lin and Sun, 2009). The development of online shopping has enabled people to buy anything at any time, opening up new opportunities for customers, who can easily compare product characteristics and prices, thus making it the most flexible way of purchasing. According to Soto-Acosta et al. (2014), customer purchase intention can be affected by information overload during online shopping. Therefore, the importance of investing every aspect of web identity and web services in online shopping has been recognised, because clients’ perceptions of a company and its character are likely to be derived from their perceptions of its website. Thus, it is important to understand how the design of e-commerce and e-service systems affects customer reactions and customer loyalty (Gao, 2005; Mummalainen, 2005). This paper seeks to explore the transition from physical service locations to e-service processes in the nature of the “servicescape” (Bitner, 1992) that customers encounter, and to understand perceived e-shopping value and customer loyalty. The servicescape is one of the leading factors in customers’ perception of value in physical settings, so the relationship between physical servicescapes and perceived e-shopping value and customer loyalty has been documented in the literature (Babin et al., 2004; Hightower et al., 2002; Harris and Ezeh, 2008). On the other hand, the transition of physical servicescape to e-servicescape is worthy of further investigation. Even if a consistent number of studies have been conducted in the field of e-servicescape, the connections between e-servicescape with the perceived e-shopping value and loyalty have been underexplored. This study examines the effect of e-servicescapes on perceived e-shopping value and customer loyalty, and may shed new light on the transition of the buying process from physical shop to online shop. Prior studies have demonstrated the relationship between e-servicescapes and e-satisfaction (Zymanoski and Hise, 2000; Ballantine, 2005), later shopping behaviour (Menon and Kahn, 2002), pleasure (Eroglu et al., 2003), trust (Papadopoulou et al., 2001), trust and purchase intentions (Harris and Goode, 2010), and trust and word-of-mouth (WOM) (Tran and Strutton, 2016). This paper’s contribution is in the analysis of all three dimensions of servicescape on the perceived value and, consequently, on the multidimensional construct of loyalty. Despite the agreement that a more comprehensive explanation is needed in order to better understand customers’ perceived e-shopping value and loyalty, no study so far have examined this concept with the complete dimensions of the e-servicescape suggested by Harris and Goode (2010).

Therefore, in order to overcome this gap, this paper aims to improve the existing knowledge to enhance the understanding of e-servicescape, following the stimulus–organism–response (S–O–R) framework: environmental stimuli, consumer’s inner organism and behaviour responses. In this research, the e-servicescape manipulation represents the “stimuli” and is examined in each of its dimensions, the “organism” part is represented by the perceived e-shopping value, and the “response” part is customer loyalty. This approach contributes to the marketing service sphere, analysing customer loyalty as a response and a multidimensional construct advancing the understanding of customers perceived e-shopping value and loyalty starting from e-servicescape. Filling this gap The moderating variables between e-servicescape and other constructs are tested not only for gender, age and income, but also for the educational background. Nowadays, the digital technology allows new interactions between customers and
suppliers, so testing the dimension of the e-servicescape affecting the perceived e-shopping value opens up new important insights of customers' perception in the changed digital and social environment. This relationship has not yet been investigated, and its meaning in creating customers' loyalty is a challenging task to understand and implement in the online businesses. The present study involves research that is of interest to website marketing managers working in a global environment.

To ensure the success of research efforts, the focus was placed on fashion products, a rapidly emerging market branch and the second most popular products among online purchases (Kawaf and Tagg, 2012).

This paper is divided into five parts: after this introduction, the second part provides the literature review regarding the constructs of perceived e-servicescape, perceived e-shopping value and customer loyalty. The methodology of the research is then explained, followed by the results. In the closing remarks, which make up the fifth part, the academic implications are discussed, as well as managerial implications, research limitation and some suggestions for future research.

Literature review and hypotheses development

Fashion shopping

Some of the many obstacles to purchasing fashion products online are the impossibility of trying on a product, seeing its quality before buying it and consulting with helpful staff (Kawaf and Tagg, 2012). New technologies have to offer ways to overcome these obstacles to online shopping. According to Ha et al. (2007), there are three visual merchandising features of the traditional offline store that have to be implemented in online apparel websites: online path finding assistance (search engines, site maps and categorisation), environment atmospherics (music, videos, display, background colours and colours surrounding the products) and manner of product presentation (view and display method, colour and methods of presentations, detailed views, swatch and mix and match). Another dimension of fashion shopping is that this shopping is not limited just to spending money on products, but rather shopping is also an important socialising and engaging exercise that provides opportunities to be with others (Kang and Park-Poaps, 2011). The social dimension of the online shopping experience gives growing attention to the significance of social network sites, virtual communities and customers review (Kim and Gupta, 2012).

According to Blázquez (2014), the fashion industry has been slower to adopt e-commerce than other sectors, but the internet has transformed the fashion industry and the retail environment with intense competition. The innovation in digital technologies has allowed the multisensory inputs in the online environment, so the e-servicescape has an important role in translating the atmospheric cues in the virtual store.

$S\rightarrow O\rightarrow R$ model

The $S\rightarrow O\rightarrow R$ paradigm suggests that in any environment three basic emotional states, known as PAD, mediate approach-avoidance behaviours (Mehrabian and Russell, 1974). The first empirical test of the effects of the retail atmosphere was provided by Donovan and Rossiter (1982) based on environmental psychology. Subsequent work in this area focussed on the impact of specific atmospheric cues operationalized as the “stimuli”, shoppers' emotional reactions as the “organism” and approach/avoidance behaviours as the “response” (Eroglu et al., 2003). The $S\rightarrow O\rightarrow R$ model was developed for consumers' purchases of items in offline stores (Buckley, 1991). The studies in traditional store environment has investigated the influence of the servicescape on customers' expectations, cognition and emotions, but with the emergence of the internet and the growth of online shopping, researchers focussed on various aspects of the $S\rightarrow O\rightarrow R$ model in this new medium (Kawaf and Tagg, 2012; Peng and Kim, 2014). This paper presents the e-servicescape,
a consolidated scale for online environmental cues, further explained in the following section, as external stimuli (S). Online environmental cues affect the customers’ emotions and intentions (Koo and Ju, 2010), and a significant relationship exists among web aesthetics, online shoppers, perceived service quality and satisfaction (Wang et al., 2010) and consumers’ flow experience and positive affect (Huang et al., 2017). Peng and Kim (2014) operationalized the stimulus as an internal influence (hedonic and utilitarian shopping value) and external influence (environmental stimuli). Previous researchers found that real and online store stimuli elicit both emotional and cognitive responses within organisms (Ha and Im, 2012). The organism, represented as an intermediary state and process that mediate the relationship between the stimuli and the response (Hsin Chang and Chen, 2008), is analysed as a cognition because the perception in the online shopping environments is more cognitive than in offline environments (Demangeot and Broderick, 2007). In the S–O–R paradigm, avoidance or approach behavioural outcome variables were operationalized as the response behaviour or behavioural intention (Ha and Lennon, 2010). The extension of the S–O–R model is widely presented by exploring the satisfaction variable in the response field (R) (Ha and Lennon, 2010), extending it from cognitive and emotional reactions that occur in the organism (O) such as intention to WOM (Ha and Im, 2012).

This research investigates the behavioural and attitudinal loyalty as a response (R), extending the applicability of the S–O–R paradigm into customer loyalty in the online context. The perceived e-shopping value is a mediator between the perceived e-servicescape (S) which affect customer loyalty as response (R). Therefore, the S–O–R model is the most appropriate theoretical framework for this study because it aims to explore the relationships between these three dimensions.

E-servicescape

The services marketing literature labels an organisation’s service setting as a “servicescape” (Bitner, 1992), the set of tangible, physical cues that represent an organisation to its clients. Zeithaml et al. (2002) define servicescape as the environmental cues that affect customers impersonally. According to Bitner (1992) servicescape comprises ambient conditions (temperature, air quality, noise, music, odour [...]), space/function (layout, equipment, furnishings [...]), and signs, symbols and artefacts (signage, personal artefacts, style of décor [...]). Kauppinnen-Raisanen et al. (2014) cite different divisions of servicescape elements: exterior, general interior, store layout, interior displays and human variables (Turley and Milliman, 2000); physical, social, social-symbolic and natural (Rosenbaum and Massiah, 2011). While advances have been made in conceptualising offline servicescapes, the online context has been comparatively neglected because there are not enough studies and surveys into the online buying environment to keep abreast of the growing importance of the internet as a sales channel for numerous retailers. Since this research field is designated as one of the leading emergent areas for future studies (Mari and Poggesi, 2013), more studies have to shift their attention from physical servicescape to the virtual one, referred to as an “e-scape” (Koernig, 2003), “cyberscape” (Williams and Dargel, 2004), “online servicescape” (Harris and Goode, 2010), “virtual servicescape” (Vilnai-Yavetz and Rafaeli, 2006; Mari and Poggesi, 2013), “digital servicescape” (Ballantyne and Nilsson, 2017) and “e-servicescape” (Hopkins et al., 2009; Huang et al., 2017; Teng et al., 2018). The conceptualization of e-servicescape is presented by Harris and Goode (2010) as comprising three dimensions: aesthetic appeal, online layout and functionality, and financial security. In this way, starting from the original work of Bitner (1992), two dimensions can be broadly equated, while the third dimension (signs, symbols and artefacts) is replaced by financial security, in order to stress this more important element in an online environment.

The design of the virtual site that customers encounter in an e-service setting can be defined as the designing of web environments to create positive effects in users in order to
increase favourable consumer responses (Dailey, 2004). Internet web pages as an electronic servicescape have marked their significant presence in the business world and are predicted to continue to gain popularity (Lai et al., 2014). Previous studies have proven the importance of environmental cues influencing website trust and purchase intention (Harris and Goode, 2010), trust and relationship building (Papadopoulou et al., 2002), customers’ cognitive and affective responses as well as specific shopping behaviours in terms of both approach and avoidance outcomes (Eroglu et al., 2003). The connection between offline servicescape variables and loyalty intention was conceptualised in the model by Harris and Ezeh (2008) and later discussed in an online context by Allmér (2014). Empirically, it was researched in travel and tourism websites (Sreejesh and Ponnam, 2016), to which the authors suggested the study should be extended to cover different services, and this paper is proposed in the fashion shopping field.

The three sub-dimensions of the e-servicescape (Harris and Goode, 2010) were used for testing the online shopping intention and purchasing experience (Wu et al., 2016), customer evaluation of the website and shopping behaviour (Adiwijaya et al., 2016), website trust in airline industry (Kühn et al., 2015), and consumers’ flow experience and positive affect (Huang et al., 2017). However, these studies reflect only partially the importance of the environment on the customer because no research yet applied the multidimensional scale of e-servicescape for investigating the perceived e-shopping value and loyalty in the S–O–R model. In addition, this paper seeks to explore each dimension of the e-servicescape in order to understand which dimension of the online servicescape is the best predictor for perceived e-shopping value and loyalty.

**Perceived e-shopping value**

Consumers’ perceived e-shopping value is the core element in a relational exchange, and it is the perception of what expected gains and losses are in the repurchase process (Wu et al., 2014). According to Zeithaml (1988), perceived e-shopping value is defined as the consumer’s overall assessment of the utility of a product based on perceptions of what is received and what is given. Customer perceived e-shopping value has received considerable attention in the field of marketing strategy because it achieves sustainable competitive advantages (Lindgreen and Wynstra, 2005), has an important role in predicting purchase behaviour (Chen and Dubinsky, 2003) and affects relationship management (Payne et al., 2001). A valuable exchange of the salient sacrifice/give and benefits/get components is an important argument not only in offline environments but also in online ones. Customer perceived e-shopping value in the online shopping environment is of crucial importance and that is why it is necessary to understand the role of customer perceived e-shopping value in online shopping behaviour (Hsin Chang and Wang, 2011). Existing research does not analyse the multifaceted concept of value, and Harris and Goode (2010) suggest detailed analysis of values in future studies. In this specific field of fashion products, online retailers have to address their efforts to perceived e-shopping value because during the exchange perceived e-shopping value determines consumers’ decisions about purchasing and future behaviour regarding shopping and after-purchase activities. According to Nilsson and Ballantyne (2014, p. 377) in real shopping atmosphere, “the servicescape […] become part of the meanings customers associate with any value proposition, thus influencing their service expectations”. In this way, the experience created by servicescape generates value through increased customer satisfaction (Voss et al., 2008). Therefore, the servicescape concept can be adapted to digital media and online shopping to support customer perceived e-shopping value. Investigating the connection between the dimensions of e-servicescape and perceived e-shopping value is important to predict the process of creating the perception of utility from the three dimensions of e-servicescape. As stated in
the research of Sreejesh and Ponnam (2016, p. 15), “researchers should examine the impact of each dimension of the e-servicescape separately”. Hence, the authors predict that the perception of the three dimensions of e-servicescape has a positive influence on perceived e-shopping value:

**H1.** Perception of online aesthetic appeal is positively related to consumers’ perceived e-shopping value.

**H2.** Perception of online layout and functionality is positively related to consumers’ perceived e-shopping value.

**H3.** Perception of online financial security is positively related to consumers’ perceived e-shopping value.

**Customer loyalty**

Customer loyalty can be defined as the preferential, attitudinal and behavioural purchase process towards one or more brands expressed over a period of time, where loyalty is the result of delighted customers delivering superior value from excellent services and quality products (Yang and Peterson, 2004). According to Yang and Peterson (2004), from the attitudinal perspective, customer loyalty is a specific desire to continue a relationship with the service provider, while the behavioural view defines customer loyalty as repeat patronage or the proportion of times a purchaser chooses the same product or service in a specific category compared to the total number of purchases made by the purchaser in that category. With the increase of internet transactions and purchases, a new construct has been frequently used: e-loyalty. It has been classified as a customer’s favourable attitude towards the e-retailer that results in repeat buying behaviour (Srinivasan et al., 2002). According to Gremler (1995), loyalty has two dimensions: attitudinal and behavioural, both of which are incorporated in this research. In addition to the virtual environment, one of the most important capabilities of the internet relative to other mass communication technologies is its bidirectionality, where customers can make their personal thoughts and opinions easily reachable to other internet users. That’s why the intention of positive WOM is defined as our recommendations from others. The EWOM is an extension of traditional WOM in a virtual environment (Matute et al., 2016). According to the research of Tan and Chang (2015), the advantages of EWOM over the traditional WOM can be seen in the different scenarios generated by the psychological distance dimensions and destination familiarity. It is frequently ranked as the most influential source of repurchase information (Kim and Hardin, 2010) and, in this study, is incorporated in the dimension of customer loyalty, because positive EWOM influences customer behaviour, such as expenditure level and loyalty (Bowman and Narayandas, 2001; Godes and Mayzlin, 2004). Also, according to Soderlund (2006), the WOM is part of attitudinal loyalty which explains the feelings towards online store. In previous studies, a positive correlation between perceived servicescape experiences and EWOM exchanges among customers has been found (Thevenot, 2007; Kim and Hardin, 2010). Carlson et al. (2015) investigate the customers’ perceived online channel value to online channel satisfaction and the online channel satisfaction to the online channel loyalty. Correspondingly, we predict that the perceived e-shopping value is understood as a necessary condition to establish loyalty. As such, these hypotheses are tested for the two dimensions of loyalty:

**H4.** Perceived e-shopping value is positively related to consumers’ attitudinal loyalty.

**H5.** Perceived e-shopping value is positively related to consumers’ behavioural loyalty.

With regard to the five above-mentioned hypotheses, the research model is shown in Figure 1.
Methodology

Questionnaire design and measurement

The questionnaire was designed using ten closed questions collecting demographic data and 41 items on a seven-point Likert scale. According to Harris and Goode (2010), who are considered “the pioneers to conceptualize and validate a comprehensive scale of e-servicescape in the online services literature” (Sreejesh and Ponnam, 2016), the construct of e-servicescape perception is composed of three measures, nine scales, and 52 items: “aesthetic appeal” is formed by visual appeal, originality of design, and entertainment value; “layout and functionality” includes usability, relevance of information, customization/personalisation and interactivity; “financial security” comprehends ease of payment, and perceived security. To enhance the usefulness of the measure, the authors themselves (Harris and Goode, 2010) suggest a shortened 24-item version of the scale which has been used in this research.

Perceived e-shopping value is analysed as an overall assessment of expected benefits during online shopping, including time and effort saving, economic value, etc. The scale was developed according to Sirdeshmukh et al. (2002) and Wu et al. (2014).

Shopper loyalty is conceptualised and measured by attitudinal loyalty and behavioural loyalty (Srinivasan et al., 2002; Ha and Im, 2012; Söderlund, 2006).

The research is based on a quantitative approach, using a questionnaire-based survey as a method of data collection, since it was considered prudent to focus on consumers’ interpretation in their natural setting, according to Wu et al. (2014). The sample was made up of active online shoppers who have at least one online purchase in the past six months in the sphere of fashion shopping. The researchers first asked the respondents to choose their favourite website where they purchase fashion products and with which they are familiar.

Sample

The research sample was selected from a population of respondents/participants of major forums and social network groups in Croatia that bring together fashion product buyers. Invitations were sent to forum and social network participants by e-mail or by private messages to the forums, containing a URL leading to the website hosting the survey. A screening question was posed to all respondents, asking whether they had purchased any fashion products online in the previous six months. A total of 221 fully completed

Figure 1. The research model

Source: By the authors
questionnaires were collected throughout April and May 2015, with a response rate of 14.75 per cent.

A non-random sample was chosen due to the unavailability of online data on respondents. With regard to the nature of the study, the authors considered it was essential that the sample consisted of online shoppers, thus providing valid data on online shopping behaviour so the sample is homogeneous. Considering that the study tests the theory of the influence of the perception of e-servicescape on perceived e-shopping value and customer loyalty, the use of maximally homogeneous samples is justified (Winer, 1999; Calder and Tybout, 1999).

