HHS Public Access
Author manuscript
Hum Commun Res. Author manuscript; available in PMC 2018 October 01.

Published in final edited form as:
Hum Commun Res. 2017 October ; 43(4): 545–558. doi:10.1111/hcre.12114.

Vectors into the Future of Mass and Interpersonal Communication Research: Big Data, Social Media, and Computational Social Science

Joseph N. Cappella
Annenberg School for Communication, 3620 Walnut St., University of Pennsylvania, Philadelphia, PA 19104-6220, USA

Abstract

Simultaneous developments in big data, social media, and computational social science have set the stage for how we think about and understand interpersonal and mass communication. This article explores some of the ways that these developments generate 4 hypothetical "vectors" – directions – into the next generation of communication research. These vectors include developments in network analysis, modeling interpersonal and social influence, recommendation systems, and the blurring of distinctions between interpersonal and mass audiences through narrowcasting and broadcasting. The methods and research in these arenas are occurring in areas outside the typical boundaries of the communication discipline but engage classic, substantive questions in mass and interpersonal communication.

Keywords

big data; computational social science; merging mass and interpersonal communication; recommendation systems

The convergence of big data, social media, and computational social and communication science allow researchers to rethink interpersonal and mass communication. The volumes of communication data at our disposal -- and readily harvested -- require computational approaches to their understanding (Shah, Cappella, & Neuman, 2015), provide access to public opinion that is unique and unprecedented (Japec et al., 2015), yield information about interpersonal connections that dwarfs previous work (Lazer, et al., 2009), and reorient the way media exposure needs to be conceptualized (Cappella, Kim, & Albarracin, 2015).

In this article, I explore some of the ways that the confluence of big data, computational methods, and social media have created insights and reformulated questions about the nature of communication in the new millennium's research agenda. The exploration takes the form of four hypothetical "vectors" – directions if you prefer – into the next generations of interpersonal/mass communication research. These vectors are derived from contributions and contributors outside the communication discipline but address important questions that, in my judgement, reside at the discipline's core. These questions and vectors include: (1)

215-898-7059 (O), joseph.cappella@asc.upenn.edu.

What structures of communicative networks enhance and retard the diffusion of social influence? The first vector describes approaches showing how networks with dense interpersonal connections yield limited social influence. (2) Can we model media effects incorporating both interpersonal connections and mass communication effects? The accompanying vector focuses on agent-based modeling techniques that move analyses "from factors to actors" (Macy & Willer, 2002) and their application in additional research on social networks, interconnectedness, and social influence. (3) How can efficiencies of mass communication campaigns begin to match the effectiveness of messages tailored to the person-as-target? The third vector, how recommendation systems offer new venues for normative influence, informs this classic social influence paradox. (4) How does the content of interpersonal communication change as it moves from personal exchanges to mediated messages that scale to broader audiences? The nature of narrowcasting and broadcasting social media messages is the vector addressing this contemporary question.

These questions and directions represent lines of research actively being pursued in network science, computational social science, marketing, and computer science, although often not using the language and terminology of communication research. Research into the now-blurred distinction between mass and interpersonal within the field of communication can, in my opinion, profit greatly from the methods and findings from other disciplines with strong computational leanings. With full recognition that other vectors exist or are in development that are outside of my ken, these four vectors offer potential for new methods and new inquiries to some of the core questions in mass and interpersonal communication.

Vector 1: What Structures of Communicative Networks Enhance or Retard the Diffusion of Social Influence?

Networks with dense interpersonal connections may yield limited social influence

One truism of mass communication effects research is that exposure is the *sine qua non* of media effects (Hornik, 2002) and communication campaigns often fail because of inadequate levels of exposure (Snyder & Hamilton, 2002). As social media have become a fact of everyday life, various voices have argued loudly that exposure to information through social – that is interpersonal – contacts is an increasingly consequential source of information and influence (Cappella et al., 2015).

