Unlocking value through an extended social media analytics framework

Insights for new product adoption

Gavin Jiayun Wu
Broadwell College of Business and Economics, Fayetteville State University, Fayetteville, North Carolina, USA

Zhenning “Jimmy” Xu
Department of Marketing, University of Southern Maine, Portland, Maine, USA

Saeed Tajdini
College of Business, Indiana University Southeast, New Albany, Indiana, USA

Jie Zhang
Department of Information Systems and Operations Management, The University of Texas at Arlington, Arlington, Texas, USA, and

Lei Song
School of Business, Stockton University, Galloway, New Jersey, USA

Abstract

Purpose – To unlock social media’s value, this study aims to integrate insights from several theoretical perspectives and the relevant literature, developing an extended social media analytics framework. It identifies the stages underlying the social media analytics process and tests the framework in three important and interconnected areas: social media (Twitter), new product adoption (iWatch and Google Glass) and social media analytic techniques (text mining and sentiment analysis).

Design/methodology/approach – Based upon a systematic review of different research approaches, theories and media types, this paper presents and tests an extended framework in three important and interconnected areas mentioned above.

Findings – This paper offers a theory-driven social media analytics framework. It validates the framework by providing concrete processes, examples, evidence and insights related to three chosen areas mentioned above, thereby helping managers create effective and efficient social media and new product development strategies.

Originality/value – This paper integrates insights from theories of the middle range (Merton, 1949), Campbell’s (1965) model of sociocultural evolution and Fan and Gordon’s (2014) social media analytics framework, developing its own extended social media analytics framework and validating it in three important and interconnected areas mentioned above. This paper demonstrates not only how the proposed framework can be applied to the context of new product development, but also how social media are transforming research approaches (qualitative, quantitative and mixed method) and the very nature of business itself (increased importance of digital business).

Keywords New product development, Sentiment analysis, Twitter, Text mining, Social media analytics, iWatch

Paper type Research paper
Introduction
A recent revolution in business brought by social media, technologies and the sharing economy, has drawn scholars’ and practitioners’ attention to the impact of earned media (user-generated word-of-mouth, social media) (Belk, 2014; De Vries et al., 2012; Hennig-Thurau et al., 2015). One reason for increased attention to earned media is that of the three types of media generally classified by marketers – paid (advertising), owned (firm website) and earned (user-generated word-of-mouth, social media) (Stephen and Galak, 2012) – only the effects of the first two have been extensively studied in the marketing literature in regard to firm performance (sales) (Belk, 2014; Hennig-Thurau et al., 2015), but the effect of earned media (e.g. Twitter data) has gone largely unexplored. While the power and impact of offline word-of-mouth should not be underestimated, one should be increasingly aware that Twitter’s and other social media platforms’ online word-of-mouth is reshaping the brand communication and even the very nature of business itself (e.g. digital marketing and business). Recent studies show that user-generated content is critical for enhancing new product adoption. For example, Hennig-Thurau et al. (2015) studied how microblogging, such as Twitter postings, can influence new movie adoption. They found that, the sooner a company starts gathering and analyzing such social media content and connections, the more it benefits.

Despite the many opportunities social media present to the business community, they nonetheless pose challenges as well. Currently, one of the primary challenges for businesses is the need for a well-developed systematic social media analytics framework to guide actionable strategies. Even though firms have been investing heavily in engaging consumers through social media, research has demonstrated that firms have difficulty gathering, analyzing and integrating social media data into their practices (Fowler and Pitta, 2013). Moreover, little insight has been offered by academia to guide strategic social media marketing and new product innovation. The lack of research highlights the imperative need for social media analytics, which refers to “developing and evaluating informatics tools and frameworks to collect, monitor, analyze, summarize, and visualize social media data, usually driven by specific requirements from a target application” (Zeng et al., 2010, p. 14). Because of the unstructured nature of social media data (texts, pictures, videos, etc.) (Kurniawati et al., 2013), they differ from traditional data. However, they are an important source for businesses seeking to gain and maintain a competitive edge (Davenport and Harris, 2007).

To address the gaps identified above, we first integrate insights from several theoretical perspectives and the relevant literature. We then propose and test an extended social media analytics framework (see Figure 1). By addressing these gaps, this study makes the following contributions. First, it contributes to the literature by integrating insights from theories of the middle range (Merton, 1949), Campbell’s (1965) model of sociocultural evolution and Fan and Gordon’s (2014) social media analytics framework, offering its own extended social media analytics framework (see Figure 1). Its extended framework delineates the underlying stages of the social media analytics process. Second, by testing the proposed framework in three important and interconnected areas – social media (Twitter), new product adoption (iWatch and Google Glass) and social media analytics techniques (text mining and sentiment analysis) – this study demonstrates not only how our framework can help address the questions at hand (e.g. new product development and adoption), but also how social media are changing and transforming research approaches (qualitative, quantitative and mixed method) and the very nature of business itself (e.g. increased importance of digital business). For example, qualitative researchers can adopt techniques such as sentiment analysis and word association analysis to better understand constructs under investigation and to advance theory development. Third, this research validates the
Figure 1.
Extended social media analytics framework
newly proposed framework by providing concrete processes, procedures, examples, evidence and insights related to chosen areas of interest, thereby helping managers create effective and efficient social media campaigns and new product development strategies.