To address a non-response bias in this study following the suggestions of Armstrong and Overton (1977), the sample was divided into two groups. The first group of 55 participants, accounting for 25 per cent of the sample, comprised respondents who answered the questionnaire first, and the second group included 25 per cent of the respondents who answered the questionnaire last. Subsequently, the value of the t-test difference of the arithmetic mean of both groups was calculated for all those that entered in the final model. t-test scores do not show significant differences (p < 0.05) in the responses of both groups. In addition, the demographic structure of respondents who responded to the questionnaire was compared with that of respondents who did not fully complete the questionnaire in terms of gender, age, income, education (Rogelberg and Stanton, 2007). Comparison of the demographic structure of these two groups showed no significant statistical differences. Finally, the structure of the sample given the various socio-demographic characteristics is similar to the total population. Therefore, it can be concluded that non-response bias does not significantly affect the results of the research.

The survey method was applied because the study involves selected aspects of online behaviour that cannot be observed directly but only indirectly using appropriately designed scales to measure attitudes. Various characteristics of the scales were tested using a variety of procedures (see the following section), thus ensuring that the conclusions reached have a satisfactory level of validity and reliability.

Of the total number of respondents, 6.3 per cent are between 18 and 21 years of age; 56.7 per cent between 22 and 30; 30.3 per cent between 31 and 40; while only 6.7 per cent of respondents are older than 41. With regard to gender, female and male respondents account for 95 and 5 per cent of respondents, respectively. According to educational background, 28.1 per cent of respondents hold secondary school qualifications; 18.1 per cent two-year post-secondary school qualifications; and 48.4 per cent higher education qualifications/master’s qualifications, while 0.5 per cent of respondents hold MScs and PhDs. Concerning total monthly household income, 24.8 per cent of respondents have an income less than EUR 750, while 23.1 per cent belong to the EUR 751–1,000 income bracket; 20.8 per cent to the EUR 1101-1500 income bracket; and 31.3 per cent to the over EUR1,500 monthly household income bracket.

Given that this is an observational study and the sample structure is uneven according to the gender proportion, four control variables are checked: gender, age, educational background and total monthly household income (see Table I).

<table>
<thead>
<tr>
<th>Control variables</th>
<th>Original sample – standardized coefficient (β)</th>
<th>t-values</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age → attitudinal loyalty</td>
<td>-0.086</td>
<td>1.582</td>
<td>0.114</td>
</tr>
<tr>
<td>Age → behavioural loyalty</td>
<td>-0.134</td>
<td>1.821</td>
<td>0.069</td>
</tr>
<tr>
<td>Income → attitudinal loyalty</td>
<td>0.011</td>
<td>0.226</td>
<td>0.821</td>
</tr>
<tr>
<td>Income → behavioural loyalty</td>
<td>0.002</td>
<td>0.034</td>
<td>0.973</td>
</tr>
<tr>
<td>Gender → attitudinal loyalty</td>
<td>0.012</td>
<td>0.256</td>
<td>0.788</td>
</tr>
<tr>
<td>Gender → behavioural loyalty</td>
<td>0.068</td>
<td>1.119</td>
<td>0.263</td>
</tr>
<tr>
<td>Education → attitudinal loyalty</td>
<td>0.058</td>
<td>1.084</td>
<td>0.278</td>
</tr>
<tr>
<td>Education → behavioural loyalty</td>
<td>0.119</td>
<td>1.750</td>
<td>0.081</td>
</tr>
</tbody>
</table>

Table I. Control variables
According to the results of Table I, no control variable performed has a statistically significant influence, so the evaluation of the measurement and structural model will continue without control variables.

**Results**

The data collected were analysed using the partial least squares (PLS) approach to analysing structural equations (SEM). PLS was applied instead of covariance-based SEM (CB-SEM) because it allows for the simultaneous analysis of the interrelationship of several latent variables, that is, the analysis of complex models with many manifest variables and theoretical constructs. In addition, PLS-SEM has no problems with identifying relations between individual theoretical constructs in small samples (< 250) (Reinartz et al., 2009) and has no special requirements tied to the distribution of manifest variables because it is a non-parameter method. Finally, PLS-SEM is better suited to situations where the objective and emphasis of analysis are focused more on prediction than on explanation (Hair et al., 2013).

SmartPLS 3 software (Ringle et al., 2015) was used for data analysis. Data were analysed in two steps. First, the measurement models of individual theoretical constructs were analysed to establish the psychometric suitability of the measurement scales. Then, a structural model was established to test the formulated hypothesis and links between the theoretical constructs. Prior to analysis, all manifest variables were checked for outliers. No outliers were identified, that is, there were no variable values exceeding the standard deviation by ±3. Because all data were collected using a single research instrument, a test was conducted for common method bias. Prior to the survey, the questionnaire was designed to clearly indicate which constructs related to which questions. In order to reduce the presence of the common method bias, the questionnaire was designed with each theoretical construct clearly separated from other constructs through a customised introductory text (Podsakoff et al., 2003).

In addition, Harman’s one factor test was also carried out (Podsakoff et al., 2003). Considering that one factor can explain, at the most, 36.7 per cent of variance of all variable indicators, it can be concluded that common method bias is not a significant problem in this study.

**Evaluation of the measurement model**

The comprehensive model of relationships between e-servicescape, perceived e-shopping value and customer loyalty have seven latent variables with reflective measurement models (Model A measurement). Reflective measurement models were used for certain constructs because they are in line with previous research (Harris and Goode, 2010; Wu et al., 2014). In addition, the reflective indicators used can be considered as a representative sample of all possible indicators of the conceptual domain of individual constructs (Hair et al., 2013), changes to indicators are caused by changes to individual theoretical constructs (Diamantopoulos and Winklhofer, 2001), and they can be considered mutually interchangeable (Jarvis et al., 2003) without jeopardising the content validity of a given construct. Finally, the assumption that certain indicators have no associated error term is not acceptable (Diamantopoulos, 2011). Hence, the following section evaluates the internal consistency reliability, and convergent and discriminant validity of the reflective measurement models of individual theoretical constructs. All analysed theoretical constructs, the e-servicescape in particular, were conceived as first order reflective models, considering that the shortened version of the scale (24 items), developed by Harris and Goode (2010), was used.

Outer loadings, composite reliability indicator (CR), Cronbach’s \( \alpha \) coefficient and average variance extracted indicator (AVE) were calculated to evaluate internal consistency reliability and convergent validity. The results of measurement model analysis are presented in Table II. Seven indicator variables, with outer loadings smaller...
<table>
<thead>
<tr>
<th>Item</th>
<th>Code</th>
<th>Outer loadings – original sample (O)</th>
<th>Outer loadings – sample mean (M)</th>
<th>SE (STERR)</th>
<th>t value</th>
<th>α</th>
<th>CR</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aesthetic appeal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The way it displays its products is attractive</td>
<td>AA1</td>
<td>0.824*</td>
<td>0.822</td>
<td>0.032</td>
<td>25.469</td>
<td>0.749</td>
<td>0.828</td>
<td>0.501</td>
</tr>
<tr>
<td>I like the way this website looks</td>
<td>AA2</td>
<td>0.849*</td>
<td>0.848</td>
<td>0.032</td>
<td>26.566</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is unadventurous</td>
<td>AA4</td>
<td>0.592*</td>
<td>0.590</td>
<td>0.087</td>
<td>6.265</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I think that this website is very entertaining</td>
<td>AA5</td>
<td>0.712*</td>
<td>0.706</td>
<td>0.056</td>
<td>12.621</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The enthusiasm of this website is catching, it picks me up</td>
<td>AA6</td>
<td>0.495*</td>
<td>0.487</td>
<td>0.084</td>
<td>5.887</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Layout and functionality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>It is not easily navigated</td>
<td>LF1</td>
<td>0.685*</td>
<td>0.684</td>
<td>0.062</td>
<td>11.121</td>
<td>0.837</td>
<td>0.876</td>
<td>0.506</td>
</tr>
<tr>
<td>Navigation through this website is intuitively logical</td>
<td>LF3</td>
<td>0.601*</td>
<td>0.799</td>
<td>0.034</td>
<td>23.739</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This website is difficult to use</td>
<td>LF4</td>
<td>0.619*</td>
<td>0.616</td>
<td>0.082</td>
<td>7.358</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This website is user-friendly</td>
<td>LF5</td>
<td>0.745*</td>
<td>0.746</td>
<td>0.044</td>
<td>17.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical details about products can be easily accessed</td>
<td>LF7</td>
<td>0.785*</td>
<td>0.786</td>
<td>0.035</td>
<td>22.732</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This website is tailored towards me</td>
<td>LF8</td>
<td>0.732*</td>
<td>0.730</td>
<td>0.047</td>
<td>15.446</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel that this is not a very engaging web site</td>
<td>LF13</td>
<td>0.581*</td>
<td>0.580</td>
<td>0.061</td>
<td>9.490</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Financial security</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payment procedures seem to take a long time</td>
<td>FS1</td>
<td>0.789*</td>
<td>0.788</td>
<td>0.041</td>
<td>19.029</td>
<td>0.815</td>
<td>0.866</td>
<td>0.566</td>
</tr>
<tr>
<td>Paying for goods is straightforward</td>
<td>FS2</td>
<td>0.841*</td>
<td>0.842</td>
<td>0.024</td>
<td>34.786</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paying for goods involves entering a lot of details</td>
<td>FS3</td>
<td>0.625*</td>
<td>0.622</td>
<td>0.039</td>
<td>10.387</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When buying from this website I am not reassured by the security procedures</td>
<td>FS4</td>
<td>0.906*</td>
<td>0.691</td>
<td>0.054</td>
<td>12.923</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall, this website seems security conscious</td>
<td>FS5</td>
<td>0.791*</td>
<td>0.791</td>
<td>0.042</td>
<td>18.679</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>E-shopping value</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This online store offers a good economic value</td>
<td>EVAL1</td>
<td>0.702*</td>
<td>0.700</td>
<td>0.053</td>
<td>13.192</td>
<td>0.856</td>
<td>0.896</td>
<td>0.635</td>
</tr>
<tr>
<td>The product/service I purchased from this online store is a good buy</td>
<td>EVAL2</td>
<td>0.782*</td>
<td>0.780</td>
<td>0.044</td>
<td>17.927</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When I make a purchase from this online store, I save time</td>
<td>EVAL3</td>
<td>0.752*</td>
<td>0.751</td>
<td>0.046</td>
<td>16.296</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>It is easy to shop for product/service in this online store</td>
<td>EVAL4</td>
<td>0.873*</td>
<td>0.875</td>
<td>0.017</td>
<td>52.646</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The effort that I must make to purchase from this online store is low</td>
<td>EVAL5</td>
<td>0.861*</td>
<td>0.862</td>
<td>0.020</td>
<td>42.548</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>Code</td>
<td>Outer loadings – original sample (O)</td>
<td>Outer loadings – sample mean (M)</td>
<td>SE (STERR)</td>
<td>t-value</td>
<td>Cronbach’s α</td>
<td>CR</td>
<td>AVE</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>--------</td>
<td>--------------------------------------</td>
<td>---------------------------------</td>
<td>------------</td>
<td>---------</td>
<td>--------------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Attitudinal loyalty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I like using this website</td>
<td>ALOY1</td>
<td>0.870*</td>
<td>0.870</td>
<td>0.017</td>
<td>51.217</td>
<td>0.923</td>
<td>0.940</td>
<td>0.724</td>
</tr>
<tr>
<td>To me this website is the best retail website to do business with</td>
<td>ALOY2</td>
<td>0.764*</td>
<td>0.762</td>
<td>0.034</td>
<td>22.299</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I believe that this is my favourite retail website</td>
<td>ALOY3</td>
<td>0.757*</td>
<td>0.755</td>
<td>0.041</td>
<td>18.404</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would recommend this website to other people</td>
<td>ALOY4</td>
<td>0.921*</td>
<td>0.929</td>
<td>0.013</td>
<td>71.119</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would recommend this website to my friend</td>
<td>ALOY5</td>
<td>0.925*</td>
<td>0.925</td>
<td>0.011</td>
<td>83.587</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I will encourage people to do business with this site</td>
<td>ALOY6</td>
<td>0.852*</td>
<td>0.851</td>
<td>0.028</td>
<td>30.033</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioural loyalty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I seldom consider switching to another website</td>
<td>BLOY1</td>
<td>0.707*</td>
<td>0.702</td>
<td>0.053</td>
<td>13.464</td>
<td>0.854</td>
<td>0.891</td>
<td>0.621</td>
</tr>
<tr>
<td>As long as the present service continues, I doubt that I would</td>
<td>BLOY2</td>
<td>0.796*</td>
<td>0.792</td>
<td>0.041</td>
<td>19.219</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>switch websites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I try to use the website whenever I need to make a purchase</td>
<td>BLOY3</td>
<td>0.742*</td>
<td>0.738</td>
<td>0.048</td>
<td>15.350</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The probability that I will use this online store again is high</td>
<td>BLOY4</td>
<td>0.804*</td>
<td>0.805</td>
<td>0.037</td>
<td>21.486</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When I need to make a purchase, this website is my first choice</td>
<td>BLOY5</td>
<td>0.879*</td>
<td>0.877</td>
<td>0.023</td>
<td>38.602</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *p < 0.05
than 0.4, were eliminated from further analysis. One of the rejected indicators referred to the construct of aesthetic appeal and six to the construct of layout and functionality. Because the rejection of six manifest variables from construct LF implies the possibility of an incorrectly specified measurement model, confirmatory tetrad analysis (CTA) (Gudergan et al., 2008) was carried out according to the approach of Bollen and Ting (2000). CTA is based on the concept of tetrads, representing the difference between the products of two pairs of co-variances among various indicator variables. In a reflective measurement model, it is assumed that each tetrad vanishes, that is, equals zero (Hair et al., 2016, p. 286). The CTA conducted established that not one tetrad is statistically different from zero, making the tetrads appropriate for use in a reflective measurement model for the construct LF. In addition, all servicescape constructs were designed as formative models and it was established that for most of the indicator variables the outer weights were not statistically significant, suggesting the use of a reflective measurement model. For the most part, the rejected indicator variables of construct LF refer to the sub-dimension customization (LF9, LF10 and LF12) and, having similar semantic content, can be replaced with the variable LF8 – “This website is tailored to me”. Finally, the rejection of the stated variables can be ascribed to the shortened version of the servicescape scale, so the results obtained indicate the need of using the full version of the servicescape scale for future research in the mentioned areas.

The results of measurement model analysis suggest that all seven reflective measurement constructs possess a satisfactory level of internal consistency reliability and convergent validity. Outer loadings of all indicator variables of individual constructs are statistically significant at the 5 per cent level. The statistical significance of outer loadings was established by the bootstrap procedure based on 5,000 sub-samples (as recommended by Hair et al., 2016, p. 149). Furthermore, the outer loadings of most of the variables exceed the recommended level of 0.7, and the indicator variables with outer loadings ranging from 0.4 to 0.7 were retained because CR values and Cronbach’s α coefficient were above the recommended level of 0.7, while the AVE of indicators was above 0.5, and the elimination of these variables would not bring about a considerable increase in the values of the mentioned coefficients (Nunnally and Bernstein, 1994; Hair et al., 2013).

Discriminant validity was assessed by using the Fornell–Larcker criterion, comparing cross-loadings and calculating the heterotrait–monotrait ratio (HTMT) of the correlations. Table III presents the assessment of the Fornell–Larcker criterion. The table provides the square roots of the indicators’ AVE on the diagonal, and the correlations of the indicators to each other below the diagonal. The total square roots of the indicators’ AVE for all constructs are greater than the correlation of these constructs with other constructs in the model, thus meeting the Fornell–Larcker criterion (Fornell and Larcker, 1981).

The discriminant validity of individual constructs is confirmed when the indicator’s outer loadings with its associated construct are higher than all of its cross-loadings with the remaining constructs, which is proved in this model. The HTMT ratio of correlations is a

<table>
<thead>
<tr>
<th></th>
<th>Aesthetic appeal</th>
<th>Attitudinal loyalty</th>
<th>Behavioural loyalty</th>
<th>Financial security</th>
<th>Layout and functionality</th>
<th>EVAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aesthetic appeal</td>
<td>0.708</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attitudinal loyalty</td>
<td>0.473</td>
<td>0.851</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioural loyalty</td>
<td>0.356</td>
<td>0.680</td>
<td>0.788</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial security</td>
<td>0.385</td>
<td>0.555</td>
<td>0.359</td>
<td>0.753</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Layout and functionality</td>
<td>0.605</td>
<td>0.627</td>
<td>0.467</td>
<td>0.582</td>
<td>0.711</td>
<td></td>
</tr>
<tr>
<td>EVAL</td>
<td>0.438</td>
<td>0.702</td>
<td>0.435</td>
<td>0.596</td>
<td>0.633</td>
<td>0.797</td>
</tr>
</tbody>
</table>

**Table III.** Fornell–Larcker criterion
relatively new indicator for assessing the discriminant validity of constructs in PLS models and research has shown it to be superior to the Fornell–Larcker criterion and to cross-loadings testing (Henseler et al., 2015, p. 121). The HTMT ratio for each pair of analysed constructs is lower than the recommended value of 0.9, and the upper limit of the bias-corrected confidence interval for the HTMT ratio is less than 1, providing sufficient evidence of the existence of the model’s discriminant validity.

Structural model analysis

After checking for internal consistency reliability and convergent and discriminatory validity, the structural model was analysed and the formulated hypotheses tested. The statistical significance of the structural model’s parameters was established using the bootstrapping technique, running 5,000 sub-samples (the option no sign changes was used). Table IV shows the results of structural model analysis.