One obvious consequence of the exposure–effects linkage is that the greater the connections among persons in a social network (i.e., the more dense the network), the greater the chances for information sharing among network members and, therefore, the greater the impact of the information shared. Networks that are densely interconnected should yield more information sharing and influence. The greater the connectivity between persons within a network's communities and between communities would lead to information sharing and social influence under simple assumptions of diffusion and retransmission within and between members of the communities.

However, these commonsense conclusions have been challenged by recent data. The availability of large-scale behavioral social media data allows unprecedented network analyses of who communicates with whom, about what, and to what degree. Rather than simply assuming or simulating patterns of information transmission and retransmission, recent network research has been able to focus on large-scale networks with certain kinds of behavioral data allowing empirical re-evaluation of core ideas about network structures and social influence processes.

These big data studies of network influence have produced some conclusions suggesting that dense networks can yield less social influence than sparse ones. A variant on this theme is the observation that those on the periphery of a network may be as important in some contexts as those who are highly connected and "reside" in central communities of the network. We will explore both of these insghts below.

Short versus long chains—The core idea behind most studies of diffusion is that the adoption of an innovation or an idea (i.e., social influence) will move throughout a network beyond several steps removed from the initial source cascading to other recipients from the initial adoption or acceptance of the idea. Goel, Watts, and Goldstein (2012) take issue with this notion in network research. They argue that modeling the transmission of an idea or an innovation like an infection moving through a population is misleading when applied to more complex contagions. This core assumption of information transmission (that it is like transmitting an infection – passive) in a connected network is simply incorrect when applied to a variety of examples of information transmission for adoption and acceptance in the online world.

The authors examine seven different diffusion data sets, three having to do with the acceptance of an action by another person in the network (e.g., playing a game together or watching a video together while synchronously sending messages). The other four involve transmission of news stories, videos, political views, or voice credits for a new protocol. The most important finding for the world of interpersonal-mass communication is that the overwhelming proportion of transmissions occur in one step in the network diffusion process. This result is consistent across all seven data sets. There is a very small proportion of diffusions that occur beyond that initial transmission, but these number far less than 1%.

The authors consider various objections to their findings including the idea that there are data sets in which transmission will occur beyond one step and through several generations as is the case with biologically based viral infections and pandemics. They consider, for example, some particular Twitter elements and events that had a virality greater than some of the other examples that they used, but still find no strong evidence for transmission beyond one step. They also reject the hypothesis that the usual transmission assumption in network influence models (e.g., a simple viral infection model that assumes that passive exposure is sufficient) does not work for most cases of real social influence which require acceptance on the part of the receiver and deliberate acts of transmission and retransmission by senders.

Effects research in communication is concerned with innovations and ideas and memes and their acceptance and rejection, and not simply passive exposure. As network researchers have gained access to larger sets of behavioral data from social media, the random mixing models of early diffusion research and their implications about interpersonal communication have fallen by the wayside in favor of quite different assumptions about how interpersonal communication is important in social influence processes and network. Goel et al. (2012) acknowledge that much of the early work on the transmission of media messages in a two-step flow model may in fact be reasonably accurate and all that is necessary to make the modeling of social influence work to capture media plus social influence (Katz & Lazarsfeld, 1955).

Activists vs. slacktivists—Another line of research that raises questions about the importance of dense interpersonal connections for social influence is the research on political movements and political activism under the name of slacktivists vs. activists (González-Bailón & Wang, 2016; see also Neubaum & Krämer, this issue). The implicit assumption in network studies has been that social influence advances when interpersonal connectivity is high. The more sparse the connections, the less likely the social influence. While this would certainly be true in the extreme case where the complete absence of social connectivity makes influence impossible, recent work has suggested that macrosocial influences (at the societal level) can occur even with relatively non-dense connections at the periphery of networks.

González-Bailón and Wang (2016) studied a 1-month period close to the development of the Occupy movement in the United States and in response to a similar movement that developed in Spain. The data for the networks they studied come from the Twitter sphere involving tweets and retweets in the networks that are created by sending and resending information about the movements.