Extended social media analytics framework
By integrating insights from several theoretical perspectives and the relevant literature, this research extends and improves Fan and Gordon’s (2014) three-stage-process framework (i.e. capture, understand and present) by making it a five-stage-process one (i.e. formulate, determine, capture, understand and present). Fan and Gordon’s (2014) framework is based on broad input-process-output models consisting of only three stages. Therefore, their approach and framework might be most appropriate for exploratory research or studies with loosely or broadly defined research objectives and scopes. However, because Fan and Gordon’s (2014) approach is based on broad input-process-output models and their framework starts with the capture stage, their framework may appear to be purely data-driven or exploratory, thereby potentially preventing their framework from being appropriately understood and applied. To the contrary, starting with the formulate stage, our proposed framework helps us to embrace the nature of most social media data (e.g. unstructured, high velocity, veracity and volume), to articulate typical social media research’s differing objectives (to extract insightful patterns from the majority and the minority outliers; to facilitate conversations and interaction; and to monitor processes) and more importantly, to recognize one of the primary advantages of a study with social media data: to identify an opportunity (e.g. to extract insightful patterns and intelligence), which is a vital requirement for innovation. Therefore, in some cases, the seemingly “data-driven”, exploratory or descriptive social media studies are actually well justified and, in fact, appropriate.

In addition, Fan and Gordon’s (2014) approach and framework might not be sufficiently elaborate to study complex modern business issues. To improve measurement in practice, Sir Ronald Fisher implores, “[M]ake your theories elaborate” (Cochran, 1965, p. 5). Applying Fisher’s dictum, we believe that Fan and Gordon’s three-stage framework is an excellent starting point (most appropriate for exploratory research or studies with loosely or broadly defined research objectives and scopes) that nonetheless requires further elaboration. Therefore, we extend their framework to enable social media’s emergent business fields to grow. For instance, as discussed in detail in the sections relating to the formulate and determine stages below, our more elaborate approaches and framework help us identify another condition (i.e. to identify an opportunity) that justifies adding the formulate stage in a social media analytics framework. These approaches and this framework also explain why the social media research process makes convergent parallel mixed methods more feasible, as well as why and how the traditionally clear differences between the primary and secondary data may become less clear when referring to social media data.

Moreover, our chosen research approaches and contexts align with the developmental orientation of the theories of the middle range (Merton, 1949). Theories of the middle range, as the rubric implies, apply to various aspects of social phenomena (e.g. the unanticipated consequences of goal-directed behavior, reference groups and social control) that are not so removed “from particular classes of social behavior, organization, and change [that these aspects cannot] account for what is observed” (Merton, 1949, p. 39). At the same time, neither are these aspects such “detailed orderly descriptions of particulars that [they cannot be] generalized at all” (Merton, 1949, p. 39). Ever mindful of Merton’s (1949) developmental orientation, this research does not attempt to abruptly develop an entirely new encompassing theory to explain issues associated with social media. Instead, it integrates
insights from three theoretical perspectives – theories of the middle range (Merton, 1949), Campbell’s (1965) model of sociocultural evolution and Fan and Gordon’s (2014) social media analytics framework – and focuses on three important, interconnected areas (social media, new product adoption and social media analytics techniques). This research offers a reasonably representative assessment of social media and new product adoption as its research setting and research topic, respectively, including ways in which they actually interact and proceed.

Specifically, by identifying and integrating Campbell’s (1965) model of sociocultural evolution and Fan and Gordon’s (2014) social media analytics framework, we propose and test an extended social media analytics framework (see Figure 1). The reasons for incorporating this sociocultural theoretical perspective (Campbell’s 1965 model) are as follows. Applying Bagozzi’s (1992) view and logic regarding the consumer research field, we consider the social media research field to be a social system. It therefore adheres to essential “can-and-should-be-understood” principles, such as “[r]ather than haphazard or random growth...[its] process is guided by natural selection at the social level where negative feedback functions as a control mechanism” (Bagozzi, 1992, p. 355), bringing about adaptation and evolution. For example, legitimate inference derived from online imprints is only one aspect of human behavior (Tufekci, 2014). Some of what users do to add a layer of invisibility to their online behavior (e.g. mock re-tweeting, sub-tweeting, use of “screen captures” for text, etc.) makes interpretation of social media big data quite difficult. Another difficulty comes from field effects – universally influential events that affect not only the network under study but also the entire society (Tufekci, 2014).