According to the results of structural model analysis, H2–H4 are accepted, while H1 is rejected. The constructs of aesthetic appeal – AA (H1: $\beta = 0.101; [-0.021–0.295]$, $p < 0.1$) has no statistically significant impact on perceived e-shopping value. The other two servicescape constructs, layout and functionality – L&F (H2: $\beta = 0.339; [0.217–0.464]$, $p < 0.05$) and financial security – FS (H3: $\beta = 0.375; [0.243–0.512]$, $p < 0.05$) have a statistically significant positive influence on perceived e-shopping value and, together, they explain 48% of its variances ($R^2 = 0.48$), which can be considered moderate. In explaining the perceived e-shopping value construct, the $f^2$ effect sizes of the constructs AA, FS and L&F are 0.012, 0.147 and 0.134, respectively, which, according to Cohen (1988), can be considered a small effect size for the construct AA and a medium effect size for FS and L&F. Also discovered was a statistically significant positive effect of perceived e-shopping value on the constructs of attitude loyalty (H4: $\beta = 0.702; [0.625–0.777]$, $p < 0.05$), behavioural loyalty (H5: $\beta = 0.435; [0.332–0.546]$, $p < 0.05$). The construct of perceived e-shopping value explains 49% per cent of the variances of the construct AttLoy ($R^2 = 0.49$), and 19 per cent of the variances of the construct BehLoy ($R^2 = 0.19$). The $f^2$ effect size of the construct of perceived e-shopping value in explaining the constructs AttLoy, BehLoy amounted to 0.97 and 0.23, respectively. In other words, perceived e-shopping value has a large effect on AttLoy but a small effect on BehLoy.

The predictive relevance of the structural model was assessed by the blindfolding procedure using the cross-validate redundancy approach, and Stone–Geisser’s $Q^2$ values were calculated (Geisser, 1974; Stone, 1974). The $Q^2$ values of all endogenous constructs (perceived e-shopping value, AttLoy, and BehLoy) were higher than 0, indicating the model has a sufficient level of predictive relevance. Certain constructs of the e-servicescape had small predictive relevance for the construct of perceived e-shopping value; the values of $q^2$, the indicator of relative predictive relevance, ranged from 0.01 to 0.06. The relative predictive relevance of the perceived e-shopping value for AttLoy (0.06) and for the BehLoy (−0.03) construct was small.

The standardized root mean square residual value was calculated to assess the quality of the structural model. It allows assessing the average magnitude of the discrepancies between observed and expected correlations as an absolute measure of (model) fit criterion, which was 0.085 for the composite factor model, less than the recommended value of 0.1 (Hu and Bentler, 1998; Henseler and Sarstedt, 2013), indicating a sufficient level of structural model adjustment to empirical data. Results of the study are presented in greater detail in the discussion section and in recommendations for future research.

Closing remarks

Discussion and academic implications

The development of online shopping is opening up new opportunities for customers because online competition is much more intense and fierce, geographical distances are eliminated,
<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Original sample – standardized coefficient (β)</th>
<th>t-values</th>
<th>95% confidence intervals**</th>
<th>$R^2$</th>
<th>$f^2$ effect size</th>
<th>$Q^2$</th>
<th>$q^2$</th>
<th>Accepted hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H1$ AA → value</td>
<td>0.101</td>
<td>1.564</td>
<td>[−0.021 – 0.230]</td>
<td>0.48</td>
<td>0.012</td>
<td>0.276</td>
<td>0.01</td>
<td>Rejected</td>
</tr>
<tr>
<td>$H2$ FS → value</td>
<td>0.339*</td>
<td>5.238</td>
<td>[0.217 – 0.464]</td>
<td>0.147</td>
<td>0.05</td>
<td>Accepted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H3$ L&amp;F → value</td>
<td>0.375*</td>
<td>5.381</td>
<td>[0.243 – 0.512]</td>
<td>0.134</td>
<td>0.06</td>
<td>Accepted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H4$ Value → AttLoy</td>
<td>0.702*</td>
<td>18.161</td>
<td>[0.625 – 0.777]</td>
<td>0.49</td>
<td>0.972</td>
<td>0.300</td>
<td>0.06</td>
<td>Accepted</td>
</tr>
<tr>
<td>$H5$ Value → BehLoy</td>
<td>0.435*</td>
<td>7.976</td>
<td>[0.332 – 0.546]</td>
<td>0.19</td>
<td>0.234</td>
<td>0.100</td>
<td>0.03</td>
<td>Accepted</td>
</tr>
</tbody>
</table>

Notes: *p < 0.05; **bias-corrected
and it is easier for customers to compare product characteristics. These features, together with the growing computer literacy of customers, are making e-commerce all the more important for almost every category of product, especially fashion products.

This study analyses the transition of the buying process from a physical shop to an online shop starting from the e-servicescape perception. Generating new associations and relationship in the virtual environment, this paper seeks to understand the construct of the perceived e-shopping value affecting loyalty. It is important for the online business to understand which dimension of the online servicescape is the best predictor for perceived e-shopping value and loyalty. Therefore, this study applies the shortened version of the multidimensional scale for e-servicescape by Harris and Goode (2010) for investigating the impact on the perceived e-shopping value and, consequently, the customer loyalty in the S–O–R framework. Although the application of the S–O–R model is similar to that seen in previous studies (Wu et al., 2014, 2016), some of the insights of this work can be important for the exploration of the perception of e-servicescape, A concept which is recently highlighted as one of the leading emergent areas (Mari and Poggesi, 2013; Eroglu et al., 2003). Aligned with the finding of Harris and Goode (2010) and Kühn et al. (2015) for predicting customers’ website trust, the most influential e-servicescape dimensions for perceived e-shopping value is Financial security. Contrary to the study of Wu et al. (2016) which had pointed out that the aesthetic appeal had the greatest impact on the purchase intention, this research demonstrates that aesthetic design of the website has no statistically significant impact on perceived shopping value. The explanation of this unexpected result can be found in Lai et al. (2014) who pointed out that, due to the online experience gained by shoppers through multiple exposures to shopping websites, visual messages may not contain information pertaining to buying. This finding in the sphere of fashion products is specifically important because stems that the online shoppers need buying information from verbal messages, since they are more aware because of their buying frequency and experience. In accordance with the results of Hernandez et al. (2011), this research has proven that the socioeconomic variables of gender, age, educational background and total monthly household income have no statistical significance on perceived e-shopping value and loyalty for those e-shoppers who make frequent online purchases.

This paper contributes to the literature also by providing a better understanding of the process of how the perceived e-shopping value affects the multidimensional construct of customers’ loyalty. The perceived e-shopping value in online shopping research is proven to be important as supported by Wu et al. (2014) and Hsin Chang and Wang (2011). Accordingly with the authors, who argue the perception of value affected by the servicescape in offline shopping (Voss et al., 2008; Chen and Dubinsky, 2003, Payne et al., 2001), this study proves that also in the online reality the e-servicescape influence consumers’ perceptions and intentions. The results of the current study support the multidimensionality of the shoppers’ e-loyalty construct (Srinivasan et al., 2002), enriched with the intention of positive WOM integrated in behavioural loyalty dimension (Ha and Im, 2012).

Since previous studies have ignored the impact of e-servicescape on customers’ loyalty, this study provides additional evidence in the S–O–R model that dimensions of e-servicescape as stimulus affect the perceived e-shopping value, which influences the customers’ loyalty. In this way, different dimensions of the e-servicescape need to be managed in order to build a positive perceived e-shopping value of the website which should result in customers’ loyalty.

Recommendations to managers
Research results indicate the existence of the positive effect of certain dimensions of e-servicescape on perceived e-shopping value as a core concept, the enhancement of which should be the focus of marketing efforts of online fashion product shops. According to
research results, this will help to create positive attitudes towards an online shop, ensure repeat purchase intention and spread positive information about the online shop to others, thus improving the shop’s image and reducing the risk perceived by online customers. Managers responsible for the design of fashion product websites should therefore consider these dimensions of the e-servicescape and incorporate them in the design of their online strategies to build perceived e-shopping value and loyalty. Accordingly, online shops need to invest in all dimensions of e-servicescape, in particular in layout and functionality and financial security, dimensions that had a stronger effect on perceived e-shopping value. Analysing online financial security, sellers of fashion products should focus on perceived security and payment options to increase trust and website value, also providing privacy guarantees against misuse of financial information.

To enhance perceived e-shopping value in the fashion product context, online shops should make efforts to increase the usability of their websites, manage the information required by online customers, customise various parts of their web pages to meet customer needs, and enable interaction with customers. Furthermore, special attention should be placed on ensuring transaction security, investing in customer personal data protection and informing customers regarding the use of their personal data.

Although the study showed that the aesthetic appeal dimension of e-servicescape has no influence on perceived e-shopping value, online shops should not neglect this aspect of the offering. This in particular refers to the visual aspect of web pages that creates the first impression and which can affect risk perception. The weak effect of the aesthetic appeal dimension can be explained by the perceived “similarity” of various online shops with regard to web page appearance and originality. In addition, some 34 per cent of respondents most often purchased from online shops that are organised as auctions, where aesthetic appearance and having fun while shopping are not dominating features.

The authors seek to enhance interest in research into, and care about every aspect of B2C communication, especially on the web, with customers as the focus of every marketing-minded organisation.

Research limitations and future research
The research conducted has several limitations that could serve as fertile areas for future study. First, to lessen the strain on respondents, the shortened version of the original scale of Harris and Goode (2010) was used in measuring e-servicescape dimensions. As a result, it was not entirely possible to compare the results of this research with those from studies that used the whole scale, thus limiting the possibility of generalising results. Second, the research neglected the dynamics between individual constructs. An experimental approach (e.g. conjoint analysis) could be taken to investigate the relationships between e-servicescape and other important marketing constructs in an online context, such as perceived risk, online service satisfaction, e-service quality, price perception and product quality. In this way, the causal interrelationship between various constructs could be established to a large extent, which is not possible with cross-sectional studies. Third, because the sample structure predominantly included women (about 95 per cent), this reduces the ability to generalise results even though women are mostly the ones buying fashion products. Four, generalisation is also reduced considering that the majority of respondents (over 64 per cent) rated four online shops and that their online purchases largely involved clothes (76.3 per cent), footwear (50.5 per cent) and jewellery (35.2 per cent). Future research should expand both the type of fashion products to include, for example, electronic devices, perfume, etc., and the number of analysed online shops. Five, the weaker effects of certain e-servicescape components on perceived e-shopping value, and perceived e-shopping value on the construct of behavioural intentions, suggests the need to include other theoretical constructs mentioned earlier, and to analyse the effects of moderating variables on the relationship between constructs.
Moderating variables between e-servicescape and other constructs in future studies could include the online shopping experience of customers, the attributes of online shops (e.g. the level of shop personalisation, online path finding assistance, environment atmospherics, and manner of product presentation), etc. Although the language was not taken into consideration in this study, its importance as a moderating effect in online servicescapes has been proven (Alcántara-Pilar et al., 2017) and future research could investigate the impact of language on e-servicescape perception. The construct of perceived e-shopping value could be analysed from additional content on the internet, such as social networks or forums. On the other hand, it would be interesting to make a comparison of servicescape perception between online and real shopping to investigate the differences between elements of servicescape in a real atmosphere and a web-based environment. Also, the weaker links between certain constructs suggest the possible existence of unobserved heterogeneity (Hair et al., 2013), and future research could address this issue, which the present paper did not due to limited scope of research.

References


**Further reading**


**Corresponding author**
Ana Cuic Tankovic can be contacted at: anact@fthm.hr

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 and co-creative service innovation: an empirical study

Bijoylaxmi Sarmah and Shampy Kamboj

Department of Management Studies, Indian Institute of Technology Roorkee, Roorkee, India, and
Jay Kandampully
Department of Human Sciences, The Ohio State University College of Education and Human Ecology, Columbus, Ohio, USA

Abstract

Purpose – Online information research on hotels is gradually emerging as a key area of research with the increasing use of social media as a platform for co-creative service innovation (CCSI). The purpose of this paper is to examine the relationships between the key drivers of co-creation intention in the social media context. Understanding relationships between key drivers of customers’ co-creation intention will prove valuable in advancing current knowledge about service innovation using social media. The key drivers examined in this study are – customer innovativeness, attitude toward CCSI on social media, subjective norms and perceived behavioral control. This knowledge will be of considerable value for its practical application in the hotel industry.

Design/methodology/approach – Data were collected from 346 hotel guests using survey method. Structural equation modeling with a bootstrapping estimation was used to analyze the data.

Findings – The results show that customer innovativeness, attitude toward CCSI on social media, subjective norm and perceived behavioral control positively influence both co-creation and adoption intention. Further, it was also found that co-creation intention mediates the relationship between its two driving factors, namely, customer innovativeness, attitude toward CCSI on social media and adoption intention.

Research limitations/implications – The findings provide theoretical implications for hospitality discipline. The findings also provide various strategies hospitality firms can use to co-create service innovation through the effective use of social media.

Originality/value – The relationships examined in the present study have not been tested previously; this is the first attempt of the kind. Thus, the associations established in this study form an important contribution to the existing body of knowledge in co-creation, service innovation and social media literature.

Keywords Service innovation, Social media, SEM, Hotels, Co-creative

Paper type Research paper

1. Introduction

Modern sustainable business practices and strategies are influenced by changes in information technologies (ITs) (Hua, 2016; Law et al., 2015), which are also considered crucial for gaining competitive advantage in service sector industries including the hospitality industry (Bilgihan et al., 2011; Pereira-Moliner et al., 2016). In recent years, the impact of online information research on hotels has been acknowledged (Buhalics, 2003) with the growing use of social media as a promotional, selling and service innovation platform for hotels (Chan and Law, 2006; Sarmah and Rahman, 2017). Although the extant literature is rich about the role of hotel websites as a part of the digital strategy (Baloglu and Pekcan, 2006), effective utilization of social media to develop new services (O’Connor, 2010) through customer–firm co-creation still lacks research attention.

Social media have emerged as an important platform for interactions and exchange of multidimensional information, including comments, reviews, invitations, images, photos and videos (Urutnik, 2016). These online communities are flexible, fluid and represent a
broad range of cultural concerns and social affiliations (Brown et al., 2007). In recent years, the use of social media has encouraged firms to engage with prospective consumers (Lee and Suh, 2016; Filieri and McLeay, 2013), as well as increased the firms’ ability to influence buying behavior (Vermeulen and Seegers, 2009). As the proliferation of social media applications continues, hotel firms have an opportunity to determine customers’ wants and needs and to invite customers to co-create new services with them. Technology, including IT, online platforms, and social media, has revolutionized the global market and, perhaps even more importantly, provided new opportunities to engage customers to disseminate valuable ideas, feedback and other useful information (Zhang et al., 2015). However, the ubiquity of social media is changing user interaction in the context of service innovation (Uratnik, 2016). Social media is enabling customers to co-create new services in collaboration with the service firm by engaging through social media such as Facebook, Instagram and Twitter (Kaur, 2016). In this study, we extend the service-ecosystem approach (Vargo and Akaka, 2012) to study co-creative hotel service innovation, and further to explain how it can provide a theoretical lens for understanding technology-mediated co-creative service innovation (CCSI) activities using social media in the context of the hospitality industry.

The study of co-creation is an integral part of service research and focuses mainly on exchange networks, highlighting interactions among customers, service firms and technology (Maglio and Spohrer, 2008). Highlighting the growing importance in service, Vargo et al. (2008) proposed the concept of service-dominant (S-D) logic and defined it as “the application of resources for the benefit of another that centers on the concept of value co-creation.” Since its introduction in 2004, S-D logic has been revised, elaborated, and extended to include the “service-ecosystems perspective” (Vargo and Lusch, 2011). A “Service-ecosystems” framework is applied allowing for “understanding service systems or the interaction and co-creation among multiple service systems” (Vargo et al., 2010). The ecosystems view emphasizes the importance of networking among various stakeholders, institutions, technology and social norms for CCSI (Aal et al., 2016).

The emerging literature review in the CCSI area reveals that most studies have emphasized the nature of customer value co-creation (Roberts et al., 2014), adoption intention (Arts et al., 2011; Morosan, 2015), satisfaction (Hoyer et al., 2010), purchase intention (Franke et al., 2008), repurchase (Dong et al., 2008), willingness to pay (Franke and Piller, 2004), etc. A few studies have also highlighted the contribution of social media and mobile technologies in the co-creation process (Cheung and To, 2016; Morosan et al., 2016). According to Kai-Wai Chu and Kennedy (2011), collaborative activities lead to the development of a shared understanding of concepts. A few factors such as the quality and depth of the interactions, together with the communication tools used, often affect the effectiveness of co-creation.

In literature, co-creation intention has been discussed in several contexts such as hotels (Morosan, 2015), new product development (NPD) (Füller et al., 2009) and online course registration and internet setup scenarios (Dong et al., 2008). Additionally, the role of antecedents of co-creation intention includes: customer involvement (Morosan, 2015), customer participation in service recovery, customers’ ability for future co-creation, role clarity for future co-creation, perceived value for future co-creation, satisfaction with service recovery (Dong et al., 2008), etc.

Previous studies in the hotel industry have applied the theory of diffusion of innovation (Rogers, 2003) and theory of planned behavior (TPB, Fishbein and Ajzen, 1975; Ajzen, 1991) to examine customer behavioral intention. However, very few studies have used these two theories to address hotel customers’ future co-creation intention while participating in the hotel’s new service development activities in social media context (Chang and Taylor, 2016). More specifically, an empirical explanation regarding the formation of hotel customers’
future co-creation intentions to develop new services has been atypical. It is therefore important to understand factors that lead to customers’ future intentions to co-create and to gain insight into their decision-making processes. Overall, this study aims to test the applicability of theory of diffusion of innovation and TPB to explain hotel customers’ intention to co-create a new hotel service in the future.