Algorithmic methods identify communities are identified using algorithmic methods. Individuals who bridge between communities serve to connect them. The relevance and importance of weak ties and the brokering of information between communities within the network occurs through the “interpersonal” mechanism of tweeting not necessarily as a mechanism of social and political influence but rather as an almost indiscriminate broadcasting (see vector 4 below) of available information that may or may not be influential. Nevertheless, communities pass on the information perhaps in part due to the low cost of doing so.

Barberá et al. (2015) reach the empirical inference that those on the periphery of a network are crucially important in social protests. The very small number of individuals who are at the core of the social protest are extremely active but have limited reach because of their numbers and connectivity. The periphery -- which is behaviorally and cognitively less involved in social activism -- is much larger in its numbers but less committed. They are sometimes called “slacktivists” in service of the idea that they are not active in the world of social protest but instead are passive–yet not disinterested. Their sheer number and at least passive involvement in the particulars is important to the protest movement. The number of peripherally connected individuals outweighs their lowered centrality within the network, and their influence in the social system is their number not the density of their connections. The authors show these effects in three different protest movements.

Large-scale social media data that are primarily behavioral rather than self-reported allow the examination of interpersonal connections and communication in ways that are unprecedented. These data, in turn, allow an examination of established assumptions about the importance of dense connections, the likelihood of the activation of long chains of connection, the appropriateness of assumptions such as “simple contagion” (= “acceptance and influence”). These new data sources and methods allow research into communication that blurs the distinction between interpersonal and mass influence allowing the evaluation of established ideas such as simple contagion. The direction depicted in this first vector is itself too simplistic as “what is accepted” (ideas, innovations, information, memes, etc.) will likely differ by topic and network (Simmons, Adamic, & Adar, 2011; Weng et al., 2013).

The particular findings discussed in this section, while provocative and in some senses at odds with conventional wisdom, are less important than the fact that macrosocial influence (diffusion of information) and exposure (knowledge) occur through mechanisms that are obviously at the interface of interpersonal communication, targeted to individuals and to broader reach through the same social media mechanisms. Unprecedented and detailed data are available to examine the micro- and macrosocial influences and patterns of exposure through these mass-interpersonal sources. The possibility of re-examining our long established hypotheses about networks and interpersonal influence at a much larger scale is a real option.

Vector 2: How Can our Theories and Models of Media Effects Simultaneously Incorporate Both Interpersonal Connections and Mass Communication Effects?

Moving from factors to actors

This vector has two prongs. The first is the general idea about how to best model human social behavior using modern techniques such as agent-based modeling (Wilensky & Rand, in press) and the second is using such applications to support the claims of vector one, namely that dense network structures may not be the most likely routes to social influence. The typical approach to statistical modeling that most of us have used for eons uses variables (i.e., factors) that the targeted population measures and models. One can build statistical, graphical, or computational models around such variables. Variables or “factors” are prominent and focal. The actors or persons are assumed to be independent of one another rather than connected and the research is usually designed to insure independence. In these cases, the factors approach is appropriate.

But when the actors are connected or interact in some way, not only is there statistical dependency in need of correction but, additionally, a modeling opportunity focused on actors, each of whom possesses the factors (or variables) of interest. Developments in software and modeling technology allows complex person-by-variable models. The basic idea here is that every person in the system under investigation is a “carrier” of all of the variables that are a part of the modeling effort. So, instead of having an N variable model applied to K independent people, we instead have an N-by-K variable model in which every person interacts with those to whom they are connected through the N variables that are part of the modeling process. What is unique about such computational models is that the interpersonal influence components and degrees of connectivity must be modeled explicitly rather than bypassed in the modeling endeavor or statistically adjusted (Raudenbush & Bryk, 2002).

Macy and Willer (2002) have argued that core questions in sociology have or will (or should) move in the direction of studying the interaction of actors, influence patterns and connectivity, rather than the aggregating interaction of variables across actors, treating clustered actors as if they were not connected, or controlling connected actors' connectivity out of the modeling. Agent based modeling provides a route to such theorizing. Important publications using this approach make clear that both the structures of connectivity and the variables pertinent to social influence must be considered simultaneously. The invention and dissemination of powerful modeling tools such as agent-based modeling (Wilensky & Rand, in press) and statistical approaches that model at both the factor and actor levels within the same analysis (i.e., multilevel models) offer methodologies to bridge the factors of mass communication effects and the actors who carry out interpersonal influences.

