What Communication Scholars Write About:
An Analysis of 80 Years of Research in High-Impact Journals

ELISABETH GÜNTHER
University of Münster, Germany

EMESE DOMAHIDI
Leibniz-Institut für Wissensmedien, Germany

Research topics, as indicators of the profession's development, are central to the evaluation of academic practices in communication research. To investigate the main topics in our field, we trace the development of research topics since the 1930s by evaluating more than 15,000 articles from 19 academic journals based on an automated content analysis. Topic modeling reveals a high diversity from the early years on. Only a few journals show the tendency to focus on one topic only, whereas most outlets cover a broad variety and thus represent the field as a whole. Although our discipline is strongly interconnected with the changing media landscape, results show that communication research is characterized by high consistency. Although they have not provoked a revolutionary change, Internet and social media have become the most monitored media, parallel to—not displacing—classic media such as newspapers and TV.

Keywords: research topics, high-impact academic journals, quantitative analysis, automated content analysis, computational methods, topic modeling, Internet, social media

The theoretical and empirical investigation of research practices is an emerging topic in a variety of disciplines. This "metaknowledge" (Evans & Foster, 2011) strives to gain knowledge about knowledge to be able to identify the process of scientific production and to define academic gaps as well as future directions for research. Researchers' data-driven insights in a particular field have recently been enhanced by the availability of digital archives that document the scientific process (Evans & Foster, 2011) and by the ever-advancing development of methods in computer science that allows analysis of the vast amount of information stored in these archives (Grimmer, 2015; Wallach, 2016).

The perspective of metaknowledge research in communication studies is often directed at the history of the field (Katz, 1983; Löblich & Scheu, 2011; Meyen, 2012; Rosengren, 1983; Schramm, 1983) with a focus on cross-national as well as scientific–cultural differences (Koivisto & Thomas, 2008;

Elisabeth Günther: elisabeth.guenther@uni-muenster.de
Emese Domahidi: e.domahidi@iwm-tuebingen.de

Simonson & Park, 2016; Thussu, 2009), the rise and fall of interest in certain research topics (Cherry, 1957; Craig, 1993; Neuman & Guggenheim, 2011), the standardization and the use of certain (new) methods (Scharkow, 2013; Zamith & Lewis, 2015), and the identification of important manuscripts as well as the existence of cooperation between certain authors (Chang & Tai, 2005; Pasadeos, Renfro, & Hanily, 1999; Rice, Borgman, & Reeves, 1988). This self-monitoring of the discipline unfolds into two main research perspectives or methods. The historical perspective often uses narrative reviews or qualitative methods and focuses on the biography of specific persons (or institutes) and their views on the field (Katz, 1983; Löblich & Scheu, 2011; Meyen, 2012; Rogers, 1997; Rosengren, 1983; Schramm, 1983; Wiedemann & Meyen, 2016). Other works concentrate on the development in a specific area or in the whole field and prefer quantitative methods, which include reporting the frequency of certain articles or citation analysis to delineate the relationship between particular works or authors (Chang & Tai, 2005; Neuman & Guggenheim, 2011; Pasadeos & Renfro, 1992; Rauchfleisch, 2017; Rice, Chapin, Pressman, Park, & Funkhouser, 1996). Until today, however, a broad empirical analysis of the topics covered by high-impact publications in communication research over the years is missing. To expand the metaknowledge of our discipline in this respect, we investigated the emergence and development of research topics in high-impact research articles on communication science.

Three challenges become apparent when we consider publications and topics in communication research: (1) The boundaries of the field are blurry because communication research is interdisciplinary and open for a variety of topics; (2) the growing body of research output and the fragmentation of this output across a vaster number of research topics complicate an overview; and (3) new developments further push those boundaries, triggering more output and even higher fragmentation of the field. One of these developments goes back to the emergence of the Internet around 1991, specifically the WWW, which is widely believed to be a game-changer for communication studies (D'Urso, 2009; Leiner et al., 2009; Valkenburg & Peter, 2013; Vorderer & Kohring, 2013). Over the past 25 years, Internet use has become widespread among the global population and has nearly reached saturation in Western civilizations such as the United States (Parks, 2009; Perrin & Duggan, 2015). The increasing popularity of Internet usage has caused the proliferation of online applications and research associated with these applications (Lenhart, Purcell, Smith, & Zickuhr, 2010; for a review, see Rice & Fuller, 2013). Whereas some scholars expected unique effects, others found that the Internet is now so deeply embedded in our everyday activities that it might one day become barely visible (Parks, 2009). Considering the importance of computer-mediated communication (CMC)[1] as a potentially transforming development for communication research, we expand our research interest to examine its role for our field. In doing so, we take the bird's eye view on the development of our research field. Such a perspective enables us to discover broad research topics, to show how this relatively young discipline has transformed over time, and to highlight future directions for our field.