We discuss each stage’s framework in detail below which focuses on the theoretical and conceptual perspectives, as well as each stage’s implementation strategies which emphasize their operational and empirical aspects. Additional justifications for adding the formulate and determine stages to Fan and Gordon’s (2014) framework follow.

Stage 1: formulate

Framework

This research includes formulate as the first stage in our extended social media analytics framework for several reasons. First, Campbell’s (1965) model (see Figure 2) justifies including not only the formulate stage but also the determine and present stages. Applying Bagozzi’s (1992) logic, Campbell’s (1965) model demonstrates that three basic mechanisms (variation, selection and retention; see Figure 2) comprise the socio-evolutionary process, “each connected through two throughput and two feedback processes” (see Figure 2; Bagozzi, 1992, p. 355). Variation (change) arises in the form of new information or knowledge, such as social media data. Therefore, variation in Campbell’s (1965) model is analogous to the “input” in the input-process-output models while corresponding more closely to the capture stage in Fan and Gordon’s (2014) framework. Because variation initiates growth in a social sense (Bagozzi, 1992), greater variation is expected to enhance the impact and importance of social media research; therefore, “the more numerous and the greater the heterogeneity of variations, the richer the opportunities for an advantageous

![Figure 2. Outline of Campbell's (1965) model of sociocultural evolution](attachment:image.png)
innovation” (Campbell, 1965, p. 28). This observation further supports choosing social media as our research setting and new product adoption as our research topic.

As variations like those occurring in social media data can be potentially overwhelming, selection’s mechanism serves as a social system’s coping response in some sense and “brings order out of chaos” (Bagozzi, 1992, p. 356). Accordingly, selection in Campbell’s (1965) model is analogous to the “process” in the input-process-output models while corresponding more closely to the understand stage in Fan and Gordon’s (2014) framework. However, according to our guiding research perspectives mentioned above (e.g. theories of the middle range), we extend Fan and Gordon’s (2014)’s capture and understand stages to the formulate, determine, capture and understand stages, to be more elaborate.

Second, this research makes further contributions by identifying another condition that justifies adding the formulate stage in a social media analytics framework. The formulate stage serves two purposes: to solve a problem and to identify an opportunity (e.g. insightful patterns). The main purposes of a large body of research (e.g. descriptive and causal research) are to describe patterns and to solve problems. Nevertheless, because typical social media research has different objectives (to extract insightful patterns from the majority and the minority outliers; to facilitate conversations and interaction; and to monitor processes), and due to the nature of most social media data (e.g. unstructured, high velocity, veracity and volume), one of the primary advantages of a study with social media data is that it can identify an opportunity (e.g. to extract insightful patterns and intelligence from both the majority and the minority outliers), as opposed to merely solving a problem, which is an important requirement for innovation. For example, Zeng et al. (2010, p. 14) argue that social media analytics research’s purposes include “extracting useful patterns and intelligence to serve entities that include, but are not limited to, active contributors in ongoing dialogues”. Some of these purposes align with those of the qualitative research approach (e.g. general patterns or key issues that govern human behavior) (Nestor and Schutt, 2014). Therefore, in some cases, the seemingly “data-driven”, exploratory or descriptive social media studies are well justified and, in fact, appropriate. To the contrary, in addition to the aforementioned troublesome issues associated with Fan and Gordon’s (2014) framework, starting with the capture stage may nonetheless lead to justifiable criticism or misleading results.

Implementation
The present study’s formulate stage relates to three areas: social media (Twitter), new product adoption (iWatch and Google Glass) and social media analytic techniques (text mining and sentiment analysis). A synthesized literature review relating to these three areas, briefly described and justified in the introduction, follows.

Although previous studies have examined the transition from one-way communication to two-way conversation (Kumar and Mirchandani, 2012), studies on deriving new product insights from two-way conversations on social media are few, despite practitioners’ call for more research in this area (Nielsen, 2012). Typical advertising as a type of paid media is often one-way communication, marketer-driven, celebrity-endorsed and mass-distributed. Due to the proliferation of social media, a two-way conversation provides a unique opportunity for firms to understand their customers in real time, to modify existing products and even to design new products that serve customers better than their competitors’ products (Xu et al., 2016). From the perspective of relationship marketing, the individually driven, idiosyncratically derived and exclusively distributed information acquired from social media, facilitates effective dialogue. In terms of media richness and strength, social media possess rich functions and abilities (e.g. use of multimedia information), enabling consumers to easily share content with friends and even with strangers across media.
platforms to create social “buzz”. Their conversations increasingly diffuse new product information through both strongly and loosely connected social media members, thereby enriching information even further (Xu et al., 2016). Most importantly, insights from social media may prove valuable for new product managers because consumer opinion, as well as trending topics, can reflect likes and dislikes, potentially determining whether or not particular innovative product features are desired by the market. Therefore, ignoring social media and their underlying sentiments might affect firms’ marketing management and new product development.