One application of the factors and actors approach is that of Centola and Macy (2007). They carried out formal models of complex forms of social influence (in contrast to simple passive ones where contact equals influence). They were able to show that complex contagions require multiple sources mimicking norms, credibility, and other components of social influence, and that complex innovations will not move as readily across weak ties as simple ones do, instead requiring "wider" bridges across network enclaves. The upshot of these studies argues that the classic "strength of weak ties" (Granovetter, 1973) is itself problematic when considering complex rather than simple contagions. Other formal modeling efforts have shown that even highly unpopular norms can in fact cascade throughout a network even when the vast majority of actors oppose the convention (Centola, Willer, & Macy, 2005) and that very high and very low levels of consolidation and homophily are problematic in the diffusion of behaviors (Centola, 2015).

These findings are not confined to the world of formal modeling. Centola (2010) tested the idea that networks with tight clusters of interpersonal connection, and separation between clusters, would be worse than networks with more connectivity (like random graphs) for complex behaviors such as adopting a health behavior. The resulting models and data show the opposite to be true. This study involved a decision to register for an Internet-based health forum. Participants were randomly assigned to one of two experimental conditions based on the kinds of contacts that they had available to them. One was a clustered lattice network (small clusters of people in communication with few connections between clusters) and the second was a random network of contacts. The number of contacts that each person had regardless of network that they were assigned to was the same. The set of connections was not able to be altered by the participants in any way and the communications that they got from each other were mediated by the experimenters. So *de facto* everybody in each condition had the same social neighborhood so to speak as everyone else. Any differences in the acceptance of the idea of participating in the Internet-based health forum would be the result of the structure of the networks to which people were assigned.

The key findings show that the more clustered networks yielded more rapid and more extensive adoptions than the random networks across six different trials. This conclusion obtained with networks of different size as well. Centola concludes that the diffusion and adoption of complex health behaviors are best suited to clustered networks rather than casual contact networks and that this would especially be true for behaviors that are costly, difficult, or contrary to existing social norms.

The factors to actors approach has two consequences for the mass/interpersonal distinction. The formal modeling of factors affecting social influence within various network structures has challenged core assumptions in interpersonal communication, such as the strength of weak ties and too-simple assumptions that contagion equals influence for complex issues. The vector also can have an orienting effect inviting researchers to employ available formal and statistical models that simultaneously treat factors and actors. The consequence can be more representative models and tests of the real world of social influences that have both mass (factor) processes and interpersonal (actor) structures.

Vector 3: How Can the Efficiencies of Mass Communication Campaigns Begin to Match the Effectiveness of Messages Tailored to the Person-as-Target?

Recommendation systems offer new venues for normative influence

Recommendation systems refer to a broad array of input from users of online content. Included are comments in response to online news, products, or other social, political, cultural, or entertainment stimuli; simple likes and dislikes for these items; and other forms of “flagging” (e.g., ratings) of these same items. One can view each of these forms as information indicative of existing social norms approving, interpreting, and disapproving sociocultural entities as diverse as news articles, restaurant menus, and commercial products of every stripe. In this sense, these recommendations are kinds of interpersonal communication, although displayed in and through mediated platforms. They are and should be studied as representations of the world of social influence through pseudointerpersonal forms (Shi, Messaris, & Cappella, 2014; Walther & Jang, 2012). This space of horizontal communication representing views by elements of the public has already been widely studied within the communication discipline (de Zúñiga, Veenstra, Vraga, & Shah, 2010; Marichal, 2012) and even extended to consider the possibility that this arena has the potential to be a Habermasian deliberative space (Habermas, 1989) rather than an uncivil free-for-all (Dahlberg, 2011).

There is another form of recommendation system whose connection to the world of interpersonal communication is less obvious and that offers the kind of recommendations provided through algorithmic computations such as collaborative filtering. The most well-known examples of such approaches are the recommendations provided by commercial vendors such as Netflix and Amazon. But what is collaborative filtering and what does it have to do with interpersonal communication?

