



Article

Developing and validating the A-B-C framework of information diffusion on social media

new media & society
2018, Vol. 20(1) 272–292
© The Author(s) 2016
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1461444816661552
journals.sagepub.com/home/nms



Yuhua (Jake) Liang and Kerk F Kee

Chapman University, USA

Abstract

This research addresses the problem of promoting information diffusion, the extent to which information spreads, on social media platforms. Utilizing the number of views, comments, and shares as indicators of diffusion, we developed and validated an original research framework based on the big data approach (using all the blog posts in a university in the year 2013; $N=4120$). This A-B-C framework (1) analyzes the textual features of blog posts using linguistic inquiry and word count (Study 1), (2) applies the former results to build message concepts (Study 2), and (3) creates validated instructional material based on message concepts to promote message diffusion among blog readers (Study 3). This framework supports operational strategies for developing strategic and corporate communication material aimed at increasing diffusion.

Keywords

Big data, blogs, comments, information diffusion, linguistic analysis, research strategy, shares, social media, views

Diffusion is the communication process through which a new idea, behavior, and/or technology spreads through certain channels from a person, an organization, or any unit of adoption to another within a social system over time (Rogers, 2003). Communication researchers have long been interested in understanding information diffusion for

Corresponding author:

Yuhua (Jake) Liang, School of Communication, Chapman University, One University Drive, Orange, CA 92866, USA.

Email: liang@chapman.edu

strategic idea exposure, brand awareness, product marketing, technology adoption, social influence, behavioral change, leadership vision, and so on (Compeau et al., 2007; Dearing, 2009; Rice, 2009; Rogers, 2003; Weigel et al., 2014; Wisdom et al., 2014). However, diffusion research investigating the spread of social media messages is only emerging (Southwell, 2013).

Katz and Lazarsfeld (1955) proposed a model to describe the flow of ideas from mass media through influential individuals in a social system. Under this model, mass media are the primary information sources, but near peers are powerful in reframing and/or filtering the information as secondary channels. The model challenged the common assumption that mass media directly influenced the audience. Granovetter (1973) argued that people seldom act on information from mass media unless it is also passed on by near peers. His work has recently been re-examined in light of a range of media landscape of “ties,” such as weak ties (De Meo et al., 2014), strong ties (Kee et al., 2016), intermediary ties (Grabowicz et al., 2012), and dyadic ties (Michelfelder and Kratzer, 2013).

In the 1990s, communication researchers turned to online discussions as a context for studying information diffusion and Internet users as the primary information source. Notably, Rafaeli and LaRose (1993) researched electronic bulletin boards and found that active adoption was positively related to symmetry in user contribution. Their findings suggested that interactivity was key to information diffusion in a vibrant online community.

Furthermore, Berthold et al. (1998) examined the distinguishing features of messages that received a high number of subsequent references in discussion threads, leading to the diffusion of the message content. They concluded that these messages typically contained 11–25 lines of original text, appropriate subject lines, statements of facts, and no quoted text from inside or outside the threads. Also, these messages addressed other readers and did not include questions/requests, emoticons, or punctuation devices to express emotion. Finally, highly referenced messages were authored by users who identified their gender via names/signatures, and they were mostly males.

Given the integral role of social media in contemporary society (Duggan, 2015), they provide a timely context for studying today’s information diffusion. Social media contents are generated by users for the consumption of both their interpersonal ties and the mass audience. However, the contents are unlikely to diffuse widely unless they are passed on by users to their near peers.

Previous research showed that user interaction with social media can lead to a form of information cascade (Galuba et al., 2010). In some exceptional instances, content can achieve a viral status (Chu, 2011; Kaplan and Haenlein, 2011) or a global cascade (Watts, 2002), perhaps through a critical mass of easily influenced individuals (Watts and Dodds, 2007). This research examines how message use by contributing users drives the diffusion of social media content.

The competitive and saturated information environment

Communication and diffusion efforts today compete to attract the attention of a mass audience (Hutter et al., 2013). In fact, contemporary diffusion efforts need to become more strategic in order to attract and maintain viewers’ attention in the noisy online

marketplace. Therefore, strategic information diffusion refers to the extent to which information spreads over time, beyond its initial appearance.

Diffusion literature on social media provides useful insights regarding factors promoting information spread. One conceptualization regards diffusion as word-of-mouth (e.g. Berger, 2014). A wealth of research has approached understanding diffusion based on users' characteristics (e.g. Cheung and Thadani, 2012). Research also links diffusion to users' viral behavioral intentions (e.g. Alhabash and McAlister, 2015). Kozinets et al. (2010) examined message characteristics of information diffusion and found that bloggers utilized specific blogging strategies to promote readership. In short, the literature connects various user characteristics and motivations, as well as tests message strategies related to diffusion.

Diffusion is a systemic process based on a number of individual adoption behaviors (Kee, 2017). We distinguish three different adoption behaviors as diffusion outcomes. First, *viewing* a social media message is information adoption, where argument quality and information usefulness are considered by readers. Second, *commenting* on a social media message is information engagement, which also triggers the attention of adopters' near peers to the source—the diffusing message. Third, *sharing* a social media message within and across platforms is information dissemination, similar to telling others about a product. Essentially, we focus on diffusion in light of how users expose themselves to new information (e.g. views), interact with the information (e.g. comments), and/or pass along the information to others in their social networks (e.g. shares).

The literature does not yet point to a systematic process that can be utilized to identify the most unique and important predictors of information diffusion in a given social media environment. This study contributes to the literature by developing and testing a three-step process model, enabling such a process of predicting information diffusion using the available data gleaned from a social media environment with a given domain of content (e.g. blog posts in higher education, customer reviews for products). The goal of the study is not to predict adoption or non-adoption; the goal is to predict the extent of three kinds of diffusion (i.e. views, comments, and shares). We identify message features that promote the three distinct social media behaviors, which we propose as a threefold way to measure information diffusion.

The A-B-C framework for information diffusion on social media

Based on the first letters of the three steps' key words, we name this the A-B-C framework of information diffusion on social media. This framework functions by (1) analyzing and linking online content (i.e. blogs) to their linguistic features (i.e. variations in the language on the blogs), (2) building and developing message concepts out of the linguistic features and empirically validating these concepts through content analysis, and (3) creating and testing the instructional material (i.e. a one-page training guide) to support the development of more diffusible content for bloggers.