The Indian hotel industry is a major contributor in the tourism and hospitality industry and is expected to generate 2.3m jobs by 2025 (India Brand Equity Foundation, IBEF, 2016). Revenues from foreign and domestic tourism reached US$21.08bn in 2015, respectively. The Indian tourism and hospitality industry is the third largest sub-segment of the Indian service sector which accounted for 12.5 percent of India’s GDP in the 2014–2015 timeframe (India Brand Equity Foundation, IBEF, 2016). This growth further signifies increased importance for customized and specially designed service offerings to satisfy the tourists and hotel guests. Hotels are using advanced technologies to enhance user-friendly interfaces on social media to enable hotel guests to enjoy their role as co-creator of new services (Frochot and Batat, 2013). Although hotel guests may show separate set of behaviors during their stay in a hotel, it will be most useful to know their adoption intention behavior toward newly developed services (Roberts and Shea, 2017). Hotels are encouraging sharing photos and videos on social media with reference to their experience at the hotel (e.g. hotels may recommend the most Instagram-able suite or cuisine to patrons) (Kim et al., 2016). Many hotels provide topics of conversation among the customers and reward guests for their active participation (see Figures 1 and 2). The leading travel company, Make My Trip (MMT), offers e-cash to be credited in customer MMT account wallet if they write reviews about their experience at a hotel reserved over their website. The feedback may include any suggestions for consideration and approval by MMT for service improvement. Once approved, customer’s suggestions are published on the company’s website to further motivate other customers to participate.

In this way, hotel firms can use social media as platforms to connect current and prospective customers while generating awareness about services and maintaining a customer relationship (Facebook.com, 2011). The Marriott Hotels launched a global marketing campaign named “Travel Brilliantly” and concurrently launched a new website – www.travelbrilliantly.com – as a co-creation platform to solicit user-generated ideas to facilitate service innovations (http://marriott-hotels.marriot.com/). This website depicted Marriott’s past and future innovations and welcomed innovative ideas from its customers. Visitors to the website were encouraged to share their ideas to improve the tourists’ experience in areas such as: layout, cuisine, wellness and technology. Financial and non-financial incentives were offered to the providers of the best ideas.

In the NPD, “My Starbucks Idea” was another example of using customers’ ideas to design new products and services. The ideas were shared, discussed and voted on via social media (https://ideas.starbucks.com). In India, Tanishq, a jewelry brand under the Tata group, used crowdsourcing practices on social media as a tool for customer co-creation in ideation stages of NPD through an engagement platform called “My Expression” (Sarmah and Rahman, 2017).

Prior research in CCSI has attempted to test drivers influencing customers’ future co-creation intention such as: innate innovativeness (Lowe and Alpert, 2015; Morosan, 2015), attitude toward co-creation and subjective norm (Cheung and To, 2016). Conversely, there is no reported empirical study that has tested a complete model of the possible relationships between innate innovativeness, attitude toward co-creation, subjective norm or perceived behavioral control, and how they impact the intention to co-create on social media specifically in the hospitality industry. To address these critical lacunas, and considering the fact that customer co-creation is crucial for service innovation, this study proposes and validates a conceptual model that explicates customers’ adoption intention toward co-creativity to develop new services in hotels.
This study’s objectives are to:

1. investigate how the driving factors of co-creation intention (customer innovativeness, attitude toward CCSI in social media, subjective norm and perceived behavioral control) relate to co-creation intention on social media in hotels;
2. examine the relationship between co-creation intention and adoption intention behavior on social media in hotels; and
3. investigate the mediating roles of co-creation intention and concern between customer innovativeness, attitude toward CCSI in social media and adoption intention in hotels.

To achieve these objectives, this study has applied the theory of diffusion of innovation (Rogers, 2003) and the TPB (Fishbein and Ajzen, 1975; Ajzen, 1991) to identify the underlying factors and conditions that drive hotel guests to co-create new services using social media. The extensive use of social media by customers provides an opportunity to undertake this study pertaining to the technology-mediated service ecosystem contributing to co-creative hotel service innovation.
This study is organized as follows: the ensuing section provides the review and theoretical background of this research and proposes the research hypotheses for this study. Next, based on the theoretical underpinning, a conceptual model is presented as a means to quantify hotel customers’ future co-creation intention. Afterward, the methodology used and the data collection procedures are explained along with the data analysis. Subsequently, the results of the study are presented. Finally, conclusions, implications, limitations and future research avenues are discussed in Table 1.

2. Background and hypotheses

2.1 CCSI in social media

Vargo et al. (2015) define innovation as “the combinatorial evolution of new, useful knowledge in the context of service ecosystems” and viewed service ecosystem as “a relatively self-contained, self-adjusting system of resource-integrating actors that are connected by shared institutional logics and mutual value creation through service exchange” (Vargo and Lusch, 2016).
It is viewed that an innovation involves any institutionalized practice that enables actors to integrate resources and co-create value for both themselves and others in novel and useful ways (Edvardsson and Tronvoll, 2013). Recently, Vargo and Lusch (2016) explored the role of institutions and institutional arrangements in service ecosystems to uncover ways to respond to business uncertainties that seek resource integration as a method to create value for the involved actors. Lusch and Nambisan (2015) proposed that “innovation as a collaborative process occurring in an actor-to-actor network, service as the application of specialized competences for the benefit of another actor or the self and as the basis of all exchange, the generatively unleashed by increasing resource liquefaction and resource density, and resource integration as the fundamental way to innovate.” Customers are considered as innovators in CCSI and study their transformation from being passive recipients of products or services to active co-creators of new services (Magnusson, 2009). This transformational journey of customers as co-creators is regarded as a new and emerging paradigm in service innovation literature which warrants further consideration.

Previous studies have addressed service innovation in the hospitality and tourism industry. Jones (1996) outlined a 15-step approach to the innovation development process in the hospitality industry. Zhang et al. (2015) examined motivations for customer engagement in online co-innovation communities. There have also been many case studies on individual hotels, for example, Phan (2007) identified key factors for successful new service development. Literature also focused on technology application and adoption in various hospitality firms that principally focused on website adoption (Scaglione et al., 2009), e-innovation in small UK hospitality firms (Martin, 2004) and the effect of technological innovations on the relationships with customers in hospitality firms (Khan and Khan, 2009).

In this study, the role of the customer’s innate characteristics, “customer innovativeness,” is adopted from the theory of diffusion of innovation. Customer innovativeness is examined as the antecedent of the customers’ intention to co-create and their adoption intention of those co-created new hotel services on social media. Other antecedents of adoption intention are adopted from the TPB. The other antecedents include: attitude toward CCSI, subjective norm and perceived behavioral control. The dependent construct adoption intention is chosen to explore how well it reflects the co-creation that is fundamental to service innovation.

Online information research has been fueled by the advent of Web 2.0 (Musser and O’Reilly, 2007) and social media (Haiyan, 2010). Consequently, the hospitality industry has also observed major transformation in its business and dealings with the guests due to

<table>
<thead>
<tr>
<th>Authors</th>
<th>Definition</th>
<th>Theoretical underpinnings</th>
<th>Dimensions of co-creation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azevedo (2009)</td>
<td>A relationship between producer and consumer</td>
<td>Co-creation theory</td>
<td>Active participation/involve; integration</td>
</tr>
<tr>
<td>Kreziak and Frochot (2011)</td>
<td>The active and creative role of tourists in the process of value co-creation; tourists are co-producers</td>
<td>Consumer agency; S-D logic</td>
<td>Participation; social interaction (“socialization”) and with the organization; the experience of flow</td>
</tr>
<tr>
<td>Mathisen (2013)</td>
<td>The creation of value and views tourists as active, with a desire to use their own knowledge and skills in order to interact with other tourists, objects and environments</td>
<td>S-D logic</td>
<td>Active participation and engagement (emotional, physical and mental), social interaction and with the environment, reflexivity</td>
</tr>
<tr>
<td>Rihova et al. (2013)</td>
<td>The customer is the sole creator of value, while the firm joins in as a supporter/facilitator of customers’ value creation</td>
<td>S-D logic; C-D logic</td>
<td>The social interaction between tourist during activities, practices and experiences</td>
</tr>
</tbody>
</table>
social media. Social media has been considered as an important platform for maintaining relationships with the users. The social media brand page visitors post online reviews may affect room sales (Ye et al., 2009; Vermeulen and Seegers, 2009), service design and new service development (Gretzel and Yoo, 2008) and positive comments on social media may influence and improve customer attitudes toward hotels (Kamboj et al., 2018). It is also viewed that hotels with integration with the interactive media (i.e. social media) cause rise in online development and distribution of new services that benefit both hotel companies and customers (Bilgihan and Bujisic, 2015).

2.2 Innovation diffusion theory (IDT)
The IDT is concerned with “understanding the process by which the use of innovation spreads throughout a social system” (Rogers, 1995). Diffusion is the process by which an innovation is communicated through certain channels over time among the members of a social system. The foundation of this study is based on Everett Rogers’s (1995) study on the diffusion of innovation. This theory proposes that innovativeness is “the degree to which an individual is relatively earlier in adopting an innovation than other members of his system” (Rogers, 2003). Following Rogers and Cartano’s (1962) seminal work on diffusion, many customer studies have measured innovativeness using time of adoption method as an indicator of a customer’s innovativeness.

The applicability of the IDT and the TPB has been used by researchers in the context of CCSI in the hotel industry. For example, technological innovation in the tourist accommodation activity was discussed in the context of the hospitality service sector (Orfila-Sintes et al., 2005).

Customer innovativeness refers to “a customer’s tendency to adopt new products more often and more frequently as compared to other customers” (Moreau et al., 2001). Innovative customers tend to be more involved and responsive to firms’ offerings than non-innovative customers (Walczuch et al., 2007). Studies show that in the hospitality industry, social media can be used effectively to entice customers who are innovative and inclined toward technology to co-innovate. Customers thus may actively engage with the hotel to co-create and innovate hotel services (Morosan, 2015).

Customer innovativeness is a characteristic of humans reflecting the extent to which they accept novelty (Midgley and Dowling, 1978) and which helps to explain their adoption intention toward innovative products, services, and technologies. Moreover, innovativeness has been used extensively in the marketing literature, especially in segmentation, providing tools for marketers to distinguish between innovative and non-innovative customers (Agarwal and Prasad, 1998). Thus, it is possible for hotel firms to attract customers who show interest in new services to participate in co-creation and innovation of hotel services through social media. We propose the following hypotheses:

\[ H1. \] A guest’s innovativeness positively influences his/her intention to co-create new services with hotels on social media.

\[ H2. \] A guest’s innovativeness positively influences his/her adoption intention toward co-creatively developed new services on social media in hotels.

2.3 Theory of planned behavior (TPB)
The TPB (Ajzen, 1985, 1991) explains how an individual’s attitude toward a behavior, subjective norm, and perceived behavioral control factors influence his/her intention to perform a given behavior (Ajzen, 1991). Ajzen and Fishbein (1980) proposed the theory of reasoned action (TRA) to explore human behaviors during the decision-making process (e.g. Ajzen and Fishbein, 1980; Fishbein and Ajzen, 1973). This theory suggests that
individuals are rational and affected by motivation in their decision-making processes, and make a reasoned choice among various alternatives (Fishbein and Ajzen, 1975). The TPB was developed as an extension of the TRA. TPB considers both volitional and non-volitional control to explain the behavior of an individual. In addition, it has been found that both the TRA and TPB provide the most accurate prediction of a particular intention of a person’s behavior (Fishbein and Ajzen, 1975). Therefore, it has been suggested that behavioral intention is a precursor of the actual behavior (Ajzen et al., 2009). In the hotel service context, co-creation intention has been conceptualized as a hotel customer’s readiness/willingness to co-create a new hotel service in future. Specifically, intention is based on such variables as attitude toward the behavior, subjective norm and perceived behavioral control.

TPB has been previously used in the extant literature to study ethical behavior (Flannery and May, 2000) and internet activity (Hsu and Chiu, 2004). Han et al. (2010) applied TPB and showed its applicability in the hospitality context. Kim et al. (2016) applied TPB to analyze the selfie-posting behavior on social networking sites. Chen and Lu (2011) studied the modeling e-coupon proneness as a mediator in the extended TPB model to predict consumers’ usage intentions. Barsky (1992), Ryu and Han (2010) and Morgan and Hunt (1994) also studied the significance of the variables to explain customer post-purchase behaviors.

TPB has been used to explain customers’ behavior and their adoption of new concepts and systems (Ajzen, 1993). In the context of this study, it is assumed that the three antecedents of intention to co-create (attitudes, subjective norm and perceived behavioral control) influence hotel guests’ adoption intention toward co-creating new services on social media.

2.3.1 Attitude toward CCSI through social media. Attitude toward CCSI describes “a positive or negative evaluation or appraisal of the behavior” (Kraus, 1995). The TPB explains attitude toward a behavior as “the degree to which a person has a favorable or unfavorable evaluation or appraisal of the behavior in question” (Ajzen, 1991). In this study, the target behavior is the adoption intention toward co-creation to develop new hotel services on social media. It is proposed here that customers’ positive attitudes toward co-creating new services with the service provider will positively influence the intention to adopt co-creation to develop new services. According to Fishbein and Ajzen (1975), an individual’s attitude toward a behavior is one of the most significant predictors of both his/her intention to engage in that behavior and his/her actual behavior. Therefore, we propose that when customers develop a favorable attitude toward co-creation on social media, they may show interest to initiate a strong feeling of fulfilling the co-creation beliefs. We propose the following hypothesis:

H3. A guest’s attitudes toward co-creative hotel service innovation positively influence his/her intention to co-create on social media.

2.3.2 Subjective norm and intention to co-create. Subjective norm and intention to co-create suggests that subjective norm is another determinant of intention and defined it as “the perceived social pressure to perform, or not to perform, the behavior” by the individual. It is the support or advocacy of others in the mutual society to perform a specific behavior. However, building social relationships is also possible in social media context via social network technologies (Ellison et al., 2007). In the hospitality context, the subjective norm may influence hotel guest’s intention to co-create service innovation. Hotel guests may consider it socially desirable and permissible within their social network to co-create new services. Prior studies on subjective norm were applied to explore users’ online purchase intention (Pavlou and Fygenson, 2006), adopt blog (Hsu and Lin, 2008), use advanced mobile services (López-Nicolás et al., 2008) and online community users’ participation intention (Zhou, 2011).
This study is conducted to understand subjective norm in the context of CCSI behavior on social media. We propose the following hypothesis:

$$H_4.$$ Subjective norm positively influences a hotel guest’s intention to co-create new hotel services on social media.

2.3.3 Perceived behavioral control and co-creation on social media. The construct ‘perceived behavioral control’ is explained as “the degree to which an individual perceives the performance of a specific behavior as easy or difficult, referring to the perception of control he or she has over the behavior” (Ajzen, 1991). It was also found that it “reflects past experience as well as anticipated impediments and consequences” (Ajzen, 1991). Two factors (e.g. control beliefs and perceived facilitation) comprise “perceived behavioral control” and an individual’s perception toward the presence or absence of external constraints also determine his/her behavioral intention (Ajzen, 1991). An individual’s co-creation intention therefore is influenced by his/her perceived control over performing a behavior on social media (Chu et al., 2016). Thus, this study emphasizes that hotel guests perceived behavioral control influences their intention to co-create and adoption intention toward co-creatively developed new services on social media. We propose the following hypotheses:

$$H_5.$$ Perceived behavioral control positively influences guests’ intention to co-create new hotel services on social media.

$$H_6.$$ Perceived behavioral control positively influences hotel guests’ adoption intention toward co-creatively developed new services on social media.

2.3.4 Intention to co-create and co-creation adoption intention in social media. Behavioral intention is the core of the TPB and explains an individual’s intentions to perform, or not to perform, one certain behavior and proposes intention as an immediate antecedent of actual behavior (Ajzen, 1991). Fishbein and Middelstadt (1995) consider a person’s intention to be the most appropriate predictor for its behavior. This study focuses on hotel guests’ intention to co-create new services on social media as an important predictor for the actual adoption of co-creatively developed new services. TPB further suggests that by developing a stronger intention to co-create, hotel guests will devote more effort to co-create services (e.g. sharing information and knowledge or providing feedback) that will result in positive adoption intention toward co-creation of new services on social media. We propose the following hypothesis:

$$H_7.$$ A guest’s intention to co-create new hotel services will positively influence his/her adoption intention toward co-creatively developed new services on social media.

2.4 Mediating role of co-creation intention

It is noteworthy that co-creation intention has attracted the attention of the researchers due to its predictive power on final adoption or purchase of new services. In extant literature, an increasing number of studies have examined customers’ co-creation intention as a behavioral intention in IT and information systems (IS) (see Heidenreich and Handrich, 2015). Prahalad and Ramaswamy (2004) and Morosan (2015) viewed co-creation intention as a study to examine the dimensionality of co-creation such as effort, personalization, information sharing, etc. Hess et al. (2014) and Legris et al. (2003) consider co-creation intention instead of actual behaviors as the best surrogates of actual behavior (Morosan, 2015).

Customer intention toward future co-creation is defined as a customer’s willingness to participate in service production and delivery in the future. As customers participate more
in service recovery, the skills and confidence they need to complete the task are improved (Dong et al., 2008). Previous studies have examined antecedents of co-creation intention, for example, personal innovativeness, trust and perceived personalization (Heidenreich and Handrich, 2015; Moorman et al., 1992), positive perceived value and satisfaction with the service provider (Dong et al., 2008). However, most of the studies of customers’ co-creation intention were conducted based on theories of innovation, and the intention is studied as a predictor of customers’ intention.