Collaborative filtering is one element of prediction used to design recommendation engines whose purpose is to offer selections of future items based on a person's past selections and on selections made by others who are computationally similar—that is, are structurally similar in some way to the targeted individual. ("Clones" is descriptive but too strong.) This process is a surrogate for direct social influence processes in which individuals actually recommend similar choices by their own choices, but do not advocate their recommendations explicitly—only implicitly recommend them through similar patterns of prior choices. Many commercial vendors, including some news sources, use such recommendation algorithms to offer suggested products and content.

Recommendation systems all have the same goal: to estimate ratings of items in a universe by a target person, even though that item has not been seen by the target user. The research on effective recommendation machines in commercial applications has been extensive; Adomavicius and Tuzhilin, (2005), for example, classify these machines into three categories: those based on content of the items being recommended, those based on collaborative recommendations (also known as social filtering or ratings), and hybrid or combined methods using both types of data. In content-based recommendation, a user will be recommended items similar to the ones preferred in the past. The value of a new item for a user is estimated based on the value of items of similar content as assigned previously by the same user. The items recommended are typically those that have a high degree of similarity to past preferences. In collaborative recommendation (e.g., ratings); items recommended are those that people with similar tastes and preferences selected in the past. The value of a new item is estimated based on the values assigned to that item by other raters who are similar to the user. For example, in order to recommend a message to a user, a collaborative recommendation system tries to find the "peers", i.e., those other users that have similar preferences. Only the messages preferred by the user's "peers" would be recommended.

Collaborative recommendation methods collect ratings of artifacts from many individuals, and use nearest-neighbor techniques to make recommendations to a user concerning new artifacts. The recommender system compares the user's ratings to those of other users, finds the "most similar" users based on some criterion of similarity, and recommends items that similar users have preferred.

The types of collaborative recommendation fall into two broad groups: memory- and model-based. Memory-based collaborative algorithms develop a set of recommendations based on the ratings of other users. "Other users" can include all other users or the set of other users who are most similar to the target user in their pattern of previous ratings. Scholars have proposed a variety of measures of similarity that include different clustering procedures for users, weighting of individual ratings, and adjustments for variation in rating scale use, among many other possibilities (Adomavicius & Tuzhilin, 2005). Model-based collaborative recommendation is based on various approaches including machine learning, Bayesian, linear regression, and others. While commercial practices widely deploy collaborative filtering approaches, generating much research, their application to problems in communication is just beginning (Cappella, Yang, & Lee, 2015). Our lab has successfully tested a message selection recommendation system for antismoking PSAs. The models have been able to add unique variance to preferences over and above that due to aggregate message content and targeted features (Kim, Yang, Kim, & Cappella, 2016).

The promise of collaborative recommender systems lies in the fact that items selected for a given user (e.g. messages for health, political or other goals) are a good fit to the user because a clone of the target person hs previously selected or positively rated them. The process of tailoring the choice is automatic once the prediction algorithm is developed. So, the process optimizes the effectiveness of tailored choice with the efficiency of automatic selection from a pool of choices. The collaborative filtering process takes the place of personalized recommendations delivered by peers with similar tastes (see Kreps, this issue).

Miller and Steinberg's (1975) approach to interpersonal communication highlighted the importance of the tailoring of information in unique ways to those with whom personal relationships developed. The person-specific nature of communication made trust (and its loss) possible. Subsequent work in tailored communication (e.g. Kreuter et al., 2013) made clear that tailoring made for more effective forms of social influence. But tailoring is intensive in terms of information gathering and requires the cooperation of the person targeted in tailoring. It is very effective but costly and inefficient.

Algorithm-based recommendation systems are in some senses substitutes for the information-gathering and fitting that takes place in more traditional interpersonal tailoring allowing the selection of information well-suited to the target. The fact that a version of tailoring is able to be deployed automatically (through algorithm-based selections) means that the effectiveness of influence delivered at the interpersonal level using surrogates of person-specific selection can be accomplished efficiently with much larger audiences. The ideal of a mass distribution of person-specific choices can at least be entertained.