This research involved a university (hereafter, the University) as a research site and test case to develop the A-B-C framework. The research represented an endeavor between

the University's Strategic and Marketing Communications (SMC) office and faculty to achieve the practical goals of diffusing the university's branding message, inherently embedded in the social media contents related to the University. In order to achieve the stated goals, we describe how integrating big data in the data extraction approach informs the analytic goals of understanding diffusion in this context.

Information diffusion in the Web 2.0 ecosystem with big data

The broad concept of information diffusion can evoke a number of theoretical views spanning across literature in numerous academic disciplines. Following Rogers' (2003) definition, we conceptualize information diffusion on social media as the extent to which individuals engage in communication behaviors that expose themselves to, interact with, and/or pass on online content via one form of social media (e.g. WordPress and YouTube) to another (e.g. Facebook and Twitter) over time. Therefore, diffusion was conceptualized essentially as an interaction between users, information, and systems.

One way of bolstering the predictive effects of a research framework is to reduce or eliminate sampling error. Utilizing big data aids in such a reduction (Parks, 2014). The concept of big data was defined by what has become known as the three V's of big data: volume, variety, and velocity (Kuziemsky et al., 2014; Laney, 2001). Volume refers to the size of the data (measured in terabyte and petabytes) and the number of messages. Variety denotes the diversity of the data (structured and unstructured data). Velocity is the speed at which data are generated and analyzed (real-time or near-real-time processing). Recently, Gandomi and Haider (2015) added three more V's of variability, veracity, and value. Variability describes data flow rates as inconsistent since data flow stems from multiple sources and can create periodic peaks and troughs. Veracity suggests that big data contain important insights despite being imprecise and uncertain. Value means the significant worth stemming from data's large volume.

While volume in terabytes and petabytes is a common criterion for big data, Gandomi and Haider (2015) argued that the notion of size for volume is relative—what is considered big today may not be big tomorrow. Therefore, we turn to another useful definition of big data by Mayer-Schönberger and Cukier (2013): when the size of a sample is equal to the size of the population (i.e. $N = \text{all}$). When $N = \text{all}$, the need for inferential statistical tests is no longer relevant. This study utilizes the population data instead of sampling, with the units of analysis as the individual blog posts rather than the bloggers. We use all the data available for a period of 12 months (i.e. all the blog posts at the University in the year 2013; $N = 4120$). The posts averaged 299 words each, totaling 1,323,571 words. The blog posts provide the basis for predicting diffusion.

Furthermore, this research addresses an important issue from previous research on message diffusion (and similar variants in word-of-mouth, virality, and information cascade). Previous research typically produced correlational results that were not verified through demonstrable message effects using an experimental test. We address this issue by developing an analytic framework that finalizes with actual communication material empirically verified to aid in increasing diffusion. Specifically, the three-step A-B-C framework integrates big data from social media content to ultimately analyze, build,

and create more diffusible content through three inter-related studies carried out in this research.

Linguistic inquiry and word count analysis

A large amount of content is available across thousands of blog posts at the University. Thus, an automated initial approach can yield an expedited and objective analysis regarding message patterns. Specifically, these patterns stem from linguistic variations and how these variations affect the three diffusion outcomes. Linguistic inquiry and word count (LIWC) offers a useful and automated tool to delineate variations among language use. After numerous validation studies, Pennebaker and King (1999) demonstrated that LIWC can identify variations in language use by counting the frequency of words used (for a complete description of each category and validation, refer to <http://www.liwc.net>). Specifically, these variations implied underlying social processes, affective processes, perceptual processes, and personal concerns. For example, positive emotion, an affective process, may be revealed through the counting of words such as love, nice, and sweet.

Most diffusion studies on social media to date employed survey methods (e.g. Compeau et al., 2007; Liu-Thompkins and Rogerson, 2012), network analyses (e.g. Cao et al., 2012; Kee et al., 2016), and/or qualitative methods (e.g. Kozinets et al., 2010; Page et al., 2013). This study uses LIWC to examine the connection between language and information diffusion on social media. LIWC's main advantage is that it objectively searches content for language use based on previously validated categories. LIWC has been applied across numerous research application in examining social media content (e.g. Liang et al., 2014), including those that aim at identifying sharing behavior (Berger and Milkman, 2012). The recent controversial experiment by Facebook that affected its users' emotional states also utilized LIWC to generate positive and negative emotional content (Kramer et al., 2014). Applying LIWC, we start presenting the A-B-C framework by describing the analyze step.

Step A / Study I—analyze the linguistic predictors of information diffusion

Method

WordPress (<https://wordpress.org/>) served as the official online publishing platform to create and publish blog posts at the University. The University blogs were publicized through online channels including email campaigns, personal and professional social media accounts, and the University homepage. Blogs were also retrievable by readers via Google search results.

To track online readership, the SMC office at the University developed an original and proprietary software plug-in for WordPress to collect data about each individual blog post in order to measure diffusion of each blog entry. The original purpose of these data was to idiosyncratically provide feedback to individual staff writers on which pieces of content were diffusing (i.e. readers' viewing, commenting, and sharing). This study utilized these data for the empirical analyses in Study 1.

The University also produced an original blog platform that aggregated all university-related posts onto one single website. This website contained all blog posts written by the SMC office staff, staff writers across various colleges and academic units, and students. Therefore, each unit of analysis in this step concerns a single blog post. All posts ($N=4120$) in 2013 were extracted and formatted for LIWC analysis by reducing to relevant textual content. For each post, LIWC produced a percentage of word use for all 66 linguistic categories (subsequently referred to as simply LIWC categories). The titles of the blog entries were not included in the analysis. For a full description, see <http://www.liwc.net/descriptiontable1.php>.

Information diffusion measures

Views. Google Analytics tracking code provided the number of blog views for each entry. Only pageviews directly tied to a specific blog post were counted; a view was not counted if the user saw a preview of the story but did not click on the preview to read the full story. Repeated views were counted as the data obtained did not assess unique visitors. However, the assumption was that most readers will typically read the blogs once. Data were exported from Google Analytics using the Google Developer API (application programming interface) to synchronize the Google data with the WordPress database.

Comments. The number of comments was collected on each blog post using the Disqus commenting system. Readers who chose to write a comment directly on the blog posts were included in these data. Comments written on other social media platforms that referenced or linked to the extracted blog posts were not included. The final measure of comments referred to the total number of comments made for a given blog post.