---

[1] In this study, we use the term *computer-mediated communication* to refer to the broad developments, which include communication on the Internet, in social media, human–computer interaction/communication, database search, work interfaces, and so on. Later, we focus on the Internet and social media in more detail.

Topics and High-Impact Publications in Communication Research

The outlets of a given field, namely its journals and other publications, are the venues to observe the field's changing academic practices, and are particularly important for identifying the "foci of scientific interest" (Merton, 1938, p. 397) as represented by research topics. Publications are essential for academic work and a central element in defining science itself (Kuhn, 1962/2012). Luhmann (1990) emphasizes that the production of knowledge through scholars has to be completed by publishing the results to enable accessibility of this knowledge and the accompanying "scientific objectivity" (p. 405). Even though the definition, function and benefits of so-called high-impact journals are controversial (e.g., Anseel, Duyck, De Baene, & Brysbaert, 2004; Garfield, 2006), this measure remains, across disciplines, the best indicator for the most influential scientific journals. Regarding the challenges for a comprehensive description of relevant journals, as outlined, we describe the field of communication research as represented by its high-impact journals and ask the following:

RQ1:   *How have high-impact journals in the field of communication research evolved over time?*

In answering our research question, we focus on journals' evolution in terms of the quantity of published articles, topical variety, and similarity in and between them. According to the research literature, the general number of journals and published articles has increased over time (De Solla Price, 1965; Schramm, 1983). In a special issue of the *Journal of Communication* on the topic "Communication as a Field of Study," Schramm (1983) described how the quantity of outlets grew from a handful of journals in the 1940s and 1950s to about 50 in 1983. We expect that this general development has further increased with the appearance of digital publishing outlets around the 1990s. Beyond this role as a new publication venue, CMC is itself a research topic. Regarding the quantity of CMC-related research for the years 1990–2006, Tomasello, Lee, and Baer (2010) confirmed that the diffusion of new media articles had reached a critical mass and that a core set of 14 communication journals had published research on CMC. Given that the Internet serves both as a new digital outlet for publications and as a new topic, we expected that the number of journals (H1.1) and journal articles (H1.2) would have increased (a) in general over time and (b) especially since the rise of research on CMC from the 1990s onward.

With more and more outlets, strategic leadership decisions might lead to an increasing specialization of academic journals. Thus, as more journals develop within a field, we expected that individual journals would become more consistent paradigmatically and less diverse, representing a specific field of research. This implies that individual journals demonstrate less topical variety over time. Parallel to this development, we expected that the topical diversity between journals would have increased over time, as the special focus of each individual journal would make it more distinguishable from other journals. This means that journals today would be less similar to each other with respect to published topics. Accordingly, we expected that the topical variety within individual journals (H1.3) and the topical similarity between journals (H1.4) would have decreased over time.

Communication scholars have often investigated the topics of our field when provoked by developments that were perceived as problematic or challenging. Being a relatively young and interdisciplinary field, the diversity and heterogeneity of communication research were perceived as

precarious. Berelson's 1959 article "The State of Communication Research in *Public Opinion Quarterly*" can be seen as one of the first contributions to a long-lasting discussion about the topics covered by communication research. In this work, Berelson identified four major approaches: "the political approach, represented by Lasswell; the sample survey approach, represented by Lazarsfeld; the small-groups approach, represented by Lewin; and the experimental approach, represented by Hovland" (p. 2). More than three decades later in 1983, Schramm described a wider variety of topics in communication research, such as "social change, business and industrial relations, political power and political organization, the stewardship of rich countries over the telecommunications they own, intercultural and international relations, preparation of young members of the society for adult roles, and countless others" (p. 16). As a broad overview on the topics covered over time is missing until today, we asked the following research question:

RQ2:   How have the topics in high-impact communication research journals evolved over time?