Figure 3 incorporates the drivers influencing customer adoption intention as variables in a model showing how customers and hotels co-creatively develop new services. In this conceptualization, co-creation is the central mechanism by which hotels improve their exploration, creation and delivery of new services. The logic of the proposed model is that service innovation requires a customer’s innovativeness, attitude toward CCSI on social media, subjective norm and perceived behavioral control to develop co-creation intention that finally results in positive adoption intention. Therefore, this research examines the effect of co-creating customers’ innovativeness, attitude toward CCSI social media, subjective norm and perceived behavioral control as drivers that will lead to positive behavioral intention, which further ensure the service firm’s competitive advantage in the market. We propose that co-creation intention acts as a mediator between the antecedents of adoption intention (customer innovativeness, attitude toward CCSI on social media) and adoption intention toward co-creating new services. We propose the following hypotheses:

H8. Customers’ intention to co-create mediates between customer innovativeness and adoption intention.

H9. Customers’ intention to co-create mediates between attitude toward CCSI on social media and adoption intention.

2.5 Hypotheses
Based on the theoretical framework discussed above, the following nine hypotheses are proposed:

H1. A guest’s innovativeness positively influences his/her intention to co-create new services with hotels via social media.

H2. A guest’s innovativeness positively influences his/her adoption intention toward co-creatively developed new services via social media in hotels.

H3. A guest’s attitudes toward co-creative hotel service innovation positively influence his/her intention to co-create via social media.

H4. Subjective norm positively influences a hotel guest’s intention to co-create new hotel services via social media.
H5. Perceived behavioral control positively influences guests’ intention to co-create new hotel services via social media.

H6. Perceived behavioral control positively influences hotel guests’ adoption intention toward co-creatively developed new services via social media.

H7. A guest’s intention to co-create new hotel services will positively influence his/her adoption intention toward co-creatively developed new services via social media.

H8. Customers’ intention to co-create mediates between customer innovativeness and adoption intention.

H9. Customers’ intention to co-create mediates between attitude toward CCSI on social media and adoption intention.

3. Methodology
The main purpose of this study is to measure the effect of driving factors of co-creation intention on social media by estimating the relationships between the customer innovativeness, attitude toward CCSI in social media, subjective norm and perceived behavioral control, co-creation intention and adoption intention. The theoretical framework of present research is illustrated in Figure 3. The latent variable of customer innovativeness was measured using four items based on the work of Goldsmith and Hofacker (1991) and Bruner and Kumar (2007). Following Taylor and Todd (1995), attitude toward CCSI on social media was measured using four items. Adapted items from Taylor and Todd (1995), both subjective norm and perceived behavioral control was measured by five items, subjective norm by two items and perceived behavioral control by three items. Co-creation intention on social media was measured by four items borrowed from the study of Prahalad and Ramaswamy (2004). The seven items for adoption intention were adapted based on prior studies of Handrich and Heidenreich (2013) and Zwass (2010). All adapted items were modified to fit into the hotel context, and were presented in a seven-point Likert-type scale with “Agree–Disagree” anchor statements (see Table AI).

By following Mohsin and Lockyer (2010), a structured questionnaire was developed for this study. The questionnaire was pre-tested to confirm both its reliability and validity. Thus, to ensure that the survey instrument was consistent, an expert opinion was taken, and a pilot study was conducted. The expert panel consisted of five members (two professors who have expertise in tourism and hospitality and three senior doctoral fellows pursuing their PhD in tourism and hospitality). All experts had more than three years of experience in their respective fields. The pilot study was conducted using 42 students (master- and doctoral-level students) from a large university in India. All selected students were social media users and had stayed or slept for at least one day and one night in a star category hotel in the three months preceding the date of survey. Most of panel and participants commented that the instrument was easy to read and understand whereas some suggested that statements should be more specific. Thus, in the statements, some wordings were changed to reflect the comments from the pilot study.

3.1 Data collection and sample
Data were collected for this study from the State of New Delhi where the highest number of luxury five-star category hotels are located in India. The Indian luxury hotel business has observed continued growth in last few years, and it is projected that the number of hotels will increase from 750 (2012) to 1,338 (2017) and 2,457 (2022) (India Tourism Statistics, Ministry of Tourism Government of India, 2012). There are a total of 92 five-star luxury category hotels in India; New Delhi has six of them (India Tourism Statistics, Ministry of Tourism Government of India, 2012).
Tourism Government of India, 2012). Luxury hotels are categorized into five-star deluxe, five- and four-star and Heritage hotels (India Tourism Statistics, Ministry of Tourism Government of India, 2012). We approached all luxury hotels; however, only five-star deluxe category luxury hotels agreed to participate in the data collection process. Thus, this present research provides an opportunity to gauge a largely under researched area of customers and their participation in the co-creation of service innovation in the five-star deluxe category hotels in India. Additionally, this study also offers a possibility to compare and contrast with other similar studies undertaken in other countries.

To ensure the eligibility of respondents, hotel guests were asked an initial set of screening questions: whether they engage in and actively participate in social media site (e.g. Facebook, Instagram, Twitter, etc.; whether they have subscribed, liked or joined any e-travel service companies’ brand page using any social networking site; and whether they ever posted or considered reviews and ratings of hotels listed on e-travel service companies’ websites via their official site or via mobile app while planning their hotel bookings'. Additionally, to ensure respondents’ co-creation in hotel service innovation, questions were asked: whether they have shared their actual experience about hotel using social media; and whether they have ever provided their feedback regarding a new service offering or modifying any existing service on e-travel service companies’ website in return of earning e-cash or points to be used for booking of hotels in future.

The survey comprised of three parts: first, the initial screening questions; second, the investigative questions, probing hotel guests’ feedback regarding hotel service offerings and their suggestions (e.g. how was their stay at a particular hotel? Have they written reviews on a travel community website using survey link and rated their stay? Have they been featured on the hotels’ newsletter & blog? etc.); and the third part was used to collect demographic details of the respondents in terms of gender, age, type of trip (business or holiday) and country of residence.

A questionnaire including 24 statements was used and translated from English to Hindi and then translated back to English to insure the accuracy of translation (Mount and Back, 1999). All of the hotel guests’ reported variables were measured using a seven-point Likert-type scale.

Data were collected from the hotel guests while the guests are in the hotel (lobby, front office, reception counter, room service and inner café/restaurants or other convenient locations inside the hotels). Throughout the survey, all comments made by hotel guests were written down and the survey team provided clarification as needed to the guests while they completed the survey questionnaire. A number of guests indicated their preference to complete the survey in their room; these guests returned the completed survey questionnaire to the hotel reception desk. Hotel guests had the option to withdraw their involvement at any stage during the survey. However, the respondents were considered to be familiar with the hotel and the services offered to them as guests. Finally, the presence of the researchers further ensured the clarity and specificity of the responses received.

A total of 439 questionnaires were distributed to hotel guests, and 346 usable responses were obtained, with a response rate of 78 percent that fulfilled the general range for using “structural equation modeling” (SEM; in which 5–10 per item ratio is suggested; Bollen, 1989). Hair et al. (2002) also suggested that 10:1 ratio (sample to item) is required for the multivariate data analysis. For this study, the sample size was above the minimum required ratio, i.e. 10:1. The non-response partiality was tested as well as a comparison of “early” and “late” respondents' responses based on the date of the received replies to the questionnaire (Armstrong and Overton, 1977). Further, to apply the t-test (sample mean comparison), the first 30 percent of responses was compared with last thirty percent and thus, the independent sample of t-test displayed an insignificant difference between these two groups, implying no case of unit non-response bias with the collected data.
3.2 Analytical strategy

SEM was used with AMOS version 20 to analyze the collected data and to test the proposed hypotheses. The SEM technique provides a “maximum-likelihood estimation” of the entire system in a proposed model, which allows the measurement of variables aligned with the data collected (Joreskog and Sorbom, 1982, 1983). In this study, the researchers adopted a two-step strategy as outlined by Anderson and Gerbing (1984) to test the hypotheses and research model. For this data, researchers first used a confirmatory factor analysis (CFA) to assess the measurement model. Second, a structural model using SEM was used to estimate the goodness-of-fit indices of proposed hypotheses and research model. In addition, researchers reported the value of chi-squared ($\chi^2$) as an absolute fit index to determine how well this model fit the data, with a number of fit indices (“comparative fit index [CFI], Tucker–Lewis index [TLI], root mean square error of approximation [RMSEA]”), which permitted the researchers to compare the hypothesized model with the alternative models. The more than or equal to 0.90 value of CFI and TLI shows a good fit to the data (Hu and Bentler, 1999). Similarly, the “RMSEA” is a measure of “average standardized residual per degree of freedom,” and its value equal to and lower than 0.08 shows that model appropriately fits the data (Hu and Bentler, 1999). Scale items are presented in Table AI. The descriptive statistics with correlation matrix is described in Table II.

4. Results

Along with testing the hypothesized model, we also examined the measurement model’s goodness-of-fit indices, reliability, convergent and discriminant validity.

4.1 Confirmatory factor analysis

Initially, to determine the measurement model, CFA was performed using AMOS version 20, and the significance of the entire model with relationships between all items and scales were tested statistically. CFA was performed with six factors including 24 items. We conducted CFA in order to evaluate the quality of factor structure through statistical testing of the full model (difference between scales) and linkage between scales and its items. In order to test the proposed hypothesis, SEM was applied to the primary data obtained via survey method, which has been used in similar context with similar topics (e.g. in this study, each construct along with their corresponding items depicted a high level of internal consistency, as the value of Cronbach’s $\alpha$ for each of them were found to be more than the minimum acceptable value, i.e. 0.70) (Nunnally, 1978). The final questionnaire’s overall internal consistency was found satisfactory. The composite reliability (CR) value was observed to be above the standardized cut-off value, i.e. 0.70 (Fornell and Larcker, 1981), and ranged from 0.77 to 0.93 (Table IV). Thus, findings recommend adequate reliability of the constructs. The 24 items loaded significantly on six factors, and were confirmed an overall goodness-of-model fit to the data (see figure 4). All these values were within the standardized limits, thus showing a good fit between the observed data and the model. The value of average variance extracted (AVE) was found to be above 0.50 for all six factors, and thus met the cut-off range suggested by Bagozzi and Yi (1988) and Fornell and Larcker (1981). Findings are depicted in Table III.

4.2 Assessment of construct reliability

The reliability of constructs was assessed using Cronbach’s $\alpha$ and CR. For the CR estimate, the rule of thumb states that the value at least 0.70 or more than indicates adequate reliability. The value of CR depicts the internal consistency among all items measuring the particular construct (Fornell and Larcker, 1981). The values of Cronbach’s $\alpha$ and CR
were more than 0.70, representing that there was internal consistency (i.e. all measures consistently indicated reliability). The values lie within suggested threshold criteria suggested by Nunnally (1978). Table III depicts these values for each construct.

4.3 Convergent and discriminant validity analysis

In addition to analyzing the measurement model in Table IV, it was also used to assess the validity (i.e. convergent and discriminant) of the survey instrument. Convergent validity is assessed through following criteria: at least or above 0.70 value of “composite reliability (CR)”
greater than the 0.50 value of “average variance extracted (AVE)” (Fornell and Larcker, 1981) and in CFA greater than the 0.70 value of all items loadings (Nunnally, 1978). All these criteria for convergent validity are met in this study (see Table III, Figure 4), thus there are no convergent validity concerns. The discriminant
validity examined by the following criteria: the value of AVE should be greater than 
"Average Shared Variance (ASV)" and "Maximum Shared Variance (MSV)." The criteria for 
discriminant validity are also fulfilled in this paper (see Table III); consequently, there is no 
issue for convergent validity.

4.4 Structural model

In this study, the establishment of the measurement model followed SEM. Further, 
considering Moghavvemi and MohdSalleh (2014), the conceptual model of this study was 
evaluated through goodness-of-fit index for the measurement model (Table IV). All seven 
indices were found suitable with respect to their threshold value, suggesting that responses 
were fitted best in the proposed model, making it valid and adequate for validation reasons. 
The result demonstrated that the empirical model revealed a good overall fit with the data. 

The SEM analysis facilitated the standardized path coefficients that were shown in Figure 5. 
In sum, the respective \( t \)-value was given in Table V. This table depicted that all proposed 
hypotheses were found to have significant relation between constructs at 0.001, 0.01 and 
0.05 levels. Thus, all proposed hypotheses of this study were found to be accepted (see Table V).

Customer innovativeness, one of the key variables that significantly influences hotel 
guests’ co-creation intention (e.g. Morosan, 2015), is expected as the guests were willing to 
contribute in the service development process. Another important result of this study 
depicts that customer innovativeness would also significantly affect their adoption intention 
(e.g. Im et al., 2003). This is obvious if something innovative is presented to the customers, 
they show more willingness to become involved or to co-create with the service provider. 
In addition, the results of this study also show that attitude toward co-creative hotel service
innovation on social media significantly affects co-creation intention (Cheung and To, 2016) that further results into hotel guests’ adoption intention toward new services developed through co-creation (e.g. Taylor and Todd, 1995). Another finding concerns the positive significant relationship between subjective norm and guests’ intention to co-create (Cheung and To, 2016). Perceived behavioral control significantly affects guests’ intention to co-create (Cheung and To, 2016). Intention to co-create acts as an antecedent and significantly influences guests’ adoption intention (Cheung and To, 2016).

4.5 Mediating effects testing

Mediation analysis was performed to determine whether co-creation intention mediates the relationship between two predictor variables (customer innovativeness and attitude toward CCSI on social media) and adoption intention as an outcome variable. A bootstrap process using SEM was employed through AMOS to analyze the mediation effect, which also helps to overcome certain key problems associated with the Baron and Kenny (1986) approach (i.e. unable to explain and provide statistical test for indirect effect caused by independent variable on dependent variable via proposed mediating variable). On the other hand, the Sobel test for mediation facilitates undefined assumption of normality in case of small samples and possibility of Type I error (Preacher and Hayes, 2004). The SEM model was tested using these bootstrap procedures to produce 1,000 estimates of each path coefficient. Additionally, the indirect effect of predictor variable is estimated through the output from 1,000 estimates of each path coefficient.

Results of the bootstrap analysis (see Table VI) indicated that co-creation intention partially mediates between both predictor variables (customer innovativeness and attitude toward CCSI on social media) and outcome variable (adoption intention).

The direct effect of customer innovativeness to adoption intention was found to be significant at the 0.001 level. Similarly, when analyzed, the indirect effect between customer innovativeness to co-creation intention to adoption, it was also found significant at the
same level. Thus, the study result provides evidence regarding the existence of co-creation intention as a partial mediation between customer innovativeness and adoption intention. Similarly, co-creation intention mediates the relationship between attitude toward CCSI on social media and adoption intention. The degree of mediation significance was at the 0.001 level. Both the direct and indirect path found to be significant and thus confirm the existence of co-creation intention as a mediator between the aforementioned relationships. The detail of mediation analysis is provided in Table VI.

5. Discussion and implications
A conceptual model was designed to highlight a set of relationships in-between the constructs that were finally tested and validated. This validation provided evidence that customers’ innate characteristics (consumer innovativeness) can be used to predict the manner in which customers become involved in service innovation through co-creation in hotel service innovation contexts. This validation extended the scope of SDL to the social media facilitated service ecosystem in co-creative hotel service innovation (Table VII).

The findings showed strong support for the proposed model, in which co-creation intention partially mediates the link between customer innovativeness and adoption intention; similarly, it partially mediates between attitude toward co-creative hotel service innovation and adoption intention on social media. Our results indicate that the customer innovativeness is a positive determinant of adoption intention on social media. This finding implies that customers require the prerequisite of the belief that social media is useful in obtaining timely and valuable information. Customers also consider social media in assisting them in making strategic decisions on adopting services before the customers develop a positive attitude toward co-creation. These finding suggests that tourists/hotel guests who are accessing social media place a higher value on having a stronger interest in social media and will be highly motivated to use social media as a key means of co-creating new and better services. This finding further suggests that active role of co-creating customers during hotels’ service innovation process.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Hypothesis path</th>
<th>Coefficient</th>
<th>t-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Consumer innovation → co-creation intention</td>
<td>0.47***</td>
<td>9.39</td>
<td>Supported</td>
</tr>
<tr>
<td>H2</td>
<td>Attitude toward CCSI in social media → co-creation intention</td>
<td>0.13*</td>
<td>2.26</td>
<td>Supported</td>
</tr>
<tr>
<td>H3</td>
<td>Subjective norms → co-creation intention</td>
<td>0.23***</td>
<td>4.77</td>
<td>Supported</td>
</tr>
<tr>
<td>H4</td>
<td>Perceived behavior control → co-creation intention</td>
<td>0.11*</td>
<td>1.96</td>
<td>Supported</td>
</tr>
<tr>
<td>H5</td>
<td>Co-creation intention → adoption intention</td>
<td>0.32***</td>
<td>4.95</td>
<td>Supported</td>
</tr>
<tr>
<td>H6</td>
<td>Consumer innovation → adoption intention</td>
<td>0.19**</td>
<td>3.04</td>
<td>Supported</td>
</tr>
<tr>
<td>H7</td>
<td>Perceived behavior control → adoption intention</td>
<td>0.28***</td>
<td>4.84</td>
<td>Supported</td>
</tr>
</tbody>
</table>

Notes: *,**,***Significant at the 0.05, 0.01 and 0.001 levels, respectively

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Hypotheses</th>
<th>Direct effect</th>
<th>Indirect effect</th>
<th>Total effect</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H8</td>
<td>Consumer innovativeness → intention to</td>
<td>0.141***</td>
<td>0.145***</td>
<td>0.286***</td>
<td>Partial</td>
</tr>
<tr>
<td></td>
<td>co-create → adoption intention</td>
<td></td>
<td></td>
<td></td>
<td>mediation</td>
</tr>
<tr>
<td>H9</td>
<td>Attitude toward CCSI in social media → intention to</td>
<td>0.390***</td>
<td>0.058***</td>
<td>0.448***</td>
<td>Partial</td>
</tr>
<tr>
<td></td>
<td>co-create → adoption intention</td>
<td></td>
<td></td>
<td></td>
<td>mediation</td>
</tr>
</tbody>
</table>

Notes: Model fit: $\chi^2$ (425.0), $p < 0.00$; GFI = 0.91, NFI = 0.92, CFI = 0.95; RMSEA = 0.06. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001
This study findings support all nine hypotheses. Innovative customers show more intention to co-create (H1) which leads to adoption intention (H2). Attitude toward co-creative hotel service innovation (CCSI) leads to intention to co-create (H3). The subjective norm leads to intention to co-create (H4) and perceived behavioral control leads to intention to co-create (H5) and affects adoption intention (H6). Intention to co-create significantly affects a customer’s adoption intention (H7). Additionally, customers’ intention to co-create partially mediates between customer innovativeness and adoption intention (H8). Similarly, attitude toward CCSI on social media and adoption intention are partially mediated by customers’ co-creation intention (H9).