Shares. Share counts referred to the total number of times users interacted with a particular URL on any of the following social media platform: Facebook, Google Plus, Twitter, LinkedIn, Digg, Delicious, StumbleUpon, and Pinterest. This proprietary score was developed by the SMC office. It included the number of times where a user shared a URL to a post, clicked the “Like” button on a post that had previously been shared, or re-shared a post which had already been shared by another user. Share count represented an aggregated data of shares across the various social media platforms. Share counts were reported directly from each social media API service, making such data publicly available for any given web URL. We used the web service sharedcount.com to aid with data collection. Because the data were associated with the URL of the blog posts, they did not dependent on a user clicking on the social interaction buttons displayed next to the posts.

Results

Multiple regression analyses applied the 66 LIWC categories to separately predict information diffusion in terms of views, comments, and shares. An analysis of the frequency distribution revealed that each of the outcome variables appeared to follow an exponential decay function. As a result, the three diffusion variables received log-transformation

Table 1. The strongest LIWC predictors in each diffusion outcome (Study 1).

Diffusion outcome	LIWC categories	Unstandardized β	Example words
Views ($R^2 = .58$; $R^2_{adjusted} = .34$)	Achievement	2.33	Earn, hero, win
	Dash	1.58	–
	Function	1.54	Pronouns, articles
	Words per sentence	-2.16	
	Colon	-1.87	
	He/she	-1.54	
Comments ($R^2 = .23$; $R^2_{adjusted} = .13$)	Insight	2.00	Think, know, consider
	Function	1.90	Pronouns, articles, quantifiers
	Motion	1.75	Arrive, car, go
	Word count	1.70	
	Exclaim	1.38	!
	Article	-2.39	A, an, the
Shares ($R^2 = .10$; $R^2_{adjusted} = .03$)	Hear	1.22	Listen, hearing
	Motion	-2.18	Arrive, car, go
	Exclaim	-1.44	!
	Words per sentence	-1.16	

LIWC: linguistic inquiry and word count.

Negative values indicate that the increase in the LIWC category use led to less perspective diffusion outcome.

due to non-normality to minimize effects attributable to violation of statistical assumptions in regression. This transformation also removed entries that had zero frequency counts since the log of 0 is undefined mathematically. This removal occurred for each diffusion outcome separately. The LIWC category variables were left untransformed.

Initially, three multiple regression models included all 66 categories as predictors for each of the diffusion outcome variables (i.e. views, comments, and shares). However, due to multicollinearity concerns, non-statistically relevant categories ($t < 1$) were removed after the initial regression analysis. Then, the final multiple regression model included only statistically relevant predictors. The analyses below included only those significant predictors in the final model.

Due to the log-transformation, the unstandardized β s derived out of the regression model are interpreted differently from a typical regression. In Study 1, the equivalent interpretation is that one unit change in the predictor (e.g. the frequency of first person words such as “I” by an additional of one) led to 1% change in the dependent variable (e.g. 1% additional views). The following analyses reported the overall regression model including all relevant predictors indicated and the strongest predictors according to the t -test of slope difference. In addition, the model fit compared effect sizes, in terms of unstandardized β s and R^2 rather than conventional values of statistical significance since population data were used. Table 1 presents the strongest predictors in the model for views, comments, and shares.

Views. The final regression model included LIWC categories of colons, word count, words per sentence, numerals, functional words, I words, she/he words, adverbs, quantifiers, family words, auxiliary verbs, anger words, achievement words, filler words, periods, dashes, and quotes. The data fit the model well, $R^2 = .58$, $R^2_{\text{adjusted}} = .34$.

Comments. This regression model included LIWC categories of I pronouns, colons, word count, words per sentence, numerals, functional words, exclamations, motion words, cognitive mechanism words, insight words, and articles. The regression model was only a modest fit with the data, $R^2 = .23$, $R^2_{\text{adjusted}} = .13$.

Shares. The regression model included LIWC categories of word count, words per sentence, numerals, friend words, certainty words, hearing words, exclamations, and motion words. The data revealed that the regression model was a poor fit, $R^2 = .10$, $R^2_{\text{adjusted}} = .03$.

Discussion

Study 1 yielded empirically substantial models based on LIWC. For views, the model fit surpassed a number of previous work that utilized LIWC results to make predictions ($R^2 = .58$; $R^2_{\text{adjusted}} = .34$). Such a high R^2 value demonstrated the efficacy of applying LIWC to understand information diffusion on social media. Achievement words, dashes, and functional words positively predicted the number of views. In other words, the more use of these words led to more views. Although the interpretation for dashes and function words remained less clear, achievement words had the strongest effect on views. This relationship suggested that readers were systematically attracted to blog posts that highlighted achievements and/or personal successes, which corresponds to previous work by Nelson-Field (2013).

Alternatively, longer sentences (those with more words per sentence), more colons, and more he/she words actually reduced the number of views. Conversely, decreasing these LIWC categories led to more views, suggesting higher diffusion when simpler language was used. From an analytical standpoint, one possibility was that the log-transformation removed cases when no view occurs, reducing the error in the regression model and increasing the fit. The findings related to punctuation were interesting as they correspond to prior research showing that they impact emotion transmission in computer-mediated content (Rourke et al., 1999), especially when in the titles (e.g. Lupo and Kopelman, 1987; Whissell, 2013). However, given the goal of this study focuses on producing message concepts (Step B), the punctuation findings provided less direct interpretable application, and therefore, they were not considered for the remaining two steps.

The regression model moderately predicted the number of comments ($R^2 = .23$; $R^2_{\text{adjusted}} = .13$). Blog content related to more insights (e.g. think, know, consider), more word count, more functional words, and more motion words led to more comments. Interestingly, functional words increased both comments and views. In conjunction with other findings such as word count, the results suggested that more verbally complex discussions elicited additional discussions through comments. Oddly, the use of articles led to fewer comments. Another possibility was that blogs may have utilized more details

in the descriptions of events. Since the content of the comments was not analyzed, there was ambiguity with respect to the characteristics of comments following blog posts.

Our analysis on comments is similar to the study of online discussions by Berthold et al. (1998). Their focus on studying messages that were highly referenced in discussion threads is similar to the notion of blog posts receiving many comments in our study. Their finding that highly referenced messages typically contained 11–25 lines of original text is similar to our finding that more thoughtful and lengthy blog posts elicited more comments. Since we did not include blog titles in the analysis, future studies can examine the relationship between blog titles and comments, as Berthold and colleagues found appropriate subject lines a good message feature.