Besides the general development, we consider changes provoked by CMC. Recently, various scholars have found evidence of the increased importance of research related to CMC in communication research. Based on the *Journal of Broadcasting & Electronic Media*, Rice et al. (1996) revealed changing topics from 1956 to 1993, indicating an enhanced focus on new technologies. In a follow-up study, Rice (2005) analyzed session and paper titles as well as abstracts from the Association of Internet Researchers conferences in 2003 and 2004 and found two conceptual clusters: one that included more traditional social science research on specific online applications, and one that more specifically focused on the content and use of these applications. Kim and Weaver (2002) found 52 specific research subjects in more than 200 journals for 1996–2000 that they grouped into 12 broader categories. Their results indicated that most articles investigated general law and policy issues, followed by articles on uses and perceptions of the Internet. Cho and Khang (2006) confirmed this increase of studies on CMC and found the most important topic to be Internet usage and perception and attitude toward the Internet. Peng, Zhang, Zhong, and Zhu (2013) revealed four main research themes for the period between 2000 and 2009: e-health, e-business, e-society, and human–technology interactions. Peng and Wang (2013) demonstrated the importance of communication journals in the field of Internet-related research by comparing them with journals in seven other fields (business, economics/finance, education, information science, political science, psychology, and sociology) for the years 2000–10.

By examining existing research, we can state that the general growth of outlets (Schramm, 1983) has been accompanied by an increasing topical diversification and specialization (H2.1), (a) in general and (b) especially with the rise of research on CMC from the 1990s onward. We expected the subsequent decades would differ in terms of the most frequently covered topics (H2.2a), especially given that studies on CMC became more and more popular around the turn of the millennium (H2.2b). We expected this change would be twofold: First, we expected that this technological development would bring new research topics in its wake; their emergence would be evident starting around the year 2000 (H2.3). Second, as media are the focus of communication scholars, Internet and social media are important research topics in and of themselves. As such, we expected that the relevance of the Internet and social media would have increased from the 1990s onward, while the relevance of traditional media, such as TV and print, would have decreased (H2.4).

The Present Study

Whether or not the emergence of the Internet has led to an overall quantitative and/or qualitative change in communication research, however, remains an open question. Until today, an extensive study comparing articles with and without a focus on CMC in a broad sample over a long time period is missing. With this analysis, we aim to shed light on the processes of scientific production in our field, reveal research gaps, and identify future research directions. As such, we focus more broadly on the topics of the published works in communication research. This approach complements the research literature on the topic by revealing patterns of a changing research agenda associated with CMC.

Sample

To achieve a sample of the most influential journals in communication research, we relied on all International Communication Association (ICA) journals plus the 20 other highest-ranked outlets offered by the SCImago Journal & Country Rank for the keyword *communication*.[2] We acknowledge that by omitting journals by associations such as the Association for Education in Journalism and Mass Communication and the National Communication Association, as well as other journals from different regional associations, publications with specific importance to scholars of American, German, Spanish, or French origin might not be represented in this sample. However, the ICA is the most important international academic organization without specific thematic focus for communication scholars. Journals that are published on behalf of the ICA achieve a particularly high level of attention in the scientific community, as they are freely distributed to all members. Second, the different languages of journals maintained by the regional associations would have been a major challenge for our project. As such, we considered only ICA journals, aiming for a balanced view of communication research without preferring a certain subfield or any national preferences. However, all other journals could still enter our sample if they stood out because of a high impact factor and were published in English.

Our aim was to include as many of the selected journals in our sample as possible. Furthermore, we were interested in the main topics of the articles only. For our analysis, we accordingly relied on the abstracts (for the keyword search also including *metadata*) of all articles from these journals without a time restriction and as such followed a common approach in topic modeling (Blei, Ng, & Jordan, 2003; Griffiths & Steyvers, 2004; Wallach, 2016). To obtain this information, we used the EBSCOhost meta-database, which offers access to a wide range of databases and journals. Conducting the initial research on EBSCOhost for the journals of interest, we obtained 24,257 first hits from 19 journals.[3] The elicitation of the high-impact journals and the data-gathering procedure were conducted in July 2015. The journals included in our sample were *Communication Education*, *Communication Methods and Measures*, *Communication Research*, *Communication Review*, *Communication Theory*, *European Journal of*

---

[2] Retrieved from
http://www.scimagojr.com/journalrank.php?category=3315&area=0&year=2014&country=&order=sjr&min=0&min_type=cd&page=4

[3] The content of six originally selected journals was unfortunately not available to us based on EBSCOhost's access permissions for our home institutions.