Nevertheless, most companies still deal with social media via traditional, one-way communication. They push information to consumers rather than extract information and insights from them (Xu et al., 2016). One possible explanation for the continued use of this traditional approach is that firms either have not recognized social media’s potential effects on their brand and product success or they lack appropriate tools to analyze social sentiments. Not surprisingly, utilizing social media and big data analytics is highly rewarding (Kumar and Mirchandani, 2012).

Although traditional marketing analytics enjoy certain advantages, they face increasing challenges, primarily due to the ever-changing marketplace (e.g. social media and big data) and emerging analytical tools and techniques (e.g. text mining and sentiment analysis). For instance, traditional marketing analytics usually rely on consumer surveys and experiments to gather and analyze information relating to customers (Xu et al., 2016). However, managers are increasingly required to more quickly understand customer opinions and perceptions about products to gain and maintain competitive advantages.

To overcome the limitations of traditional marketing analytics, one promising solution is to apply social media analytic techniques (e.g. text mining and sentiment analysis) to user-generated content. Using information retrieval and natural language processing techniques, business professionals can efficiently retrieve unstructured data from user-generated content, transform them and then analyze the relationships among factors identified from the classified data (Jockers, 2014). Twitter, for example, provides a virtual world where people from almost anywhere can freely communicate and share their ideas with others. Aside from user-generated sentiments, tweets embedded on sites other than Twitter grab approximately 185 billion views per quarter (Koh, 2014).

Stage 2: determine
Framework
Similar to the traditional research process, the purpose of the determine stage is to specifically determine the research design and data collection methods and forms (e.g. primary vs secondary data, structured vs unstructured, questionnaire vs observational data collection form). However, due to the different objectives of research typically associated with social media (e.g. extracting insightful patterns and facilitating conversation and interaction) and the nature of social media data, differences exist between a traditional research framework’s determine stage and our new framework’s determine stage. These differences justify the reason for adding the determine stage to our social media analytics framework in particular and for proposing an extended social media analytics framework in general.

For example, among the three types of research (exploratory, descriptive and causal), only the first two types (exploratory and descriptive) are mainly suitable for empirical social media research studies. With regard to the research design in the traditional research process, we can use qualitative, quantitative and mixed-method approaches (e.g. convergent parallel mixed methods, exploratory sequential mixed methods, explanatory sequential
mixed methods); however, only the last two approaches (quantitative and mixed method) seem to be commonly used for research using social media data. This observation is supported by Branthwaite and Patterson (2011), who argue that by applying the three critical features identified below, qualitative research can differentiate itself from social media monitoring. These critical features are: the interactive, direct conversation or dialogue between researchers and consumers; the ease or ability to “hear” and attend to the underlying (sometimes unspoken) narrative, which links consumers’ aspirations and needs, driving forces and personal goals toward brand choice and behavior; and the interactive, dynamic qualities and characteristics of the interview that result in a “meeting of the minds” to reach a mutually shared understanding. This particular dialogue or “conversation”, philosophically speaking, lends qualitative research its authenticity and validity, making it superior to social media monitoring.

Moreover, as implied, the social media research process makes convergent parallel mixed methods more feasible. Using convergent parallel mixed methods differs from the traditional research process in which the sequence – such as ones in explanatory sequential mixed methods and exploratory sequential mixed methods, as well as ones ranging from determining research design to determining data collection method and forms – is more explicit. Also, because of high-velocity information generated and disseminated through social media platforms, the traditionally clear differences between the primary and secondary data may become less clear when referring to social media data. As such, compared to the traditional research process, the social media research process may appear to researchers to focus less on the secondary data while paying more attention to the primary data, leading them to erroneously criticize the process as being data-driven.

Primary data may be an essential element in social media research, especially when the purpose of a study is to identify an opportunity (insightful patterns). However, this does not mean that social media research is bereft of a theoretical framework or happens outside a pre-contemplated research design. In fact, a theoretical framework and research design guide primary data collection, sampling, data cleaning, the chosen method of analysis and many other operational considerations (Churchill and Iacobucci, 2010).

Implementation
To account for the research gaps and to explore and identify social media knowledge patterns, this study searches, collects and analyzes several samples of Twitter data. This allows us to determine if we can capture essential insights, such as consumer opinions, which are particularly important for new product development.