Last, the LIWC categories poorly modeled shares counts ($R^2 = .10$; $R^2_{\text{adjusted}} = .03$). In opposition to comments, motion words and exclamations were associated with a lower share count. Hearing words (e.g. hear, listen) had only a small positive relationship with share. However, due to the small effect size and the poor overall model fit, these relationships provided limited predictive utility.

A supplemental analysis examined whether findings were related to the extent of the diffusion or whether the content diffused at all (e.g. received one or more views vs no views). To do so, the prior analysis was repeated by recoding diffused versus not diffused (0/1) and then using the most prevalent LIWC categories (those appearing in Table 1) as predictors on the three diffusion outcomes through a multivariate analysis of variance (MANOVA). However, the multivariate tests for each of the predictors were not statistically significant. These results merited treating the three diffusion outcomes individually, as opposed to uniformly. The results also suggest that diffusion levels assessed in this study required some variations in the diffusion to detect their patterns. Stated differently, LIWC categories may not be helpful when conceptualizing diffusion dichotomously.

Following Study 1, a direction was that the language use aided in distinguishing blog posts that were more diffused in terms of views, comments, and shares. These findings were obtained by extracting data from existing social media, a fairly technologically unsophisticated process. Furthermore, the automated analysis aided in determining the LIWC categories of relevance in the University's communication ecosystem. This notion suggests that the analyze step can be implemented easily.

Study 1 results informed the development of the concepts regarding the message characteristics related to information diffusion. This second step focused on further understanding the linguistic differences and narrowing down to specific concepts and their related conceptual definitions. These definitions can be applied and are actionable in producing instructions material aimed to substantively increase diffusion. The following sections describe Study 2 efforts to organize and validate the concepts.

Step B / Study 2—build message concepts

Method and concept development process

This step builds linguistic findings into message concepts that describe and explain the features of diffusible content. To do so, we utilized the following sequence. First, each of

the finalized linguistic results received a careful examination. Then, the LIWC categories that were significant predictors of each diffusion outcome were identified based on effect size. We extracted the original blogs to better understand why such word usages related to the outcome in the original context. For example, the use of achievement words predicted views in the LIWC analysis. Achievement in LIWC could relate to any form of achievement and included words such as *hero*, *earn*, and *win*.

With this understanding, we further analyzed the blog posts that received a high number of views and had a high number of achievement words themes related to achievement. After assessing the blog posts, we developed the conceptual theme of achievement given the interpretation context of the social system (i.e. the University). The interpretation context was generated by reading blog posts ranked the highest on a specific LIWC category. Since the data easily facilitated ranking the blog posts on the basis of the highest percentage of each LIWC category, three independent coders read the top 10 posts from each category to generate initial conceptual definitions.

Using achievement as an example, the coders read the top 10 posts for achievement to develop and inform the definitions for the message concepts. Pulling from the linguistic findings, a concept also involved success. However, after reviewing the actual blogs, successful events occurred in an extraordinary way (e.g. someone who amazed an audience during a performance). Thus, the success definition included “the extent to which the blog posts describe an act that has been done successfully. This is typically accomplished by the effort of courage (e.g., showing that someone acted outside of comfort zone), and/or skill.”

In addition, this definition coincided with specific events at the University, such as a movie screening or a performance that was well-received. It also included acts of courage, such as when a student, faculty, or alumnus acted in an extraordinary way to help out the University or the local community. Following this process, other concepts were developed anchored to the context of the University.

After developing the initial concepts, a list containing 20 concepts were sorted into those that related mostly to views, comments, or shares. However, the three diffusion outcomes were highly correlated in Study 1. Views correlated substantially with comments ($r = .65$) and shares ($r = .59$). Comments also correlated with shares ($r = .35$). With these correlations, the 20 concepts were not separated according to the diffusion outcomes. Instead, the analytic plan involved using a multivariate regression to separate the strongest, unique predictors in an effort to identify the concepts that had the strongest effect on each diffusion outcome.

We conducted a validation study of the concepts. Three independent coders were trained based on the concepts according to the definitions. In a pilot study, they coded 25 randomly sampled blog posts in the dataset. The results showed that some of the conceptual definitions were redundant or were infrequent in the sample blog posts. Based on these results, the definitions received more revision, resulting in the 14 final concepts. To ensure the utility of the final definitions, an additional 25 randomly sampled blog posts were coded as a pilot test. The intercoder reliability using intraclass correlation reached an acceptable range ($>.80$). Table 2 outlined the final 14 concepts, with their respective definitions, means, and intercoder reliability.

Table 2. Exploratory conceptual definitions for content analysis (Study 2).

Concept name	Conceptual definition	Mean	Intercoder reliability
First person	The extent to which the blog writer is using "I" or using first person voice. The extent to which a blog post utilizes the point of view from the author	1.61	.86
Second person	The extent to which the blog writer is using "You" of addressing the reader as a second person	1.36	.83
Third person	The extent to which the writer describe subject of the blog using he/she/they without showing direct experience	4.60	.82
Achievement	The extent to which the blog post describes an act that has been done successfully. This is typically accomplished by the effort, courage (e.g. showing that someone is outside of typical comfort zone), and/or skill	2.96	.85
Award	The extent to which the blog post is describing the receipt of a credible tribute or the presentation of an award	2.19	.90
Personal connection	The extent to which the blog post is relevant to the audience. More specifically, it refers to the amount of anticipated social affiliation. Blog posts specific about students, faculty, or departments have less personal characteristics where concept that involve more social affiliations have higher personal characteristic	2.54	.79
Conversational	The extent to which a blog post encourages readers by soliciting their discussion or asking them for a direct response via commenting	1.42	.67
Part of a sequence	The extent to which a blog post has connection to another University website or blog	2.64	.66
Incentive	The extent to which a blog post encourages readers to receive some reward that is usually free	1.34	.76
Sharing intent	The extent to which the purpose or intent of the blog post is to cause a desired effect in sharing	1.41	.46
Novelty	The extent to which the blog post provides content that is not commonly viewed	1.70	.69 ^a
Details	The extent to which the message contains descriptive information specific to the blog post at hand	2.82	.87
Past tense	The extent to which the blog post describes a previous event	3.22	.93
Future tense	The extent to which the blog post describes an upcoming event	2.73	.95

All coded measures ranged from 1 (*low*) to 5 (*high*). Reliabilities indicate the intraclass correlation between all three coders.