Our findings are also supported by the result of the previous studies. Customer innovativeness significantly influences customers’ adoption intention (Morosan, 2015; Grissemann and Stokburger-Sauer, 2012). Similarly, a guest’s attitude toward CCSI positively influences his/her intention to co-create with the service firm (Morosan, 2014). Subjective norm significantly influences co-creation intention (Cheung and To, 2016). Perceived behavioral control was found to positively influence intention to co-create new services (Kim et al., 2016). Similarly, intention to co-create positively influences adoption intention (Cheung and To, 2016). Finally, the mediating role of customers’ intention to co-create was tested between predictors (customer innovativeness and attitude toward CCSI on social media) and outcome variable (adoption intention) and was found to be partially mediating. Thus, the results of mediation are consistent with the previous studies (Morosan, 2015; Cheung and To, 2016).

This study implies that “customer innovativeness,” “attitude toward co-creative service innovation,” “subjective norm” and “perceived behavioral control” are significant predictors of hotel guests’ adoption intentions toward new services developed co-creatively developed on social media. Hotel firms should find ways to equip themselves with skills and expertise in technology-mediated service development and thus, a new spectrum of co-creation can be developed in the virtual world. As Ajzen (1991) and Perugini and Bagozzi (2001) suggested that modifying the TPB model by altering paths and including additional critical constructs in different context may contribute to improve the understanding of the theoretical mechanism of the model. It will also help to predict customers intention/behavior in the hotel context. Thus, the present study has broadened and deepened present understanding of the theory by analyzing co-creation intention in hotel service innovation context (Perugini and Bagozzi, 2001). Moreover, the limited amount of previous research on the application of a blend of theories, namely, the theory of diffusion of innovation (Rogers, 2003) and TPB (Fishbein and Ajzen, 1975; Ajzen, 1991), which also motivate the researchers to apply the theories to identify the underlying factors and conditions that drive hotel guests to co-create new services and determine their future co-creation intention using social media platforms, is an unique contribution of the study to the existing literature.

5.1 Theoretical implications
The present study comprises several theoretical and practical implications and contributes to the literature in online information research and services marketing in several ways.

First, this study delivers insight on the customer–hotel firm where communication has occurred via social media. It validates the existing findings on customer participation in social media-based brand communities.

Second, it investigates the impact of consumer characteristics in technology-mediated innovation in hotel service settings (Cabiddu et al., 2013). This represents instrumental mechanisms to uniquely facilitate co-creation of service innovation and extends the scope of service ecosystem literature. This provides a useful theoretical base for understanding the mechanisms of CCSI in which human interaction coexists with technology-based interaction in the social media context.
Third, through empirical verification, this study ascertained the impact of customers’ psychological factors as key drivers of their behavioral intention to co-create and adopt. This study fills an important gap to understand how hotels could develop innovative new services using co-creation of services attributable to their customers’ use of social media.

Fourth, as social media is increasingly used in marketing, this study represents a useful theoretical basis for understanding the mechanisms that contribute to value creation. In most service contexts, the delivery of services happens when customer–firm interaction coexists with technology-based interaction (Cabiddu et al., 2013). Thus, the study provides empirical support to verify the significance of customer–firm interaction using social media to positively affect their business. This study also highlights the importance of co-creation of service innovations as it contributes to marketing and consumer behavior (Dellarocas, 2003; Litvin et al., 2008).

Finally, empirical findings confirm CCSI as an important determinant of customers’ adoption intention, which indicates that innovation activities are vital to the hotel’s success.

5.2 Practical implications

The results of this research offer numerous managerial implications that have relevance in the hospitality industry, particularly in the hotel sector. Engaging hotel customers using social media is particularly important in order to entice customers to co-create new services.

First, managers may well focus on their innovative customers and persuade them to become actively involved in CCSI. Managers can segment the market on the basis of innovative vs non-innovative customers (Agarwal and Prasad, 1998). Managers can select and target the innovators and early adopters to co-create in new service innovation as they are the opinion leaders of the new services to other potential customers.

Second, the study results also provide insightful information about how brands can use social media to enhance customer engagement and positive customer attitude toward brand community participation. Therefore, it is important for hotel managers to pay regular attention to customer-generated social media content and to be able to recognize the constructive facts within their social network.

Third, considering the importance of customers’ positive attitude and perceived behavioral control, managers should adopt measures to help customers with the co-creation of new services using social media. To overcome psychological barriers, firms should encourage customers to become actively involved in information sharing and suggest modifications on existing services and help them to co-design new services (Kim et al., 2016).

Fourth, the present study offers insightful ideas that may contribute to an organization and brand community managers to manage the organization’s social media. Brand communities should develop strategies to increase customers’ self-confidence through continuous interaction and engagement with the community. Hotel firms should adopt advance technologies for more interactive applications in user interfaces on social media; for example, they may include more photo-sharing applications or easily customized interfaces (Kim et al., 2016).

To engage customers to co-create hotels can offer incentives (e.g. rewards, discounts, or privileges) to customers who post interesting messages, valuable suggestions, share photos or videos of their experience with reference to the hotels’ new offerings on social media (e.g. “hotels may suggest the most Instagram-able suites or cuisine to patrons”; Tussyadiah, 2015). These may provide topics of conversation through “a hashtag campaign or other informational cues.” Incentives may include free vouchers for products and services. Thus, hoteliers may provide customers directions to enhance certain features of their products and services through discussion on social media. Such incentive may promote customers to co-create on social media and may simultaneously result in many social network members following each other (i.e. consumption convergence) that may result in promoting a hotel service.
5.3 Limitations and directions for future research

Although the findings expand the extant knowledge on the impact of customer co-creation on service innovation in the hotel sector, several limitations are recognized that must be taken into account when generalizing this study’s results.

Given that this study was conducted in hotels located in one country (India), one sector and one country context may limit the generalizability of the study findings. Therefore, future research should aim to undertake similar research in multiple industry sectors and in different countries. Future research can also aim to identify and distinguish between male and female affinity to co-create services. Future study may determine whether the cultural background of the customers may influence their willingness to co-create. In this study, the mediating role of two drivers of co-creation intention (customer innovativeness and attitude toward CCSI in social media) has been examined. It will be useful to understand the mediating role of other drivers of co-creation intention in the proposed model in future research. Moderating variables that influence the findings may exist; therefore, future research should undertake mediation moderation analysis. Future research may investigate co-creation between the firm, suppliers and employees. Future researchers can use this proposed model on customer CCSI in similar or other service contexts. Moreover, Wang and Wang (2010) suggested that users’ behavioral intentions cannot be entirely explained by the existing technology adoption models; that is why future research may develop and test a multidisciplinary theoretical model, rooted in combination of theories, namely, technology acceptance model (Davis, 1989), social cognitive theory (Bandura, 1986), technology task fit theory (Goodhue and Thompson, 1995) and self-determination theory (Ryan and Deci, 2000a), etc. This approach may be effective in linking important constructs of both social science research and online information research.

6. Conclusions

This study reflects the increasing use of social media as an effective medium to enhance firm-customer interaction. Hotel firms have thus been able to engage customers to co-create and develop new and innovative services through social media as an effective medium of communication. Applying Roger’s diffusion of innovation theory and Ajzen’s TPB, this study proposes a conceptual model to determine the influence of antecedents (customer innovativeness, attitude toward CCSI on social media, subjective norm and perceived behavioral control) to predict hotel guests’ intention to adopt co-creatively developed new services.

In this regard, hoteliers should proactively engage with their customers to understand their needs and wants and then respond with new services to satisfy them (Thomke and Von Hippel, 2002).

This study contributes to the existing body of knowledge in co-creation, service innovation and the technology-mediated service ecosystem. Using the hotel industry as a context, this study undertook a comprehensive conceptualization of CCSI. Findings from this study have implications and extensions possible in other sectors such as travel (e.g. tour operators, transportation and food service) and other service industries (e.g. education, financial services and health care). This study is one of the first to examine the components of the theory of diffusion of innovation and the TPB in the context of CCSI using social media as the medium to bring the firm and customers closer. Similarly, the use of social media to facilitate CCSI could represent the inception of personalized interactions leading to new service development. Finally, we conclude that firms can enjoy superior business performance by adopting CCSI strategies facilitated by social media.
References


Facebook.com (2011), “Facebook now has more than 800 million active users”, available at: www.facebook.com/notes/statspotting/facebook-now-has-more-than-800-million-active-users/204500822949549


Further reading


## Appendix

### Scale for measuring consumer innovativeness (INNO)

<table>
<thead>
<tr>
<th>INNO1</th>
<th>If I heard about a new technology, I would look for ways to experiment with it</th>
</tr>
</thead>
<tbody>
<tr>
<td>INNO2</td>
<td>Being the first to use new high-tech services is very important to me</td>
</tr>
<tr>
<td>INNO3</td>
<td>I can usually figure out new high-tech products and services without help from others</td>
</tr>
<tr>
<td>INNO4</td>
<td>Among my peers, I am usually the first to explore new information technologies</td>
</tr>
</tbody>
</table>


### Scale for measuring adoption intention (AI)

| AI1 | I usually buy products/services to be consumed during the current trip using smart phone apps to access social media |
| AI2 | I regularly make an online review of the current hotel services on SNSs via smart phone apps |
| AI3 | I give updates about my current trip on SNSs using smart phone apps |
| AI4 | I sometimes generate my own smartphone network within the hotel to access SNSs |
| AI5 | I sometimes connect to other in-room technologies to access SNSs using smart phone apps |
| AI6 | I like to share information related to hotel room atmospherics on SNSs using smart phone apps |
| AI7 | I try to provide information regarding accessing my room on SNSs using smart phone apps |


### Attitude toward CCSI in social media

| ACCS1 | I assume co-creative service innovation on SNSs via smart phone apps is a good idea |
| ACCS2 | I believe using co-creative service innovation on SNSs via smart phone apps is a wise idea |
| ACCS3 | I like the co-creative service innovation idea on SNSs via smart phone apps |
| ACCS4 | I think the use of co-creative service innovation on SNSs via smart phone apps is a pleasant idea |

| Items adopted | Taylor and Todd (1995) |

---

Table AI.

List of scale items adopted for the study (continued)
Table AI.

<table>
<thead>
<tr>
<th>Scale Item</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjective norm 2 items adopted</td>
<td>Taylor and Todd (1995)</td>
</tr>
<tr>
<td>SN1 I think users who are important to me would think that I should co-</td>
<td>create services on SNSs via smart phone apps</td>
</tr>
<tr>
<td>create services on SNSs via smart phone apps</td>
<td></td>
</tr>
<tr>
<td>SN2 I think users who influence my behavior would think that I should</td>
<td>create services on SNSs via smart phone apps</td>
</tr>
<tr>
<td>create services on SNSs via smart phone apps</td>
<td></td>
</tr>
<tr>
<td>Perceived behavioral control 3 items adopted</td>
<td>Taylor and Todd (1995)</td>
</tr>
<tr>
<td>PBC1 I have the resources, knowledge and the ability to co-create services</td>
<td>on SNSs via smart phone apps</td>
</tr>
<tr>
<td>on SNSs via smart phone apps</td>
<td></td>
</tr>
<tr>
<td>PBC2 I would be able co-create services on SNSs via smart phone apps</td>
<td></td>
</tr>
<tr>
<td>PBC3 I think co-creating services on SNSs via smart phone apps is</td>
<td>entirely within my control</td>
</tr>
<tr>
<td>entirely within my control</td>
<td></td>
</tr>
<tr>
<td>Co-creation Intention 4 items adopted</td>
<td>Taylor and Todd (1995), Prahalad and Ramaswamy (2004), Handrich and</td>
</tr>
<tr>
<td>CCI1 I intend to co-create services on SNSs via smart phone apps</td>
<td>Heidenreich (2013), and Zwass (2010)</td>
</tr>
<tr>
<td>CCI2 I think during my trip, I will access SNSs via smart phone apps to</td>
<td>buy products/services to be consumed</td>
</tr>
<tr>
<td>buy products/services to be consumed</td>
<td></td>
</tr>
<tr>
<td>CCI3 I think during my trip, I will access SNSs via smart phone apps to</td>
<td>make an online review of my current trip</td>
</tr>
<tr>
<td>make an online review of my current trip</td>
<td></td>
</tr>
<tr>
<td>CCI4 I think during my trip, I will access SNSs via smart phone apps to</td>
<td>provide updates about my current trip</td>
</tr>
<tr>
<td>provide updates about my current trip</td>
<td></td>
</tr>
</tbody>
</table>

About the authors

Bijoylaxmi Sarmah has received her MBA degree from North-Eastern Regional Institute of Management under Dibrugarh University and has served as Assistant Professor at North-Eastern Regional Institute of Science & Technology, Arunachal Pradesh. She is currently pursuing her Doctoral research (PhD) at the Department of Management Studies, IIT Roorkee. Her works have been published in journals like *International Journal of Contemporary Hospitality Management, Benchmarking: An International Journal, International Journal of Culture, Tourism, and Hospitality Research*, etc. She has also awarded as “Young Service Researcher” at 8th IRSSM held at Yonsei University, South Korea in 2017 and has been the receiver of ACR-Sheth Grant for the year 2017. Her areas of interest are co-creation, service innovation and social media. Bijoylaxmi Sarmah is the corresponding author and can be contacted at: geetsarma@gmail.com

Dr Jay Kandampully is Professor of Service Management at The Ohio State University, USA. He is Editor in Chief of the *Journal of Service Management*, and serves on the editorial advisory board of 12 refereed international journals. Jay also serves as CTF International Fellow at the University of Karlstad, Sweden; International Fellow at the University of International Business and Economics, Beijing, China; and International Fellow at the University of Namur, Belgium. He holds a PhD in service management, and an MBA, with a specialization in services marketing, both from the University of Exeter, England. Dr Jay is Co-author of the book *Service Management Principles for Hospitality and Tourism*. He is Author of the book *Services Management: The New Paradigm in Hospitality* (translated into Chinese). He is also Editor of the following books: *Service Management: The New Paradigm in Retailing* (translated into Chinese); *Service Management in Health & Wellness Services*; *Customer Experience Management: Enhancing Experience and Value through Service Management*; and Lead Editor of the book, *Service Quality Management in Hospitality, Tourism and Leisure* (translated into Chinese, Korean and Arabic). His publications have appeared in journals such as *Journal of Service Management, European Journal of Marketing, Journal of Business Research, Cornell Hospitality Quarterly, The Service Industries Journal, Journal of Services Marketing, Managing Service Quality, Journal of Consumer Behaviour* and *The Journal of Product & Brand Management*.

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
Analysis of user-generated comments posted during live matches of the Cricket World Cup 2015

Snehasish Banerjee
The York Management School, University of York, York, UK

Abstract

Purpose – The purpose of this paper is to analyze user-generated comments posted on social media while live matches were being played during the Cricket World Cup 2015.

Design/methodology/approach – Data were collected from Yahoo! Cricket (YC), a website that allows people to submit comments during live matches. The comments were qualitatively analyzed using the grounded theory approach.

Findings – The key finding of this paper is that people like to consume live sporting events in an online social setting rather than as isolated individuals. In addition, the use of the grounded theory approach helped uncover several new findings related to people’s use of social media during live matches.

Research limitations/implications – Since this paper studied the case of the Cricket World Cup 2015 and collected data from YC, caution is advocated in generalizing its findings.

Originality/value – Scholarly interest on the use of social media during live sporting events is growing. Building on such works, this paper highlights how user-generated comments posted during the Cricket World Cup 2015 – mostly by individuals within the Indian subcontinent – intersected with broader issues such as culture, identity, politics and religion.

Keywords User-generated content, Social media, Qualitative analysis, Information sharing, Sport communication

Paper type Research paper

Introduction

Advances in information technology have revolutionized the production and consumption of live sporting events. Traditionally, sports used to be broadcasted on television for passive viewing. However, social media applications now allow people to express opinions freely on the internet, and actively engage with netizens during live matches.

As information technology continues to take over the world, sporting events need to become tech-friendly in reaching out to its increasingly tech-savvy fans. Such a development will allow forging strong sporting fraternity-fan connections, which lie at the heart of not only the development of sports but also the enhancement of fans’ experience (Bhogle, 2017). Recognizing this, scholars have been studying sports-related social media use over the last decade (e.g. Kassing and Sanderson, 2010; McCarthy et al., 2014).

An emerging area of research focuses on user-generated comments posted by internet users during live sporting events (Yu and Wang, 2015). This paper considers such comments to offer a unique vantage point into real-time sentiment on the ground. Hence, their analysis would extend the scholarly understanding of the use of information technology among people during live sporting events. Moreover, it could offer insights to sports marketers into ways to enhance the overall sport experience.