This study chooses several specific new technology products (e.g. iWatch and Google Glass) to analyze for the following reasons. First, most consumers still perceive iWatch and Google Glass as relatively new. What consumers say about such relatively new technology products is critical if companies are to better optimize their new product development strategies. Toward this end, the study specifically focuses on individuals’ tweets relating to the new technology products to identify and interpret insights. Second, high-profile products, such as iWatch and Google Glass, have engendered significant word-of-mouth and therefore are appropriate candidates for this study’s context and purposes. Third, to offer new insights on competitive analysis and real-time knowledge extraction by examining the introduction of a critical discontinuous technology (e.g. iWatch) and its impact on different brands and products (e.g. Samsung, and iPhone), we incorporate Samsung and iPhone into our study as well.
Stage 3: capture
Framework
The capture stage permits a business conducting social media analytics to find out the particular social media conversations relating to its interests and activities (Fan and Gordon, 2014). This stage involves collecting large amounts of relevant data that span hundreds, if not thousands, of social media data points by exploiting application programming interfaces and news feeds and by crawling the social media platforms (Fan and Gordon, 2014). The capture stage covers popular platforms (e.g. YouTube, Twitter, Facebook, LinkedIn, etc.), as well as more specialized and niche sources (e.g. blogs, microblogs, news sites and picture-sharing sites) (Fan and Gordon, 2014). A variety of pre-processing steps may be carried out in the capture stage when preparing for a data set.

Implementation
The capture stage involves two tasks: downloading and pre-processing (e.g. filtering out irrelevant tweets). The official release date for iWatch was April 24, 2015, and this study captured and analyzed tweets before and after this date. In line with the purpose and emphasis of this study, focusing on tweets published before and after the product release sheds light on consumer attitude or behavior changes about the product, which prove critical for developing and preliminarily testing the framework.

This study uses keywords to search for relevant tweet messages via the Twitter application programming interface (API). APIs are designed to facilitate interactions with websites. Due to the current Twitter API constraints, fewer than 1,500 random tweets can be accessed at a time. To overcome this constraint, this study carries out the data retrieval process multiple times, resulting in a richer sample with a broader spectrum.

After retrieving text from Twitter using R’s TwitteR package, this study pre-processes the tweets’ text to prepare for the formal text analysis. Examples include converting tweets to a data frame; converting all letters to lower case; removing punctuations, numbers and stop words; as well as stemming and identifying synonyms (Francis and Flynn, 2010).

Stage 4: understand
Framework
The understand stage occurs after a business has gathered sufficient social media conversations relating to its operations and products. At this stage, the business evaluates the conversations’ meaning and generates metrics to make useful decisions (Fan and Gordon, 2014). Given that the capture stage gleans data from many sources and applications, a sizeable portion of the data may need to be removed before any meaningful analysis can be conducted. This particular cleaning function may use either sophisticated classifiers or simple, rule-based text classifiers (Fan and Gordon, 2014). Various statistical methods and techniques along with machine translations, network analysis and natural language processing can be used to analyze the cleaned data (Fan and Gordon, 2014). This stage provides user sentiment information, including how users feel about a company and its products, along with user behavior (e.g. how likely the user might purchase products in response to an advertising campaign) (Fan and Gordon, 2014). A sizeable number of useful trends and metrics regarding user backgrounds, interests, concerns and relationship networks can be obtained at this stage (Fan and Gordon, 2014).

Implementation
The understand stage focuses on the following tasks: clustering tweets (before and after iWatch launch); conducting analysis related to frequent terms and their associations with
iWatch, Google Glass and Samsung; and performing word cloud analysis for the search terms “watch” and “glass”. For example, this study uses the dendrograms feature in the R software’s TwitteR package to classify the topics in the tweets. We then use cluster analysis to cluster topics based on their importance and to classify main subtopics into different clusters before and after the new product launch to examine if the new product launch has any impact on consumer online word-of-mouth.

With the inception of constructs such as value co-creation, value capture and value validation, nowadays conducting marketing research or advertising campaigns both before and after launching new products can be both beneficial and cost effective. Although there may be less social media “buzz” about a new product in the pre-launch than the post-launch phase, analyzing changes in social media insights (timing, frequency and content) gathered from consumers at each phase helps interpret consumer wants and needs, thereby improving the odds of new product success. This insight-gathering approach through social media may contain more noise compared with traditional experimental studies. However, this approach aligns with our proposed framework and offers researchers and marketers a promising alternative to traditional research, for “social media is more than a place to chat with friends – it’s a place for influential audiences to congregate and be heard” (Doering, 2017, p. 1).