^aOne coder's results were removed from the analysis in this specific category due to low reliability.

Results

In this concept validation study, the three coders rated 200 randomly sampled blog posts and rated each post according to the conceptual definitions on a scale of 1 (*low*) to 5

(*high*). The intraclass correlation measured the intercoder reliability, which ranged from .69 to .95 ($r_{average} = .81$). The main analyses utilized multiple regression, by regressing each diffusion outcome onto all 14 coded concepts. The results discussed in the following section included only the strongest predictors from the dataset. Most of the concept predictors were positive, indicating an increase in the rating led to an increase in the respective diffusion outcome; a few were negative, indicating an increase in the rating led to a decrease in the respective diffusion outcome. Table 3 synthesized the results.

This section summarizes the findings. We included $\beta_{standardized}$ to compare among the effects. The $\beta_{standardized}$ could be interpreted as the unique effect of each concept, statistically controlling for the other concepts in the multiple regression performed. For views, the final regression model ($R^2 = .14$; $R^2_{adjusted} = .12$) included sharing intent ($\beta_{standardized} = .30$), details ($\beta_{standardized} = .15$), and personal connection ($\beta_{standardized} = .20$). For comments, the concepts had a poorer overall fit with the data ($R^2 = .06$; $R^2_{adjusted} = .03$). The model included first person ($\beta_{standardized} = .26$), achievement ($\beta_{standardized} = .16$), third person ($\beta_{standardized} = .20$), and personal connection ($\beta_{standardized} = -.17$). Relevant to shares, the model ($R^2 = .14$; $R^2_{adjusted} = .12$) included achievement ($\beta_{standardized} = .23$), details ($\beta_{standardized} = .12$), and third person words ($\beta_{standardized} = -.35$).

Discussion

The concepts in Study 2 arose out of LIWC analyses in the first step and received validation. Interestingly, the results in Study 2 pointed to some clear overlaps and contradictions. The concept of *details* overlapped and directly increased both views and shares. This result showed that having more specific and concrete information achieved diffusion by attracting more views, but it also elicited sharing among the audience. In contradiction, the concept of *sharing intent* actually led to an increase in views instead of facilitating readers to share content. One possibility was that explicitly asking readers to share attracted more attention to the content.

Another contradiction involved the concept of *personal connection*, which increased views but decreased comments. One possible explanation was that personal relevancy drew the audience to read the blog posts out of curiosity. However, the readers were attracted to the psychological need to satisfy this curiosity, and satisfaction concluded their interaction with the blog posts. Another possibility was that writing a public comment to a blog post can put the commenter in a complex web of relationships, especially when the commenter knew the character in the blog story as well as other readers within the university. This concept describes personal connection. Therefore, a reader might hesitate and/or choose not to comment to avoid having the comment be (mis)interpreted in various ways. One way to reconcile the discrepant results on the diffusion outcomes was to provide separate and distinct recommendations for increasing views, comments, or shares. Overall, Study 2 provided validation of the concepts from data extracted using blogs online and associated the concepts with diffusion.

Step C / Study 3—create instructional material

Applying these message concepts stemmed from LIWC results, we developed original instructional material to promote information diffusion. The recommendations were

Table 3. Conceptual categories predicting information diffusion ($n=200$) (Study 2).

Content category	Concept	Conceptual definition	Standardized β
Views ($R^2 = .14$; $R^2_{adjusted} = .12$)	Sharing intent	The extent to which the purpose or intent of the blog post is to cause a desired effect in sharing	.30
	Details	The extent to which the message contains descriptive information specific to the blog post at hand	.15
	Personal connection	The extent to which the blog post is relevant to the audience. More specifically, it refers to the amount of anticipated social affiliation. Blog posts specific about students, faculty, or departments have less personal characteristics where concept that involve more social affiliations have higher personal characteristic	.20
Comments ($R^2 = .06$; $R^2_{adjusted} = .03$)	First person	The extent to which the blog writer is using "I" or using first person voice. The extent to which a blog post utilizing the point of view from the author	.26
	Achievement	The extent to which the blog post describes an act that has been done successfully. This is typically accomplished by the effort, courage (e.g. showing that someone is outside of typical comfort zone), and/or skill	.16
	Third person	The extent to which the writer describes the subject of the blog using he/she/they without showing direct experience	.20
	Personal connection	The extent to which the blog post is relevant to the audience. More specifically, it refers to the amount of anticipated social affiliation. Blog posts specific about students, faculty, or departments have less personal characteristics where concept that involve more social affiliations have higher personal characteristic	-.17
Shares ($R^2 = .14$; $R^2_{adjusted} = .12$)	Achievement	The extent to which the blog post describes an act that has been done successfully. This is typically accomplished by the effort, courage (e.g. showing that someone is outside of typical comfort zone), and/or skill	.23
	Details	The extent to which the message contains descriptive information specific to the blog post at hand	.12
	Third person	The extent to which the writer describes subject of the blog using he/she/they without showing direct experience	-.35

Direction on the β coefficient indicates the significant relationship between conceptual definition and diffusion outcome.

synthesized into a one-page instructional guide (Figure 1). The guide created in Step C aimed to directly test the efficacy of the findings reported in the first two steps.

The last step utilized the message concepts to create a guide aimed to increase views, comments, and shares as a way to test their joint effects on diffusion outcomes. These concepts were grouped into several areas and produced direct recommendations and actions taken when writing social media or blogs. To conduct an A/B test, another mock guide served as the control. The control guide contained basic tips for grammar usage, word choice, and writing developed from online tips for better writing. The experimental and control instructions provided a direct test if the content within the experimental instructions empirically increased diffusion. The next section reports the empirical validation process.

Instructional material generation

In an original experiment, undergraduate students ($N=32$) from a departmental subject pool visited the laboratory to complete a study related to “Blogs and Communication.” Upon arrival and informed consent, the experimenter randomly assigned participants to the experimental ($n=15$) or the control condition ($n=17$). The experimenter asked the participants to read the instructional guide carefully and to write their most memorable experience at the University for approximately 30 minutes using WordPress as the blogging platform. The experimenters refrained from offering any topical or content-related suggestions as the participants worked independently in the laboratory on their own blogs. Upon completion, the experimenters thanked the participants for their time. The final dataset excluded blog entries unrelated to the prompt (e.g. the other experiences and experience with a particular instructor).