Vector 4: How Does the Content of Interpersonal Communication Change as it Moves From Traditional Personal Exchanges to Broader Audiences?

Social media messages can be narrowcast or broadcast

Social media allow options for messaging that are narrowly targeted or broadly targeted even on the same platform. In the presocial media era, interpersonal communication was to a well-defined, often well-known specific other whose needs, like and dislikes, and other idiosyncracies were clear to the communicator. As interpersonal communication has migrated into the social media sphere, targets can be more amorphous: sometimes quite specific and personal and at other times quite extensive. "Friends" on social media can be more broadly defined than friends presocial media.

Effective and ineffective messaging choices depend on appropriate targeting. Failing to take the distinction into account can create unhappy results interpersonally and at macrosocial levels as well. The distinction also affects the decision of a potential sender to transmit or retransmit information. In this sense, what has traditionally been an interpersonal decision becomes a decision with implications for larger-scale distribution (i.e., a mass communication decision).

We have all come to appreciate that the distinctions between mass and interpersonal communication are becoming increasingly blurred. Mass communication has always been associated with broadcasting—sending messages to a large number of undifferentiated individuals. Facebook and Twitter transmissions are mostly broadcasting. Interpersonal communication has always been associated with narrowcasting.

Berger (2014) examined the range of motivations that drive individuals to share information via word of mouth and, in a delightful bit of wordplay, word of mouse. He concluded that the motivations for information sharing included (1) self-enhancement, that is, the rewards of self-disclosure and sharing one's positive experiences; (2) emotion presenting intense positive or negative events that allow catharsis, disambiguation, dissonance reduction, empathy, and bonding; (3) utility leading to clarification of ambiguous events, simplification of complex events, enacting altruism and future quid pro quo; and (4) accessibility offering top-of-the-mind or salient events, ideas, and commonalities with others.

Berger's useful delineation of motivations for sharing information led to the speculation that people would communicate differently to broader and narrower audiences. This observation is certainly commonsensical but takes on greater urgency in the era of mediated social interaction (Barasch & Berger 2014). In a series of five studies, the authors showed that self-presentation of negative information was reduced in broadcasting versus narrowcasting conditions and that useful information was more likely to be shared with specific, narrow targets. This latter effect was ameliorated by increasing the psychological connection to the larger audience through name listing. Despite the analogue nature of the experimental tests, the research consistently shows that what is communicated depends on the size and closeness of the targeted audience.

In dissertation work, Kim (2015) obtained results on retransmission of health news stories consistent with Barasch and Berger's hypotheses but in more naturalistic contexts and with very different topics. Kim followed the retransmission of 760 health news stories from *New York Times* data of sending those stories via e-mail to targeted individuals (narrowcasting) versus posting them on social media sources such as Facebook and Twitter (broadcasting). He concluded that stories sent to specific sources (email targets) included more efficacy information and were coded as more useful than when the targets were broader. Stories with emotionally evocative information were more likely to go to social media than email targets as were stories with exemplars. Health news that was novel tended to be shared via e-mail while novel health stories were less likely to be posted on social media.

Together the results from Barasch and Berger and from Kim make clear that when accounting for the breadth of one's informational targets, the content of what is to be communicated as well as the motivations of the sender and target must come into play. What is not at issue is the notion that in the modern world of social media, interpersonal sharing of information is common and of potentially greater impact than the presocial media era. What exactly gets communicated and by whom, with what range of targets, is a crucial issue necessitating intensive research as we try to understand the effectiveness of social media narrowly or broadly cast. People seem to be conscious of the personalness of their communication even in the face of mediated platforms that mute the character and visibility of the target.

In the new media world where interpersonal communication contributes substantially to public information exposure and potential influence, what is communicated narrowly through platforms designating specific others and what is communicated broadly to a wider spectrum of others is more important than ever. As researchers such as Berger and Kim seek to understand the messages that circulate widely, with increased probability of sharing, the platforms that senders use and the motives driving them come into play.