Stage 5: present
Framework
Present is the last stage in our current extended framework. The meaning of present in our framework corresponds closest to retention in Campbell’s (1965) model (see Figure 2), but differs slightly in meaning from the present used in Fan and Gordon’s (2014) framework. The reason is that retention in Campbell’s (1965) model “serves to store selected variations for future application” (Bagozzi, 1992, p. 356). “In social evolution, retention exists in the form of social mechanisms such as codes of ethics, standards for publication, lines of authority in universities, journal governance, and so on” (Bagozzi, 1992, p. 356). In knowledge evolution and dissemination, retention exists through an article’s sections such as conclusion, theoretical implications, managerial implications, future research and so on. However, possibly for simplicity’s sake, the present stage in Fan and Gordon’s (2014) framework only goes slightly beyond an article’s results and conclusions sections, as they describe their last stage (i.e. present) as the “results from different analytics...summarized, evaluated, and shown to users in an easy to understand format” (Fan and Gordon, 2014, p. 7). Meanwhile, according to our aforementioned guiding research perspectives (e.g. theories of the middle range) and considering the focus of this research and the space constraints in one article, we use the present stage as the last stage in our framework and discuss retention-related information under this stage. When needed, future research can divide our present stage into more detailed stages, such as present and retention.

Implementation
The present stage involves summarizing, evaluating and presenting research findings, as well as retention-related information for future application. Note that some words retrieved from Twitter may contain typos, abbreviations or acronyms created by respondents. This study keeps such words for authenticity and adds a note or quotation marks to them for clarity. Figure 3 shows the results of clustering topics before and after the iWatch launch. Example results for the pre-launch of iWatch are as follows. “[G]ame” and “gameinsights” are clustered into one group because analysis shows that they have several tweets. The second cluster has “ipad”, “android” and “apple”. These words are clustered into one group.
Figure 3. Clustering tweets for iPhone before and after the new iWatch launch (top: before; bottom: after)
The third cluster contains words like “coins” and “gold” which relate to gameplay and product features. The fourth cluster consists of “apps” and “Instagram”, which are popular social services for customers.

In comparison, the clustering results after iWatch release show the following distinctive sentiment patterns. The leftmost cluster indicates, not surprisingly, that all three brand names, “iWatch”, “applewatch” and “apple watch”, are highly regarded and mentioned. The next cluster reveals that “ios”, “release date” and “rumors” are some keywords consumers care about when discussing iWatch. In particular, “architecture”, “art” and “features” in the third cluster indicate important cues that consumers notice. Lastly, the rightmost cluster suggests that “how to [jailbreak]” is a unique question consumers ask in the Twitter community. This implies that consumers’ decision-making processes might have moved into the new product’s anticipating, buying and using stages, taking a possible “jailbreak” or shortcut, due to the perceived high price of the product.

As iPhone and iWatch belong to the same company (Apple) and their relationship might offer insights related to marketing strategies for new product development, this study also analyzes tweets’ term frequency for “iPhone” before and after the new iWatch launch (see Figure 4 for details). These frequencies indicate several insights and concepts that marketers should consider when releasing their products. For instance, most keywords relate to serving a consumer or providing better services (e.g. “download”, “video”, “promo”, “music”, “movie”, “youtube”), which may have implications as to which product functions garner the most attention from consumers.

Table I provides a word association analysis of brands (iWatch, Google Glass, iPhone and Samsung) and their key features or other attributes that consumers are likely to “buzz” about on Twitter. As Google Glass is a new technology product featured in main technology websites, this study also seeks to determine if “Google Glass” has as good a sentiment in the Twitter community as “iWatch” does. In particular, this study finds that most keywords associated with Google Glass tend to be neutral (e.g. “tech”, “partner”, “save” and “weekend”) (Column 1 in Table I). However, most keywords highly correlated with iWatch tend to be positive (“lifestyle”, “fitness”, “stand”, “wearable” and “amazing”) (Columns 3 and 5 in Table I). Furthermore, when referring to iWatch, consumers talk about associated products like “Porche”, suggesting that how a product is being used in conjunction with other products in an ecosystem may provide reasons for people to talk about it.

Moreover, this study finds that there are more search terms highly correlated with iPhone compared to Samsung. For instance, only four terms (“bluetooth”, “fatwallet”, “phone” and “smart”) highly correlate to Samsung (Column 7 in Table I), while 14 terms (“applesa”, “southafrica”, “lifestyle”, etc.) correlate highly to iPhone (Column 5 in Table I). Also, “fitness” seems to be a keyword associated with iWatch that Twitter users found attractive (Column 3 in Table I). Other words such as “amazing”, “lifestyle” and “Porche” provide rich information regarding iWatch. These attributes or keyword associations are the main features that differentiate iWatch from competitors such as Samsung. Figure 5 demonstrates the results of the word cloud analysis for search terms “iWatch” and “glass”, which, respectively, refer to iWatch and Google Glass.