Instructional material validation

Following the blog generation process described above, a separate group of undergraduate students ($N=123$) each read an average 3.6 randomly assigned blogs from the pool of blogs generated using previously described procedures, resulting in a set of 449 blog ratings. Using an online survey system, students rated each blog using four 7-point single-item semantic differential measures. The instructions prompted participants to indicate their feelings to each of the corresponding diffusion outcomes: Views (*Not read it again/Read it a number of times*), Comments (*Unlikely to post a comment/Likely to post a comment*), Shares via Social Media (*Unlikely to share with others on social media/Likely to share with others on social media*), and Shares in Person (*Unlikely to share with others in person/Likely to share with others in person*).

We discerned the two different sub-types of sharing behavior given recent research suggesting that sharing behavior is contingent on the target audience (Barasch and Berger, 2014; Berger, 2014). Specifically, narrowcasting occurs when the target audience is small or interpersonal, and narrowcasting heightens considerations for others in the sharing process. Alternatively, broadcasting involves a larger public audience and the effect is self-focused. Given these possibilities, we bifurcated the measurement of sharing along in-person and online routes to assess their effects separately. Then, the data

DETAILS

Help readers understand the specifics of the content! Include time, place, names of related persons, reactions, adjectives to describe feelings, the chronology (sequence) of events, direct quotes, and supportive images. Use referents like Attallah Piazza or Memorial Hall. Your readers only have your words to visualize the story the way you do!

ACHIEVEMENT

People love to hear about success! Prevailing over a personal obstacle can be achievement. This can be running in a marathon, graduating college, or overcoming a personal struggle. Inspire your readers! Let them know that hard work and perseverance pays off. Describe the journey that led to the current success.

PERSONAL CONNECTION

You and your readers already have something in common- Chapman University. To establish personal connection in a blog, mention what clubs, classes or activities the person is involved in at school. Including these details will catch the attention of readers who will support someone that they know.

SHARING INTENT

Need your blog to reach a lot of people? Simply ask the reader to share your post with their friends. Stress the importance of sharing the information with others and what their awareness means. Let readers know you are interested in getting their feedback.

FIRST PERSON

You tell the story! This gives the article a personal touch. Speaking from personal experience creates positive feedback from readers. You can give advice, include quotes, describe inner feelings or thoughts to use first person as an advantage. One example is a blog entitled "How I Started My Law Firm Right After Law School."

THIRD PERSON

If you really do not have any direct experience with the subject of the article, using the "he said-she said" voice is acceptable. However, be sure to include as much detail as possible to help the reader understand the importance of the blog.

Open Wordpress.com
Click Here: [button]
Write Your Title
Start Typing Your Story!
Preview It, Save It, Post It!

(a)

CHOOSE YOUR TOPIC

Decide what you want to talk about. Your blog will be posted for readers to enjoy, so make sure you like what you are writing!

START WRITING!

Begin writing your blog! Make sure it has a beginning, middle and end. Get your point across to your readers. Be sure to tell a story.

PROOFREAD

Once you are done writing, look for any mistakes in your article. Watch out for errors in spelling, style and grammar. Read through it and make sure it flows well!

HAVE FUN!

Don't stress the small stuff. Typing your article can be a difficult process, but just have fun with it. Your blog will soon be out there for readers to see. Be proud of your work!

OUTLINE IDEAS

Organize your ideas so you can write efficiently. Having your thoughts outlined will make the writing process simpler.

FINISHING TOUCHES

After you are done with editing your blog, get ready for publishing! Add your title and preview the page. Don't forget to include your name so readers will know who wrote the article!

Open Wordpress.com
Click Here: [button]
Write Your Title
Start Typing Your Story!
Preview It, Save It, Post It!

(b)

Figure 1. Experimental and control instructions in Study 3: (a) experimental instructions and (b) control instruction.

were recoded to compare ratings in the experimental versus those in the control condition. See Figure 1 for experimental and control messages. The next section reports the findings.

Instructional material analysis and results

One-tailed *t*-tests examined whether the specified directional effect of the experimental condition was indeed more effective than the control condition on views, comments, and shares (in terms of sharing on social media and in person). All measures ranged from 1 to 7 (highest). For views, the two conditions did not differ, $t(447) = .99$, $p = .16$, $d = .09$. The experimental condition (mean [M] = 2.96; standard deviation [SD] = 1.85) influenced comments more than control ($M = 2.67$; $SD = 1.63$), $t(447) = 1.81$, $p = .036$, $d = .17$. Sharing on social media was higher in the experimental condition ($M = 3.08$; $SD = 1.90$) than the control condition ($M = 2.67$; $SD = 1.69$), $t(446) = 2.45$, $p = .007$, $d = .23$. Share count as measured by sharing in person was more influential in the experimental condition ($M = 3.56$; $SD = 1.96$) than the control condition ($M = 3.26$; $SD = 1.94$), $t(447) = 1.66$, $p = .049$, $d = .16$. To minimize any concerns with violation of assumptions, these findings were verified using nonparametric tests. Overall, with the exception of views, the data supported the efficacy of the instructional material.

Discussion

Results in the create step demonstrated the robust effect of the instructional material for comments and shares. The robustness is characterized only by asking participants to read and apply the material, without full extensive training or explanation by others. The findings also offered some evidence of external validity to support the original research goals of yielding instructional content that promotes information diffusion. However, the instructional material failed to increase views, at least in perceived likelihood to read the content again. Inclusion of a manipulation check would strengthen the analysis to assess whether the effect of the instruction was attributed to the respective concepts. Future study might further validate the measure of views used in Study 3 by providing participants a preview and asking them whether they would likely click on the preview to view the actual blog post. In addition, the sharing findings applied for both narrowcasting and broadcasting dimensions, demonstrating the robust effect of the instructional material on sharing.

A concern followed that the effect sizes of the experimental condition on comments and shares were modest. Future work might amend this shortcoming by elaborated training or discussion of the instructional material prior to implementation. Given some concept categories that conflict from Step B, a more effective testing strategy would be to produce a specific recommendation set based on each diffusion outcome and test their effects separately. The last concern was that capturing viewing behavior through self-report could only yield data about a viewing again as opposed to attracting initial view. This limitation might have accounted the lack of views in the current design.