Conclusion

The new world of computational communication science, social media, and big data are remaking what counts as interpersonal communication and social and normative influence. The size and dynamics of our interpersonal data allow us to evaluate or re-evaluate existing assumptions about connectivity, exposure, and social influence. New methods for modeling allow the testing of both persons and variables in more complex combinations than could have been imagined even a couple of decades ago. Tools such as collaborative filtering create pseudopeers providing advice to individuals that is tailored to their past choices. The interplay of media and interpersonal influence has changed at its core and continues to do so. These lines of research are the objects of inquiry in many different fields. Vibrant and relevant communication research needs to be open to these developments, because the discipline has studied these questions in some cases for decades. We will continue to contribute more complete and accurate answers by remaining open to the widest possible array of methods and models.

Acknowledgments

The author thanks the reviewers and editors for their substantial assistance in revisions. A portion of the research reported here was supported by an NIH grant to the author R01CA160226. The views expressed are those of the author alone.

References

Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering. 2005; 17(6):734–749. DOI: 10.1109/tkde.2005.99

Barberá P, Wang N, Bonneau R, Jost JT, Nagler J, Tucker J, González-Bailón S. The critical periphery in the growth of social protests. PloS ONE. 2015; 10(11):e0143611.doi: 10.1371/journal.pone.0143611 [PubMed: 26618352]

Barasch A, Berger J. Broadcasting and narrowcasting: How audience size affects what people share. Journal of Marketing Research. 2014; 51:286–299. DOI: 10.1509/jmr.13.0238

Berger J. Word of mouth and interpersonal communication: A review and directions for future research. Journal of Consumer Psychology. 2014; 24:586–607. DOI: 10.1016/j.jcps.2014.05.002

Cappella JN, Kim HS, Albarracín D. Selection and transmission processes for information in the emerging media environment: Psychological motives and message characteristics. Media Psychology. 2015; 18:396–424. DOI: 10.1080/15213269.2014.941112 [PubMed: 26190944]

Cappella JN, Yang S, Lee S. Constructing recommendation systems for effective health messages using content, collaborative, and hybrid algorithms. The ANNALS of the American Academy of Political and Social Science. 2015; 659(1):290–306. DOI: 10.1177/0002716215570573

Centola D. The spread of behavior in an online social network experiment. Science. 2010; 329(5996): 1194–1197. DOI: 10.1126/science.1185231 [PubMed: 20813952]

Centola D. The social origins of networks and diffusion. American Journal of Sociology. 2015; 120:1295–1338. DOI: 10.1086/681275

Centola D, Macy M. Complex contagions and the weakness of long ties. American journal of Sociology. 2007; 113:702–734.

Centola D, Willer R, Macy M. The emperor's dilemma: A computational model of self-enforcing norms. American Journal of Sociology. 2005; 110:1009–1040. DOI: 10.1086/427321

Dahlberg L. Computer-mediated communication and the public sphere: A critical analysis. Journal of Computer-Mediated Communication. 2011; 7(1)doi: 10.1111/j.1083-6101.2001.tb00137.x

de Zúñiga HG, Veenstra A, Vraga E, Shah D. Digital democracy: Reimagining pathways to political participation. Journal of Information Technology & Politics. 2010; 7:36–51. DOI: 10.1080/19331680903316742

Goel, S., Watts, DJ., Goldstein, DG. The structure of online diffusion networks. Proceedings of the 13th ACM Conference on Electronic Commerce; New York, NY: ACM; 2012 Jun. p. 623-638.

González-Bailón S, Wang N. Networked discontent: The anatomy of protest campaigns in social media. Social Networks. 2016; 44:95–104. DOI: 10.1016/j.socnet.2015.07.003

Granovetter MS. The strength of weak ties. American Journal of Sociology. 1973; 78:1360–1380. DOI: 10.1086/225469

Habermas, J. The structural transformation of the public sphere: An inquiry into a category of bourgeois society. Cambridge, MA: MIT Press; 1989.

Hornik, R., editor. Public health communication: Evidence for behavior change. Mahwah, NJ: Lawrence Erlbaum Associates; 2002.