Conclusions and implications for research and practice
The meteoric rise of social media and their related topics presents opportunities and challenges. This paper attempts to address some of them with regard to theory and practice. From a theoretical standpoint, this paper contributes to the literature by integrating insights from theories of the middle range (Merton, 1949), Campbell’s (1965) model of sociocultural evolution and Fan and Gordon’s (2014) social media analytics.
Figure 4. Term frequency of tweets for iPhone before and after the new iWatch launch (top: before; bottom: after)
<table>
<thead>
<tr>
<th>Term</th>
<th>r**</th>
<th>Term</th>
<th>r**</th>
<th>Term</th>
<th>r**</th>
<th>Term</th>
<th>r**</th>
</tr>
</thead>
<tbody>
<tr>
<td>tech</td>
<td>0.46</td>
<td>applesa</td>
<td>0.85</td>
<td>Applesa</td>
<td>0.82</td>
<td>bluetooth</td>
<td>1.00</td>
</tr>
<tr>
<td>luxottica</td>
<td>0.38</td>
<td>southafrica</td>
<td>0.85</td>
<td>Southafrica</td>
<td>0.82</td>
<td>fatwallet</td>
<td>1.00</td>
</tr>
<tr>
<td>googleglass</td>
<td>0.35</td>
<td>iphone</td>
<td>0.70</td>
<td>iWatch</td>
<td>0.70</td>
<td>phones</td>
<td>1.00</td>
</tr>
<tr>
<td>Anybody</td>
<td>0.33</td>
<td>lifestyle</td>
<td>0.70</td>
<td>Lifestyle</td>
<td>0.67</td>
<td>smart</td>
<td>0.82</td>
</tr>
<tr>
<td>Apps</td>
<td>0.33</td>
<td>caswellchad</td>
<td>0.60</td>
<td>Amazing</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archive</td>
<td>0.33</td>
<td>technology</td>
<td>0.60</td>
<td>Caswellchad</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrow</td>
<td>0.33</td>
<td>appleevent</td>
<td>0.30</td>
<td>Igers</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirms</td>
<td>0.33</td>
<td>applelive</td>
<td>0.30</td>
<td>Insta</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cosa</td>
<td>0.33</td>
<td>begeek</td>
<td>0.30</td>
<td>Phone</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>jmbrandonbb</td>
<td>0.33</td>
<td>ennemi</td>
<td>0.30</td>
<td>Photooft</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>partner</td>
<td>0.33</td>
<td>fitness</td>
<td>0.30</td>
<td>Porsche</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>possession</td>
<td>0.33</td>
<td>hirise</td>
<td>0.30</td>
<td>technology</td>
<td>0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>save</td>
<td>0.33</td>
<td>south</td>
<td>0.30</td>
<td>Photo</td>
<td>0.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>servirebbe</td>
<td>0.33</td>
<td>stand</td>
<td>0.30</td>
<td>Wrist</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wedesignla</td>
<td>0.33</td>
<td>tatouage</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weekend</td>
<td>0.33</td>
<td>twelve</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eyewear</td>
<td>0.28</td>
<td>wearable</td>
<td>0.30</td>
<td></td>
<td></td>
<td>wrist</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Notes: This table contains correlations between the search terms of "glass" and "watch" and different key brands and products from two series of tweets' data set retrieved from Twitter after major products, Apple iWatch and Google Glass, release. The values show correlations between the above paired terms. Note that tweets are kept in their verbatim form, and thus may contain typos, grammatical errors, abbreviations or acronyms; r denotes the Pearson correlation coefficient.
Figure 5.
Word cloud for search terms “watch” and “glass”
framework. Additionally, it offers an extended social media analytics framework that identifies the underlying stages of the social media analytics process. Scholars can apply, modify and extend both our integrated theoretical approaches and our proposed framework in their future studies relating to social media. Moreover, by testing the proposed framework in three critical and interconnected areas – social media (Twitter), marketing strategies for new product development (iWatch, Google Glass) and social media analytic techniques (text mining and sentiment analysis) – this study demonstrates not only how the proposed social media analytics framework can help address the questions at hand (e.g. new product development and adoption), but also how social media are transforming research approaches (e.g. qualitative, quantitative and mixed method) and the very nature of business itself (e.g. increased importance of digital business). In particular, this research explains and demonstrates how our framework and social media analytic techniques (e.g. text mining and sentiment analysis) can assist both qualitative and quantitative researchers in their endeavors. Specifically, qualitative researchers can adopt techniques like sentiment analysis and word association analysis to better understand constructs under investigation and to advance theory development.

With regard to managerial implications, this research validates the proposed framework by providing concrete processes, procedures, examples, evidence and insights related to the studied context, providing managers with ideas on how to create effective and efficient strategies for new product development. For example, social media provide companies with new frontiers to unlock value from and co-create value with their customers. The current study is a testament to the value of our proposed framework by providing evidence that online sentiment can be a “make or break” factor in introducing a new product to the market. When launching new technologies and products, firms might benefit from extracting knowledge and insights from unstructured social media data. This research helps practitioners better understand the value and implications of using the proposed framework along with text mining and sentiment analysis for building marketing strategies for new product development.