Our goal was to experimentally test the instructional material. The assumption was that if the material worked under the most non-ideal circumstances, they should be robust

and have stronger external validity. Regardless of the shortcomings, Step C yielded empirical support for the three-step analytic framework in producing measurable change in information diffusion.

Overall discussion

To summarize, the three steps and associated studies offered useful and original research. In Step A, the regression analyses yielded interesting results regarding how language use affected diffusion in terms of views, comments, and shares. Some LIWC categories overlapped across the different diffusion outcomes. The direction of the effects even differed, posing possible interesting interpretations for how language use may relate to different outcomes of diffusion. Step B reconciled the findings in Step A by producing concepts that were further sorted with content analysis. Based on the collective findings in these steps, we constructed instructional material that received experimental validation in comparison with a control set of instructional material in Step C. The goal of our work is the development and validation of the A-B-C framework. Step C served as a validation of the previous steps in a comparison test. In other words, this research suggests that the three-step framework can be validated in the approaches outlined in the article.

The linguistic concepts identified clearly have theoretical implications as message attributes for diffusion literature (Rogers, 2003; Walsh et al., 2004). According to Wisdom et al. (2014), diffusion research had mainly focused on studying the innovations, the individuals, the organizations, and external systems as units of analysis. This research contributes to the literature by studying messages as the unit of analysis, exploring message attributes that prompt adoption and promote diffusion. The general A-B-C framework may serve as an approach to stimulate diffusion studies across various social media platforms, developing a new theoretical understanding of message design for diffusion. Furthermore, this study proposes an operationalization of information diffusion based on the three distinct outcomes of views, comments, and shares.

Practically, future application and efforts might focus on the development of an original word-search dictionary (e.g. Donohue et al., 2014) to produce an automated feedback system that guides the production of more diffusible content. We envision this system to be an open-sourced plug-in for WordPress such that it would provide an overall score of potential diffusion as authors compose their blogs. This system would appear like a password feedback system often seen on financial websites, one that provides immediate feedback to users if their password is weak, moderate, or strong. This feedback system can be designed to tap continuous data analytics of all available and ongoing blog posts, making it a big data application. This system should be coupled with instructional material used to train bloggers. A limitation is that as more and more bloggers use this plug-in, blog posts may become increasingly similar over time. This feedback loop can potentially decrease the distinctiveness necessary to promote diffusion.

The series of three steps and studies should be interpreted in light of the limitations. The major limitation is that the studies theoretically and empirically focused on a particular population and messages from that population—The University. Although this focus might also be conceptualized as a strength given the ability to develop and validate

the A-B-C framework, caution should ensue when attempting to apply the identical linguistic results to other populations or contexts.

Moreover, the data involved only messages (blog posts) written by a number of bloggers, despite that the blogs have the potential to reach a larger population beyond the bloggers. While it is unlikely in this particular case, there exist possible sources of errors, such as to contend with bots, semantic derivatives, and problems of irony and ambiguity that are multiplied by the volume of data. In addition, the regression model fit for the concepts in Study 2 was only modest ($R^2 = .06-.14$). Additional analyses may also probe interactions and combine the analyses with LIWC results to ascertain the strongest predictors.

Another limitation involved shrinkages of overall R^2 compared to the adjusted values. In Study 1, views decreased from .58 to .34, comments decreased from .23 to .13, and shares decreased from .10 to .03. This shrinkage may relate to a sample estimation bias due to the sample-to-variable ratio when a large number of LIWC categories were applied. Interpretation of the LIWC model fit should follow caution presented here.

Furthermore, the views metric utilized in this research did not account for repeated views, which may have inflated the number of views. On the contrary, comments written on other social media platforms which referenced or linked to the extracted blog posts were not included in the data. This may have deflated the number of actual comments. Also, we do not have a complete explanation for why colons, periods, dashes, and some LIWC categories have an impact on the diffusion outcomes. Finally, the relationship among the diffusion outcomes of views, comments, and shares presents an opportunity for future research.

In conclusion, the Analyze–Build–Create framework operationalizes a clear process for producing strategic and corporate communication instructions to increase information diffusion. The big data basis of this framework also allows real-time analysis to continuously monitor how language use affects information diffusion in the ever-changing social media landscape. Although the results of big data studies may not always explain why a pattern emerges, the identification of what pattern emerges is equally powerful (Mayer-Schönberger and Cukier, 2013). It is our hope that this framework serves to accelerate and shorten the research process to create more diffusible content in the competitive social media environment.

Acknowledgements

The authors thank the Strategic Marketing and Communications Office at Chapman University for their assistance in obtaining in data used in this study.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Alhabash S and McAlister AR (2015) Redefining virality in less broad strokes: predicting viral behavioral intentions from motivations and uses of Facebook and Twitter. *New Media & Society* 17(8): 1317–1339.