Japec L, Kreuter F, Berg M, Biemer P, Decker P, Lampe C, … Usher A. Big data in survey research: AAPOR task force report. Public Opinion Quarterly. 2015; 79:839–880. DOI: 10.1093/poq/nfv039

Katz, E., Lazarsfeld, P. Personal influence: The part played by people in the flow of mass communications. Glencoe, IL: Free Press; 1955.

Kim HS. Attracting views and going viral: How message features and news-sharing channels affect health news diffusion. Journal of Communication. 2015; 65:512–534. DOI: 10.1111/jcom.12160 [PubMed: 26441472]

Kim, HS., Kim, M., Yang, S., Cappella, JN. Assessing the effectiveness of recommendation algorithms for health message design: An experiment. Paper presented at the 2nd Annual International Conference on Computational Social Science; Evanston, IL. 2016 Jun.

Kreps GL. Online information and communication systems to enhance health outcomes through communication convergence. Human Communication Research. 2017; 43 XXXX. doi:XXXXXXXX.

Kreuter, MW., Farrell, DW., Olevitch, LR., Brennan, LK. Tailoring health messages: Customizing communication with computer technology. New York, NY: Routledge; 2013.

Kreuter MW, Wray RJ. Tailored and targeted health communication: Strategies for enhancing information relevance. American Journal of Health Behavior. 2003; 27:227–232. DOI: 10.5993/ajhb.27.1.s3.6

Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, … Jebara T. Life in the network: The coming age of computational social science. Science (New York, NY). 2009; 323(5915): 721.doi: 10.1126/science.1167742

Macy MW, Willer R. From factors to actors: Computational sociology and agent-based modeling. Annual Review of Sociology. 2002; :143–166. DOI: 10.1146/annurev.soc.28.110601.141117

Marichal, J. Facebook democracy: The architecture of disclosure and the threat to public life. Farnham, UK: Ashgate; 2012.

Miller, GR., Steinberg, M. Between people: A new analysis of interpersonal communication. Palo Alto, CA: Science Research Associates; 1975.

Neubaum G, Krämer NC. Opinion climates in social media: Blending mass and interpersonal communicaton. Human Communication Research. 2017; 43 XXX-XXX. doi:XXXXXXX.

Raudenbush, SW., Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2. Thousand Oaks, CA: Sage; 2002.

Rogers, EM. Diffusion of innovations. 5. New York, NY: Free Press; 2003.

Shah DV, Cappella JN, Neuman WR. Big data, digital media, and computational social science: Possibilities and perils. The ANNALS of the American Academy of Political and Social Science. 2015; 659(1):6–13. DOI: 10.1177/0002716215572084

Shi R, Messaris P, Cappella JN. Effects of online comments on smokers' perception of antismoking public service announcements. Journal of Computer-Mediated Communication. 2014; 19:975–990. DOI: 10.1111/jcc4.12057 [PubMed: 25561825]

Simmons, MP., Adamic, LA., Adar, E. Memes online: Extracted, subtracted, injected, and recollected. Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media; Palo Alto, CA: Association for the Advancement of Artificial Intelligence; 2011. p. 353-360.

Snyder, LB., Hamilton, MA. A meta-analysis of US health campaign effects on behavior: Emphasize enforcement, exposure, and new information, and beware the secular trend. In: Hornik, R., editor. Public health communication: Evidence for behavior change. Mahwah, NJ: Lawrence Erlbaum Associates; 2002. p. 357-383.

Walther JB, Jang JW. Communication processes in participatory websites. Journal of Computer-Mediated Communication. 2012; 18:2–15. DOI: 10.1111/j.1083-6101.2012.01592.x

Weng, L., Ratkiewicz, J., Perra, N., Gonçalves, B., Castillo, C., Bonchi, F., … Flammini, A. The role of information diffusion in the evolution of social networks. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2013. Retrieved from https://arxiv.org/pdf/1302.6276.pdf

Wilensky, U., Rand, W. An introduction to agent-based modeling. Boston, MA: MIT Press; in press