Moreover, this research offers new insights on competitive analysis and real-time knowledge extraction by examining the introduction of a critical discontinuous technology (iWatch) and its impact on different brands and products (e.g. Google Glass, Samsung and iPhone) when that technology is far from nascent. Additionally, congruent with the adaptation mechanism in which social systems’ adaptation flows through the two feedback processes in Campbell’s (1965) model of sociocultural evolution (see Figure 2), another way firms might benefit from our framework, as well as from social media and sentiment analysis, is through a recurring process of launch, feedback and re-launch.

**Limitations and future research**

Despite the insights the current study provides, it does have limitations that future research can consider exploring. First, although there is overlap among some stages, and the process may go through several iterations, this study does not discuss overlapping issues and iterations in depth. With regard to overlap and iterations, readers may refer to Campbell’s (1965) and Fan and Gordon’s (2014) work for details.

Second, guided by theories of the middle range (Merton, 1949) and Campbell’s (1965) model of sociocultural evolution, this research extends Fan and Gordon’s (2014) social media analytics framework from a three-stage process to a five-stage one.
However, when applying, modifying or advancing its framework, specific conditions and factors must be considered. In theory, for instance, one may expect to conduct any type of research for various purposes. In practice, however, characteristics associated with social media (platforms, users and data) coupled with current knowledge and tools, present opportunity and challenge, often requiring us to make tradeoff decisions and to pay attention to boundary conditions. As discussed above, Fan and Gordon’s (2014) broad input-process-output approach and three-stage framework are excellent starting points (e.g. most appropriate for exploratory research or studies with loosely defined research objectives and scopes), but risk being misunderstood and inappropriately applied. Although our approach and extended five-stage framework enable us to identify several “favorable” boundary conditions suitable for social media research, such as social media’s capacity to identify an opportunity and to make convergent parallel mixed methods more feasible, this paper does not discuss potentially “unfavorable” boundary conditions in detail. For example, even when identifying a specific opportunity as a social media study’s pre-determined goal, when handling unstructured social media data, researchers should be aware of the approach’s boundary conditions. As boundary conditions relate to social media users’ capacity to articulate both their expressed and latent needs, researchers and social media analytic techniques’ limited capacity to identify insightful patterns from the majority, as well as minority outliers, and different features associated with different social media platforms, etc., they may constrain a study’s scope and overlook critical trends or insights that unstructured data might uncover.

Case in point. Although developing a framework is the primary purpose of this research, using sentiments from Twitter to validate the framework in the context of new product release may limit the generalizability of this study’s results. For instance, because Twitter sets limits on how many characters one can post and re-tweet, as well as there being software restrictions when collecting data from Twitter, conducting more advanced and sophisticated data analysis is currently challenging. However, this study hopes its chosen approaches, topics and findings can serve as a clearer starting point compared to prior research and can provide insights for future studies (e.g. value co-creation with customers in other contexts such as other new product development and advertising campaigns).

Lastly, as data coming from social media are inherently unstructured, the purification and structuring of such data may prove challenging. Moreover, individuals may tweet ironically about a product. However, with the challenge comes the prize. Alongside more structured forms of market feedback (focus groups and laboratory observations), social media analytics holds the promise of gauging customer opinion while simultaneously minimizing the effects of research and methodological artifacts.

References


About the authors
Gavin Jiayun Wu (PhD, University of Arizona) is an Associate Professor of Marketing in the Department of Graduate and Professional Studies in Business, Broadwell College of Business and Economics at Fayetteville State University, USA. His work is published in the European Journal of Marketing, Journal of Marketing Channels, Journal of Research for Consumers, Advances in Consumer Research, Accounting and Taxation, Journal of Financial Counseling and Planning, etc. For more information about his publications, please see: https://scholar.google.com/citations?hl=en&user=Hg_UCSEAAAAJ&view_op=list_works&sortby=pubdate. Gavin Jiayun Wu is the corresponding author and can be contacted at: jwu2@uncfsu.edu or jwu@email.arizona.edu

Zhenning “Jimmy” Xu (PhD, University of Texas at El Paso) is an Assistant Professor of Marketing at the University of Southern Maine, USA. He is particularly passionate about informatics and marketing analytics. His work is published in the Journal of Business Research, Journal of Marketing Theory and Practice, ACM Software Engineering Notes and Journal of Macromarketing.

Saeed Tajdini (PhD, University of Texas at El Paso) is an Assistant Professor of Marketing at the College of Business at Indiana University Southeast, USA. His work is published in the Journal of Marketing Theory and Practice.

Jie Zhang, PhD student, Department of Information Systems and Operations Management, University of Texas at Arlington, USA.

Lei Song (PhD, Drexel University) is an Assistant Professor of Marketing at the School of Business of Stockton University, USA. His work is published in Business Horizons, Journal of Retailing and Consumer Services, Journal of Marketing Channels, etc.

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