- Barasch A and Berger J (2014) Broadcasting and narrowcasting: how audience size affects what people share. *Journal of Marketing Research* 51(3): 286–299.
- Berger J (2014) Word of mouth and interpersonal communication: a review and directions for future research. *Journal of Consumer Psychology* 24(4): 586–607.
- Berger J and Milkman KL (2012) What makes online content viral? *Journal of Marketing Research* 49(2): 192–205.
- Berthold M, Sudweeks F, Newton S, et al. (1998) It makes sense: using an autoassociative neural network to explore typicality in computer mediated discussions. In: Sudweeks F, McLaughlin M and Rafaeli S (eds) *Network & Netplay: Virtual Groups on the Internet*. Menlo Park, CA: MIT Press, pp. 191–219.
- Cao L, Lin Y, Sun X, et al. (2012) Whisper: tracing the spatiotemporal process of information diffusion in real time. *IEEE Transactions on Visualization and Computer Graphics* 18(12): 2649–2658.
- Cheung CMK and Thadani DR (2012) The impact of electronic word-of-mouth communication: a literature analysis and integrative model. *Decision Support Systems* 54(1): 461–470.
- Chu S (2011) Viral advertising in social media. *Journal of Interactive Advertising* 12(1): 30–43.
- Compeau DR, Meister DB and Higgins DA (2007) From prediction to explanation: Reconceptualizing the perceived characteristics of innovating. *Journal of the Association for Information Systems* 8(8): 409–439.
- Dearing JW (2009) Applying diffusion of innovation theory to intervention development. *Research on Social Work Practice* 19(5): 503–518.
- De Meo P, Ferrara E, Fiumara G and Provetti A (2014) On Facebook, most ties are weak. *Communications of the ACM* 57(11) 78–84.
- Donohue WA, Liang Y and Druckman D (2014) Validating LIWC dictionaries: the Oslo I accords. *Journal of Language and Social Psychology* 33(3): 282–301.
- Duggan M (2015) *Mobile Messaging and Social Media 2015*. Pew Research Center. Available at: <http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/> (accessed 1 April 2016).
- Galuba W, Aberer K, Chakraborty D, et al. (2010) Outtweeting the twitterers—predicting information cascades in microblogs. In: *Proceedings of the 3rd conference on online social networks*, vol. 39, pp. 1–9. October 1, 2010. Berkeley, CA
- Gandomi A and Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. *International Journal of Information Management* 35(2): 137–144.
- Grabowicz PA, Ramasco JJ, Moro E, et al. (2012) Social features of online networks: the strength of intermediary ties in online social media. *PLoS ONE* 7(1): e29358.
- Granovetter MS (1973) The strength of weak ties. *American Journal of Sociology* 78(6): 1360–1380.
- Hutter K, Hautz J, Dennhardt S, et al. (2013) The impact of user interactions in social media on brand awareness and purchase intention: the case of MINI on Facebook. *Journal of Product and Brand Management* 22(5): 342–351.
- Kaplin AM and Haenlein M (2011) Two hearts in three-quarter time: How to waltz the social media/viral marketing dance 54(3): 253–263.
- Katz E and Lazarsfeld PF (1955) *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New Brunswick, NJ: Transaction Publishers.
- Kee KF (2017) Adoption and diffusion. In: Scott C and Lewis L (eds) *International Encyclopedia of Organizational Communication*. Hoboken, NJ: Wiley-Blackwell.
- Kee KF, Sparks L, Struppa DC, et al. (2016) Information diffusion, Facebook clusters, and the simplicial model of social aggregation: a computational simulation of simplicial diffusers for community health interventions. *Health Communication* 31(4): 385–399.

- Kozinets RV, De Valck K, Wojnicki AC, et al. (2010) Networked narratives: understanding word-of-mouth marketing in online communities. *Journal of Marketing* 74(2): 71–89.
- Kramer ADI, Guillory JE and Hancock JT (2014) Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111(24): 8788–8790.
- Kuziemsky C, Monkman H, Collins S, et al. (2014) Big data in healthcare—defining the digital persona through user contexts from the micro to the macro. Contribution of the IMIA organizational and social issues WG. *Yearbook of Medical Informatics* 9(1): 82–89. Available at: <http://dx.doi.org/10.15265/IY-2014-0014> (accessed 1 April 2016).
- Laney D (2001) 3D data management: Controlling data volume, velocity, and variety (Vol 2). Technical report, Meta Group.
- Liang Y, DeAngelis BN, Clare DD, et al. (2014) Message characteristics in online product reviews and consumer ratings of helpfulness. *Southern Communication Journal* 79(5): 468–483.
- Liu-Thompkins Y and Rogerson M (2012) Rising to stardom: an empirical investigation of the diffusion of user-generated content. *Journal of Interactive Marketing* 26(2): 71–82.
- Lupo J and Kopelman RE (1987) Punctuation and publishability: a reexamination of the colon. *American Psychologist* 42(5): 513.
- Mayer-Schönberger V and Cukier K (2013) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York: Houghton Mifflin Harcourt.
- Michelfelder I and Kratzer J (2013) Why and how combining strong and weak ties within a single interorganizational R&D collaboration outperforms other collaboration structures. *Journal of Product Innovation Management* 30(6): 1159–1177.
- Nelson-Field K (2013) *Viral Marketing: The Science of Sharing*. Sydney, Australia: Oxford University Press.
- Page R, Harper R and Frobenius M (2013) From small stories to networked narrative: the evolution of personal narratives in Facebook status updates. *Narrative Inquiry* 23(1): 192–213.
- Parks MR (2014) Big data in communication research: its contents and discontents. *Journal of Communication* 64(2): 355–360.
- Pennebaker JW and King LA (1999) Linguistic styles: language use as an individual difference. *Journal of Personality and Social Psychology* 77(6): 1296–1312.
- Rafaelli S and LaRose RJ (1993) Electronic bulletin boards and “public goods” explanations of collaborative mass media. *Communication Research* 20(2): 277–297.
- Rice RE (2009) Diffusion of innovations: theoretical extensions. In: Nabi R and Oliver MB (eds) *Handbook of Media Effects*. Thousand Oaks, CA: SAGE, pp. 489–503.
- Rogers EM (2003) *Diffusion of Innovations*. 5th ed. New York: Free Press.
- Rourke L, Anderson T, Garrison DR, et al. (1999) Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education* 14(2): 51–70.
- Southwell BG (2013) *Social Networks and Popular Understanding of Science and Health: Sharing Disparities*. Baltimore, MD: Johns Hopkins University Press.
- Walsh G, Gwinner KP and Swanson SR (2004) What makes mavens tick? Exploring the motives of market mavens’ initiation of information diffusion. *Journal of Consumer Marketing* 21(2): 109–122.
- Watts DJ (2002) A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences* 99(9): 5766–5771.
- Watts DJ and Dodds PS (2007) Influentials, networks, and public opinion formation. *Journal of Consumer Research* 34: 441–458.
- Weigel FK, Hazen BT, Cegielski CG, et al. (2014) Diffusion of innovations and the theory of planned behavior in information systems research: a metaanalysis. *Communications of the Association for Information Systems* 34: 31.

- Whissell C (2013) Titles in highly ranked multidisciplinary psychology journals 1966–2011: more words and punctuation marks allow for the communication of more information. *Psychological Reports* 113: 969–986.
- Wisdom JP, Chor KHB, Hoagwood KE, et al. (2014) Innovation adoption: a review of theories and constructs. *Administration and Policy in Mental Health and Mental Health Services Research* 41: 480–502.

Author biographies

Yuhua (Jake) Liang is an assistant professor in the School of Communication at Chapman University. Dr Liang's work examines the theoretical connections between emerging technologies (e.g. participatory systems, agents, and robots) and their practical persuasive effects (see *Communication and Influence Technology Lab*; <http://www.citlab.org>).

Kerk F Kee is an associate professor in the School of Communication at Chapman University. Dr Kee's work focuses on the diffusion of innovations in organizational, health, and environmental contexts (see www.ekerk.com).