For these reasons, the purpose of this explorative-qualitative paper is to analyze user-generated comments posted on Yahoo! Cricket (YC) (www.cricket.yahoo.com), while live matches were being played during the Cricket World Cup 2015. Given the relatively under-investigated nature of the research area coupled with the lack of a priori frameworks, analysis of the comments was done using the grounded theory approach. Consistent with
the nature of any explorative-qualitative research, the thrust of this paper lies in providing a
rich description of the specific study context – not in its strive for generalizability.

The paper is significant for both theory and practice. From a theoretical standpoint, it
deepens the scholarly understanding of how sports fans use social media during live
sporting events. The premise that sports fans will use social media even while watching live
matches is conceptually rooted in the notion of second screening, which refers to people’s
growing penchant to use multiple screens at the same time (Shim et al., 2017). Hence, it is
quite feasible for individuals to watch live matches on a television, a computer or a
smartphone; and at the same time, post online comments using a computer or a mobile
device. Thus, by studying a relatively new phenomenon in a unique context – the Cricket
World Cup 2015, the paper has practical implications for sports marketers on how to foster
online engagement among fans.

The rest of this paper is structured in the form of five sections. The next section offers an
overview of the literature on sports-related social media use. The section thereafter details
the present study context and highlights its uniqueness. It describes the sport of cricket in
general, the tourney Cricket World Cup 2015 in particular and the social media platform YC.
This is followed by another section that explains the research methods including data
collection, coding, analysis and evaluation. The penultimate section of this paper presents
the results. The closing section is devoted to discussion and conclusions. It discusses the
findings in light of the literature, and highlights the contributions to theory as well as the
implications for practice. Finally, it acknowledges the limitations of this paper, and identifies
a few future research directions.

Related works on sports-related social media use
The literature on sports-related social media use has been growing along three major
themes. One theme focuses on players’ use of social media. For example, Sanderson (2008a)
examined how Curt Schilling, a pitcher with the US Major League Baseball team the Boston
Red Sox, used his blog. Schilling was found to use self-presentation strategies to maintain a
favorable social media presence. Kassing and Sanderson (2010) investigated tweets posted
by a selection of riders during a three-week race in Italy. The riders were found to use
Twitter to provide commentary and insider perspectives for fans. In a similar vein, Hull
(2014) analyzed how PGA Tour golfers tweeted during the week of the Masters tournament.
The tweets offered interesting glimpses into the golfers’ lives for fans.

A second theme focuses on the social media presence of sports organizations. For
example, Wallace et al. (2011) examined how the National Collegiate Athletic Association
(NCAA) – an organization dedicated to the well-being and success of student athletes – used
its Facebook Fan Page. A conscientious attempt to promote brand loyalty could be
identified. On a related note, Sanderson (2011) investigated social media policies provided in
handbooks for student athletes from a selected division of NCAA schools. More recently,
McCarthy et al. (2014) analyzed how UK football clubs manage their brand presence on
social media to foster engagement with their fans. Strategic social media use was identified
as being crucial to drive not only web traffic to official websites but also commercial gains.

Another theme, on which this paper specifically builds, focuses on sports fans’ use of social
media. For example, Sanderson (2013) examined how football fans at the University of
Cincinnati reacted on Facebook when their head football coach Brian Kelly left the school to
become the head coach at the University of Notre Dame. The use of abusive language to
derogate out-groups was particularly common. Babac and Podobnik (2016) analyzed
user-generated comments posted on Facebook pages of a selection of football clubs. Both
positive (e.g. joy) and negative (e.g. anger) sentiments made their presence felt. Frederick et al.
(2016) examined the use of Twitter related to the 2014 Sochi Winter Olympic Games. Besides
focusing on sports per se, tweets were often found to deal with broader political issues.
Within this theme, an emerging area of research specifically examines user-generated comments posted during live sporting events. For example, Yu and Wang (2015) examined sentiments expressed in tweets by US football fans during a selection of matches of the 2014 Football World Cup. Yang et al. (2016) investigated online discussions about the same tournament on the Chinese social media application Sina Weibo. At a granular level, some of the concepts brought to the fore include networking among fans (Jacobson, 2016), role of emotions (Sanderson, 2016), mediated interactions with players (Frederick et al., 2012) and the vitriolic nature of comments (Galily, 2008). Building on these works, this paper explores user-generated comments posted, while live matches were being played during the Cricket World Cup 2015.

Present study context and its uniqueness

Cricket
This paper uses cricket as the test case for investigation. It is a sport played between two teams. Each team comprises 11 players, who play three major roles – batting, bowling and fielding. While batting, batsmen use bats to hit balls and score points known as runs. While bowling, bowlers throw balls at batsmen to dismiss them. While fielding, fielders chase balls hit by batsmen to minimize the runs scored. One player from each team plays as a captain, who performs the toss. The winning captain decides which team would bat first. The team batting first has a single innings in which it tries to score as many runs as possible. The innings ends when bowlers from the other team have bowled 300 deliveries at the batsmen, or when ten batsmen have been dismissed, whichever occurs earlier. Thereafter, the teams change their roles. The other team gets an innings of 300 deliveries or 10 dismissals with which it tries to chase down the runs scored by the team that batted first. The team that scores the higher wins the match. Matches are officiated by individuals known as umpires. Each match lasts for about 8 h.

Cricket was chosen for two reasons. First, unlike sports such as football or hockey that mostly comprise continuous sessions of gameplay, cricket entails distinct events of a few seconds – when batsmen hit balls thrown at them by bowlers followed by fielders chasing the balls – interlaced with temporary pauses – when fielders return balls to bowlers who prepare to bowl again at batsmen. Such pauses amid actions in an 8-h-long match give users ample time and opportunity to post comments without necessarily missing much of the live action. This in turn suggests that live cricket matches are conducive to the phenomenon of second screening (Shim et al., 2017). Hence, they offer the possibility to collect a wide pool of online comments.

Second, the researcher has been an avid cricket follower for almost 20 years. This ensures that he has sufficient knowledge to interpret the comments. This was necessary because familiarity of the researcher with the research context is crucial for conducting qualitative research (Habibi et al., 2014).

On the scholarly front, research on cricket is conspicuously rare. Among the few works, some discussed the impact of cricket on societies (Agur, 2013; Sturm, 2015), while others investigated the relation between internal communication and commitment of cricketers (Mishra et al., 2016). This paper extends cricket-related research by analyzing comments posted on social media during live matches of the Cricket World Cup 2015.

Cricket World Cup 2015
The Cricket World Cup is the sport’s biggest tournament organized by the International Cricket Council (ICC), the governing body for the sport. This flagship mega-event is played every four years. It was first held in England in 1975.

Since its inception till 2011, the Cricket World Cup had been played a total of ten times. The tournament had been hosted four times by England, thrice by the Indian subcontinent, once
jointly by Australia and New Zealand, once jointly by South Africa, Kenya and Zimbabwe, as well as once by countries in the Caribbean. Australia had won the tournament four times, India and the West Indies had won twice each, while Pakistan and Sri Lanka won once each.

The 11th edition of the Cricket World Cup was played in 2015. It was jointly hosted by Australia and New Zealand. Specifically, the CEO of the hugely successful 2011 Rugby World Cup held in New Zealand was responsible to help the Trans-Tasman nations earn the hosting right (Newshub, 2016).

In the Cricket World Cup 2015, the top 14 cricket teams in the world played 49 matches across 44 days. The final match was played at Melbourne between the two host nations. Australia emerged as the winner (ICC, 2015).

Yahoo! Cricket
User-generated comments posted during live matches of the Cricket World Cup 2015 were collected from YC. For every match, the platform offers ball-by-ball commentary. Users have the option to post comments in response to the commentary corresponding to every ball. Figure 1 shows the screenshot of the commentary corresponding to the last ball of the tournament followed by an annotation indicating the number of comments received in response to the commentary till then (31), and a link to post additional comment (Leave a comment).

The option to leave comments after every ball makes YC a unique platform. Such a functionality is able to capture internet users’ instantaneous sentiments triggered by momentary events. Furthermore, YC is mostly visited by users from the Indian subcontinent. Its regional user base allows exploring how online conversations about cricket in the subcontinent intersect with broader issues such as culture, identity, politics and religion.

Methods
Data collection
Comments posted on YC in response to every ball bowled in each match of the Cricket World Cup 2015 were collected. They were archived in a single-spaced Microsoft Office word document that contained some 93,430 words spanning across 347 pages. It specifically comprised 19,155 comments.

There were eight matches that failed to receive any comments. In general, these were the ones that had a generally strong team playing against a relatively weak side, for example, England vs Scotland, and New Zealand vs Afghanistan. This shows that matches that are expected to be one-sided receive little traction on social media.

In contrast, the highest number of comments received by a single match was 4,557. It was the semi-final match played between Australia and India. Perhaps, YC was abuzz during matches with high stakes. However, it did not receive a lot of comments during the final match played between Australia and New Zealand. This confirms that YC is largely frequented by users from the Indian subcontinent. Indeed, it was found that matches involving the subcontinent nations generally resulted in a lot of online chatter. For example, the match between India and Bangladesh received 3,299 comments while that between India and Pakistan attracted 1,735 comments.

From the initial pool of comments collected, those that were not in English and those that contained nonsensical texts were eliminated. The cleansed document was 257 pages long.
and contained 68,745 words. In particular, it comprised 15,426 comments that were admitted for coding and analysis.

Coding and analysis

The data were analyzed using the grounded theory approach by the researcher, who was trained with qualitative research methods. Each comment was the unit of analysis. The coding task started soon after data collection, and continued for a period of almost nine months. To minimize fatigue-induced coding errors, the task was limited to at most 1 h of coding per session as well as to at most three sessions per day. Codes were revisited multiple times to ensure the consistency of the analysis.

Grounded theory requires a microanalysis of data involving a detailed line-by-line scrutiny to first generate initial codes and thereafter identify relations among the codes (Glaser and Strauss, 1967; Strauss and Corbin, 1998). Therefore, the data were first microanalyzed to generate emergent codes. Due to the interpretative nature of the analysis, the researcher allowed a single comment to be assigned multiple codes. After the initial categorization of codes (open coding activity), constant comparison was used by returning to the data in order to gain insights into the usefulness of the developed codes. They were compared to ascertain similarity and inter-relationships (axial coding activity).

The data were constantly compared and contrasted. This allowed for a fluid movement between data and the theory to be built. Development and enhancement of codes continued until new observations failed to add significantly to existing codes, thereby suggesting saturation. Table I gives a snapshot of the coding process. Finally, comments’ tone – ranging from polite to unpleasant – was chosen as the core category to explain the results. The overall approach of coding and analysis is informed by the literature (Glaser and Strauss, 1967; Papathanassis and Knolle, 2011; Sanderson, 2008b; Strauss and Corbin, 1998).

Evaluation

Since grounded theory requires drawing inferences grounded in data, there was neither any predefined coding scheme nor any quantifiable results. The explorative-qualitative nature of this paper rendered inter-coder reliability analysis inappropriate (Burla et al., 2008; Papathanassis and Knolle, 2011; Sanderson, 2008b).

<table>
<thead>
<tr>
<th>Open coding</th>
<th>Axial coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Providing status updates to other users</td>
<td>Providing status updates</td>
</tr>
<tr>
<td>Developing relations with other users</td>
<td>Developing relations</td>
</tr>
<tr>
<td>Interacting with players</td>
<td>Attempting to help</td>
</tr>
<tr>
<td>Forging connection with YC commentators</td>
<td>Offering constructive advice</td>
</tr>
<tr>
<td>Helping other users</td>
<td>Venting frustration</td>
</tr>
<tr>
<td>Helping YC commentators</td>
<td>Complaining</td>
</tr>
<tr>
<td>Advising other users</td>
<td>Abusing</td>
</tr>
<tr>
<td>Advising YC commentators</td>
<td>Mocking</td>
</tr>
<tr>
<td>Advising players</td>
<td></td>
</tr>
<tr>
<td>Advising team selectors</td>
<td></td>
</tr>
<tr>
<td>Sharing anger with other users</td>
<td></td>
</tr>
<tr>
<td>Engaging in verbal duels</td>
<td></td>
</tr>
<tr>
<td>Pointing out commentary errors</td>
<td></td>
</tr>
<tr>
<td>Blaming the YC application</td>
<td></td>
</tr>
<tr>
<td>Complaining to the ICC</td>
<td></td>
</tr>
<tr>
<td>Abusing other users</td>
<td></td>
</tr>
<tr>
<td>Taunting losing teams</td>
<td></td>
</tr>
<tr>
<td>Ridiculing YC commentators</td>
<td></td>
</tr>
</tbody>
</table>

Table I. Coding process
Nonetheless, in order to ensure the validity and reliability of the coding, the researcher contacted a coder, who had a Master’s degree in Computer Science and was interested in social media research. He had knowledge about cricket but was ignorant of this paper’s purpose. In a face-to-face meeting with the researcher, the coder was asked to independently think aloud about a randomly selected sample of some 200 comments. The thoughts of the independent coder generally concurred with the themes coded by the researcher. Validity and reliability was also guaranteed by later asking the coder to review and comment on the results of this paper. These approaches to ensure validity and reliability of results obtained using grounded theory are informed by the literature (Filieri, 2016).

Additionally, the rigor of qualitative analysis is often determined through auditability, credibility and applicability (De Witt and Ploeg, 2006). Auditability refers to the clarity and the correctness in the description of research methods. To establish auditability, the researcher provided as much information as possible about the research process in the Methods section of this paper. Credibility refers to the extent to which the findings can be confirmed by other works or by participants. Given the nature of data collection employed in this paper, the findings could not be confirmed by contacting users who submitted the comments. Nonetheless, the independent coder found the findings to be generally credible. Applicability refers to the extent to which the findings can be applied in other settings. This paper leaves it to future research to examine its applicability.

Results
To present the results, this paper treats comments’ tone as the core category because it could readily encapsulate all other codes emerging from the data. Some comments were polite in tone, and could be grouped into the following themes: providing status updates, developing relations, attempting to help and offering constructive advice. However, others tended to be unpleasant, and manifested along the following themes: venting frustration, complaining, abusing and mocking.

Providing status updates
Users often provided status updates to their online peers as if YC was a social networking site like Facebook. Some highlighted delight as in “Hip Hip Hurrayyyyyyyyy.” Others sighed, “My mood is off now.”

Some comments regretted missing out on the live action, “I got busy in a call and missed […]” Comments such as “this match is going to get over before I reach home” were perhaps written while commuting. This demonstrates how social media applications have revolutionized sports coverage by allowing users to stay abreast even on the move.

One user complained about how live streaming has been blocked in the individual’s workplace by stating, “The office blocking all live stream.” An ardent Indian supporter updated:

I have to cancel the 5-star dinner treat I was supposed to give to my colleagues if India went through to the finals […] I am sad as India lost, but more than that happy on account of saving a fortune of money.

Developing relations
Several comments suggested that users enjoyed forging social connections with netizens during live matches. For example, in response to the commentary corresponding to the first ball of the tournament, a user commented, “The first ball is dedicated to all cricket fans.” The initial stages of several matches attracted greetings such as “hi everyone,” and “hello all.” After matches ended, there were comments such as “bye all dear and near friends.” At nail-biting match situations, comments such as “control your nerves dear” and
“keep calm guys” intending to calm others down were also quite common. Some users appreciated others’ comments, “I love your comments [...]”

Users occasionally took the onus on them to engage the online community. For example, one user conducted a poll asking, “Who will be the man to score in today’s match? Fast comments please.” At times, they appeared eager to entertain others. For example, when rain interrupted a match, one user started chanting, “Rain rain go away.” They often served as statisticians providing information such as “last 23 years, every world cup final saw at least one Asian team.”

Users were also found to entertain the online community by playing with players’ names and attempting to be funny. For example, when New Zealand’s Trent Boult was playing well, a user commented, “plug […] this boult with 6 inch bolt.” Likewise, when Australia’s Aaron Finch was going great guns, a comment stated, “Finch, I will pinch you.” When England’s Ian Bell was in action, a user posted, “Bell has now ring.” These suggest that users did not lack a funny bone while watching live matches.

Users also attempted to develop relations with cricketers. There were several comments that seemed to have been written with the assumption that the cricketers could read the entries. For example, one user was so impressed with the performance of a player that the individual proposed, “Can I marry you?” Taking the cue, another user commented, “marry me […]” perhaps hoping that the player would pay attention to the second proposal too.

Users collectively referred to the players of the team they were supporting as brothers as evident in the following comment, “Congratulations my brothers.” When the legendary Sri Lankan Kumar Sangakkara, who had declared his decision to retire after the tournament, played his last match, there were plenty of comments of the form “Good bye Sanga,” “Salute to you and your personality and your contribution to the game” and “We will miss you Sanga.” One user asked, “Sangakkara, why are you retiring?” almost expecting him to respond. These demonstrate parasocial interaction – development of mediated relations (Frederick et al., 2012; Horton and Wohl, 1956) – between users on the internet and players on the field.

At the end of the tournament, the online community showed their appreciation toward the YC commentary team. This was evident from comments such as “Thanks Yahoo Team for your hard work” and “good boy yahoo. will miss-bye-takecare!” It was almost as if users were bidding adieu to their acquaintances.

**Attempting to help**

Attempts to help netizens were unmistakably evident. For example, a user apparently ignorant of the rules of cricket questioned about a commentary, “What is this: 5 Wides? How can it be 5 Wides in 1 Ball?” In response, another comment clarified, “It seems like you are new to cricket if u don’t understand 5Wides. Wide with boundary (i.e. 4) so total 5. It’s a wide so mentioned like 5 Wides.”

Users were also apologetic about their incompetent writing style. They regretted causing inconvenience to others for submitting difficult-to-read texts. This is evident from comments such as “Sorry for the poor grammar.”

Occasionally, users extended their help to the YC commentary team. When one user complained that YC had mistakenly not provided the commentary of a ball – “one ball missing,” some users were quick to respond with the missing details. Users often corrected some of the inadvertent mistakes made by the YC commentators. For example, in response to a complaint about a commentary error, one comment stated:

Proper commentary: OUT! Elliot falls! Fuller length ball outside off, Elliot slashes it hard over the point region. The ball goes towards Taskin in the deep who manages to hold on to that one. It was nicely played but he found the fielder unfortunately.
Offering constructive advice

Users did not shy away from offering constructive advice to all and sundry. For example, after a series of comments that vehemently rebuked players for their poor performance, one comment advised users to calm down:

To all the viewers a humble request: sports for peace. Yes, it is meant only for peace. Be sportive. Don’t look at any game as a war. Yes, I love sports […] I love the game […] Not any country or any player […] No scolding […] Have a nice day.

Users provided plenty of advice to the commentary team of YC. As evident from comments such as “Commentary in English should be free from cliché,” and “Flowery language should be used with utmost care,” users perhaps wanted a layman perspective in the ball-by-ball updates. Jargons and grandiloquent language failed to impress.

Users often behaved as self-appointed coaches. Advising players on the field, users posted comments such as “Bowl in line and length,” “Don’t take too much pressure […] play freely” and “Now dot ball [a ball on which no run is scored]? Hurry up.” Users almost seemed to believe that their comments would inform players, who would then adjust their strategies and tactics accordingly.

To cricketers, users also offered grooming tips. For example, some users expressed their dissatisfaction with the look of players with comments such as “please shave,” almost expecting the player to run from the field to the salon.

Users often commented as if they were team selectors. For example, one comment stated, “he should not be in the team now” while another stated “please bring back Yuvraj in place of Rahane.” It was as if users were expecting the selectors to listen to their comments.

One user went on to offer an advice buttressed with detailed explanations:

Drop Finch and instead take Watson in that position because as you can see, Watson is a good opener. Just look at his record. And he used to open the innings. The team will get one bowler also […] Watson is a great opener rather than a one-down batsman [one who comes out to bat after the first dismissal], Smith is in perfect position at one-down.

Venting frustration

Several users sounded angry, and vented out frustration through comments such as “what the hell,” “I hate this” and “It’s a frustrating performance.” Some of the comments were specifically directed to cricketers. For example, when a player struggled, one comment criticized, “Ashamed of you,” highlighting parasocial interaction.

Despite the apparent camaraderie among the online community, there were occasional traces of verbal duels. For example, during a group-stage match, one user apparently serving as a fortune-teller predicted, “First Semi-final: South Africa vs New Zealand, Second Semi-final: India vs Australia, Final: New Zealand vs Australia.” In response, the user was confronted by another self-appointed fortune-teller who responded, “Fool!! First Semi-final: South Africa vs New Zealand, Second Semi-final: India vs Pakistan, Final: South Africa vs India.”

Complaining

Some users were quite critical about the errors made by the YC commentary team. One comment complained, “Du Plessis was caught by S. Dhawan. Please correct the score card, it shows caught by Mohit Sharma!” Other similar comments include “Yahoo is showing wrong score” and “Please correct it. Not good.”

Some users even compared YC with CricBuzz – another website that provides live ball-by-ball commentary. Complaining about the slow speed of the commentary update, one user commented, “Hey Yahoo! Cricket, you have to be quick, CricBuzz is faster than you.”
On a related note, another user commented, “It’s better to use CricBuzz. It updates very instantly.” One user complained about the Yahoo application, “I hate this Yahoo app. It’s very slow.”

Users did not shy away from complaining to the ICC. Some comments contained complaints about the performance of umpires. Such comments include “poor umpiring,” “umpires should wear a pair of glasses to make decisions” and “the ICC should take a visibility test of umpires.”

Some users were quite vocal in complaining about the cricketing rules that result in the dominance of batsmen over bowlers. Such complaints came especially during the knock-out matches. While some simply expressed disgust through comments such as “shame ICC,” others offered a more detailed treatment as follows:

Worst world cup!!! It’s getting worse after each match! The rules do not care about bowlers. I don’t know what these ICC fellows are doing […] Let’s see how many budding cricketers want to be a bowler in future. Batsmen scoring 200+ runs and teams scoring 400+ runs regularly. Why do we need bowlers in this game now? Place the bowling machines instead!!!!

Abusing
There were instances of cyberbullying through the comments. For example, one comment stated, “You all are f****D.” The use of the f-word suggests that users do not always follow the guidelines provided by YC – “Be considerate and respectful of your fellow posters” and “Don’t post profanity, obscenities, abusive language, or otherwise objectionable content.”

During a match between arch rivals India and Pakistan, a religiously sensitive comment stated, “It’s pretty sad to see Indian Muslims support Pakistan even though they know how badly they will be treated if they visit there!” These abusive comments were mostly posted during crunch match situations.

Nonetheless, there were some sensible users who later apologized to the community after writing abusive comments. Such an apologetic comment stated, “Sorry, by mistake […] I typed.” It seems that users can be easily caught up in jingoism during live matches.

Mocking
Users were at their sarcastic best when teams were at the verge of being knocked out of the tournament. They often taunted managers of the losing teams, “book the return flight today, the players are too homesick.” Some acerbically wished, “Have a safe flight back home.” Others even ironically quipped, “When you come, please bring some chocolates from there.” In addition, users were found to ridicule the YC commentary team for their errors with sarcastic remarks such as “hahahaaaaaaaaaaa.”

Discussion and conclusions

Findings
With information technology making rapid inroads into sports broadcast, netizens now have the privilege to post comments on social media during live sporting events. Hence, taking the case of the Cricket World Cup 2015, this paper analyzed user-generated comments posted on YC during live matches. Grounded theory was used for analysis. The results show that users posted both polite and unpleasant comments.

Five major findings could be gleaned from the results. First, netizens enjoy consuming live sporting events in an online social setting rather than as isolated individuals. As evident from numerous comments, users almost felt like watching the matches together with their online peers. This finding lends empirical support to Jacobson’s (2016) idea of networked spectatorship and Shim et al.’s (2017) concept of social co-viewing in the context of second screening. Users were keen to forge social ties with like-minded individuals passionate
about a common sport. This concurs with previous works that highlighted the role of online comments in fostering interactions, developing relations and constructing identities (Galily, 2008; Tenenboim and Cohen, 2015).

In addition, drawing from the uses-and-gratifications perspective, comments were forthcoming possibly due to users’ psychological needs for information exchange and emotional outlet (Church and Smyth, 2009; Clavio and Kian, 2010; Hsu et al., 2015; Lee, 2012). Users seemed to be in need of the information provided in the ball-by-ball commentary coupled with others’ reactions on them to stay abreast with the progress of the live matches in the online social setting. Concurrently, the plethora of comments that they posted was a testament to their penchant to share information. Their willingness to share remained active even when they were consuming live sporting events. This suggests that posting comments could hardly distract users during the matches. In fact, second screening provided users an outlet for euphoria and despair that in turn perhaps made following the matches a more pleasurable experience than otherwise (Babac and Podobnik, 2016).

The next finding is that user-generated comments posted during live sporting events are the by-products of users’ emotions. Previous works on sports-related user-generated content found both positive and negative emotions to play a significant role in influencing fans’ experience (Babac and Podobnik, 2016; Sanderson, 2016; Yu and Wang, 2015). After all, sports are about the joy of winning and the disappointment of losing. Given that users usually follow sports wanting a team to win, they experience a variety of emotions. Moreover, the psychology literature has long suggested that emotions affect social behavior (van Kleef et al., 2004). Even in the online context, emotions influence individuals’ response to situations (Chang et al., 2014). Therefore, it seems that the comments that were submitted on YC during live matches were triggered by the instantaneous emotions that users experienced during the matches.

In this vein, this paper also takes a step to respond to the call for research to study emotional shifts over the course of matches advocated by Sanderson (2016). It finds that the tone of the comments varied depending upon the stage of the matches. At the initial and the intermediate stages, most comments were friendly and cordial (e.g. “The first ball is dedicated to all cricket fans,” “hi everyone,” “welcome dear”). Specifically, the comments at the initial stages were interpersonal in nature, while those at the intermediate stages were mostly focused on match situations. At the final stages however – specifically at crunch situations, the comments started to take abusive and sarcastic shapes (e.g. “Australia we Asians hate you,” “dear Pakistani guys, be careful you have enemies everywhere,” “I hate India”).

It seems that some users calmly accept the reality about the teams that they support. Hence, they are able to maintain decorum in writing comments. In contrast, others perhaps do not hesitate to unleash their pent-up emotions to surface up to their comments, thereby occasionally making the entries unpleasant. This echoes the previous finding that sports-related online communication can occasionally lose its social polish, and become abusive whereby in-groups leave no stones unturned to belittle their out-groups (Sanderson, 2013).

The third finding is that user-generated comments posted during live sporting events foster parasocial interaction – development of one-sided bonds of intimacy through mediated interactions with those who are viewed through media (Horton and Wohl, 1956). Consistent with the literature, users were found to engage actively rather than passively in their friend-like relations with media personas – in this case, players (Frederick et al., 2012). Most sport viewers tend to have allegiances with players and teams. In consequence, parasocial interaction is expected to be conspicuous between users and their favorite players. Thus, fandom as an antecedent of parasocial interaction is conceivable.

In addition, this paper identifies parasocial interaction between users and their disliked players. This is evident from comments such as, “Please come to my home, bro. U’re so
handsome but good-for-nothing on cricket [...]” A somewhat similar finding was highlighted in Sanderson (2008b), which deemed criticism to be an aspect of parasocial interaction. Overall, users who post comments during live sporting events appear as individuals wearing their hearts on their sleeves. They are consistently boisterous in praising as well as criticizing players. In consequence, they are equally likely to develop parasocial relations with players who perform well, and those who fare poorly. In contrast, parasocial interaction appeared to be limited between users and players who were average in performance.

The fourth finding is that users can easily get into confrontation while posting comments during live sporting events. Abusive comments were particularly common during matches involving teams that have had a history of rivalry (e.g. India vs Pakistan, and Australia vs New Zealand). This was partly expected as previous works have shown sports-related online comments to touch on broader political issues beyond sports (Frederick et al., 2016). Nonetheless, the current finding adds to the literature by identifying that abusive comments were particularly forthcoming during nail-biting moments. Perhaps, nerve-racking match situations turned an ordinary sporting event into almost a life-and-death encounter in the minds of users. Viewing the situations through a gladiatorial lens, users started to abuse one another ignoring the guidelines provided by YC.

While works such as Galily (2008) found that users tend to post occasional violent comments on sports-related online articles, this paper shows that such a tendency is perhaps magnified manifold during live sporting events. As tempers flare, the abusive comments that users hurl at one another engender an atmosphere of jingoism and xenophobia. The internet’s advantage in terms of freedom of expression is unfortunately turned into freedom of incitement (Porat, 2011).

The final finding is that a number of themes were conspicuous by their absence in the comments. For one, there were no comments about the commentary on air even though there were comments related to the ball-by-ball commentary of YC. In addition, there was no discussion about the context of the tournament or the politics surrounding the event that often feature in mass media during mega-events. Perhaps, netizens remain uninterested in topics that are expected to be covered in traditional media.

Contributions to theory
By studying the case of user-generated comments posted on YC during live sporting events, this paper contributes to research on media entertainment and second screening. The extant literature abounds in evidences to suggest that individuals rely on media for entertainment purposes (Vorderer et al., 2004). Prior research also demonstrates how social media applications serve as avenues for entertainment (Dogruel et al., 2015). This paper dovetails these lines of research by demonstrating how social media applications such as YC are used for entertainment during live sporting events in order to meet social, informational as well as emotional needs.

In addition, this paper gleams several new findings that have not been exhumed in related works. For example, it identifies cyberbullying to be a severe problem in social media communication during live sporting events. At the same time, it finds that social media communication during live sporting events can be a tool to build camaraderie among internet users. Besides, this paper finds that internet users have the power to partly dictate the nature of commentary on websites such as YC in real time. Moreover, this paper extends the literature on parasocial interaction in sports by showing that it is common between users and players who give noteworthy performances – either brilliant or abysmal – yet inconspicuous between users and players who give average performance. These new findings can pique further theoretical debates among scholars about sports-related social media use.

Interestingly, some of the a posteriori findings of this paper bear resemblance with extant theories. For example, the identified gratification factors that seem to encourage users to post
comments during live sporting events – social, informational and emotional needs – are possible to explain based on the theoretical perspective of uses-and-gratifications (Clavio and Kian, 2010). Again, the ways in which users directed their comments to the players on the field are possible to explain through the theoretical perspective of parasocial interaction. These demonstrate how the findings uncovered in this paper using grounded theory can be lifted to a higher level of theoretical abstraction (Suddaby, 2006).

**Implications for practice**

This paper recommends governing bodies and marketers of sports to keep a pulse on the opinions echoed by internet users on social media during live sporting events. This is important because some users who post comments in real time tend to have good sense of the sport. Listening to their voices could help improve fans’ overall experience. For example, as indicated earlier, one comment complained to the ICC about the future of bowlers and the supremacy of batsmen in cricket during the 2015 world cup. However, the issue was officially taken up by the ICC only in the third quarter of 2016 to bring about a change in the rules (Devgan, 2016). Had user-generated comments been monitored closely by the ICC during the Cricket World Cup, such decision making might have been fast-tracked. This in turn would have resulted in better utilization of social media as a communication channel between common sports fans and sports administrators.

In addition, this paper offers two implications beyond the setting of sports. First, it demonstrates that users do not necessarily comply with the guidelines prescribed by social media applications while posting comments. In this case, even though YC requests users to be considerate while submitting comments, users were often found using abusive language. Perhaps, users did not even look through the guidelines before submitting their entries. This calls for social media applications to be equipped with automated language detection systems that will censor comments that contain abusive language. After all, while preserving the plurality of opinions on social media is essential, silencing the voices that make the online environment hostile is equally imperative.

Second, this paper highlights the potential of user-generated comments to build camaraderie among users as well as to engage the individuals in dynamic arms race. On the one hand, goodwill among users was conspicuous in numerous comments. On the other hand, anti-national and religiously sensitive comments suggested how the online community was split along the fault lines of nationality and religion. This points to how social ties are easy to be forged and destroyed on the internet (Frangonikolopoulos and Chapsos, 2012). In this vein, information and communication practitioners could explore strategies so that the ways to build camaraderie are promoted, and the fault lines diminished. Such attempts could help develop a healthier cyber-culture among the online community.

**Limitations and future research directions**

The paper is constrained by its data source coupled with the limited data collection window. Data were collected from a singular platform that is frequented by a regional user base during a single sporting event. This limits the generalizability of the findings. Drawing data from several social media platforms such as Twitter and Facebook during multiple temporally dispersed events would have facilitated a richer treatment of the research issue at stake. Moreover, non-English comments could not be coded and analyzed. Such comments might have further enriched the paper. Hence, the findings of the paper should be viewed in light of these limitations.

Nonetheless, as social media applications revolutionize sports coverage, exciting research avenues continue to unfold. In particular, this paper identifies two directions for future research. One direction involves identifying users’ motivation to post comments during live sporting events. For this purpose, in-depth interviews could be conducted.
Such an investigation would help determine why some users share while others remain lurkers. Additionally, it could also help determine ways to promote camaraderie among users, and prevent them from abusing one another during live sporting events.

A second research direction could seek to exhume the reasons for parasocial interaction during live sporting events between users and players. As evident from this paper, users were always willing to offer advices and wishes to players who were on the playing field. They almost expected the players to revert back despite knowing its apparent implausibility. By exploring psychological and cognitive perceptions of users during live sporting events, the scholarly understanding of parasocial interaction mediated by the internet could be further enriched.

References


**Corresponding author**
Snehasish Banerjee can be contacted at: snehasish_banerjee@ymail.com

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
eBook Collections

Research at your fingertips

Bringing together over 1,600 eBooks, the Emerald eBook collections are a cost-effective way of instantly expanding library holdings and increasing usage through an award winning platform alongside journals.

Visit emeraldinsight.com

Get eBook Collections for your library

Recommend eBooks to your librarian today. Find out more:
emeraldpublishing.com/ebookcollections
1009 Editorial advisory board

1010 Timely assessment of disaster and emergency response networks in the aftermath of superstorm Sandy, 2012
Jungwon Yeo, Louise Comfort and Kyujin Jung

1024 Predicting the quality of health web documents using their characteristics
Melinda Oroszlányová, Carla Teixeira Lopes, Sérgio Nunes and Cristina Ribeiro

1048 A study of antecedents influencing eWOM for online lecture website: personal interactivity as moderator
Mi sook Lee and Hongbok An

1065 To get cited or get tweeted: a study of psychological academic articles
Yingxin Estella Ye and Jin-Cheon Na

1082 Mediating impact of fan-page engagement on social media connectedness and followers purchase intention
Zoha Rahman, Sedigheh Moghavvemmi, Kumaran Suberamanaian, Hasmah Zanuddin and Hairul Nizam Bin Md Nasir

1106 Challenges for an SNS-based public sphere in 2016
Kyle John Lorenzano, Miles Sari, Colin Harrell Storm, Samuel Rhodes and Porsmota Borah

1124 The perception of e-servicescape and its influence on perceived e-shopping value and customer loyalty
Ana Cuic Tankovic and Dragan Benazic

1146 Social media and co-creative service innovation: an empirical study
Bijoylaxmi Sarmah, Shampy Kamboj and Jay Kandampully

1180 Analysis of user-generated comments posted during live matches of the Cricket World Cup 2015
Snehasish Banerjee