Communication, *Group Processes & Intergroup Relations*, *Human Communication Research*, *The International Journal of Press/Politics*, *Journalism*, *Journalism Studies*, *Journal of Advertising*, *Journal of Communication*, *Journal of Computer-Mediated Communication*, *Media Psychology*, *New Media & Society*, *Political Communication*, *Public Opinion Quarterly*, and *Research on Language and Social Interaction*. EBSCOhost offers a complete data download[4] for all research results as an .xml file. After the download, we selected all 17,858 entries of document type "Article" (other entries referred to book reviews, editorials, etc.). We then excluded all entries with a missing abstract and/or title and received our final sample of 15,172 journal articles. The availability and total number of abstracts online depend on (a) the respective journal's terms of publication, such as the number of issues published per year and the number of articles in each issue; (b) the year when the journal was published for the first time; and (c) whether or not the journal offers digitalized versions of its older articles. Accordingly, the availability and number of articles per journal were subject to variation.

Analysis

To get an overview of the body of scientific research in our sample and to answer our research questions, we identified the topics present in all 15,172 articles based on their abstracts. Given the large sample size, we used topic modeling to fulfill this task. There are a few studies in communication research (e.g., Peng & Wang, 2013; Peng et al., 2013; Rauchfleisch, 2017) and in diverse other disciplines from history to computer linguistics that have used computational methods to investigate the emergence of topics in their field of study (e.g., Block & Newman, 2011; Hall, Jurafsky, & Manning, 2008; Wang, Bendle, Mai, & Cotte, 2015). The basis for this kind of analysis is a three-level hierarchical Bayes model, which infers the hidden topic structure from the observed document collection (Blei et al., 2003). A *topic* is technically defined as a distribution over words: For every word in every document, the topic contains the estimated probability that this word occurs when the given topic is covered. Documents are analogously represented as a distribution over topics: For every document in the collection, each topic's probability of being covered is estimated, meaning that a document is modeled to always represent a mix of topics. Due to the specifics of our sample, we decided on a correlated topic model (Blei & Lafferty, 2006), which accounts for the fact that the topics present in the document collection can be correlated. This assumption fits the empirical reality of communication research well, as certain topics occur together more often (e.g., "media violence" and "children and adolescents") than others (e.g., "health" and "agenda setting").

To apply quantitative analyses to natural language text, we prepared our data set by applying several common preprocessing steps (for an overview, see Manning, Raghavan, & Schütze, 2009). In topic modeling, the number of topics has to be determined a priori. To find the value for the respective parameter k that generates the best model fit for our data, we estimated 40 topic models from $k = 5$ to $k = 200$ and systematically compared them. We then estimated our correlated topic model with the resulting parameter value of $k = 145$ topics using the R *topicmodels* package. The estimated hyperparameter values for our model suggested that there were few strong topics per document (as opposed to a high number of equally distributed topics per abstract). We therefore decided to select the two topics with the highest probability for every document, with a minimum probability of .1 to avoid

---

[4] The availability of the download option is contingent on the researcher's access to EBSCOhost.

skewed results. We conducted several runs on our data set with the same preset of 145 topics. The set of topics discussed in this article was highly consistent over these runs, with only minor changes in order.

Given that both the number of journals in our sample and their yearly output of articles have increased over time, especially between the mid-1990s and the early 2000s (see Figure 1), the selection of labels was likely skewed toward research published in recent decades. To ensure a meaningful interpretation, we validated the labels of the inferred topics by manually checking publications from every decade that contained the topic with high probability.

At this point, it is important to note that topic modeling is not an "automation" of the topic category traditionally used in a manual content analysis. What is a topic, after all? Although the notion seems self-explaining at first sight, it is striking that there is no coherent definition in communication science (e.g., Edelstein, 1993). As a category in content analysis, it has furthermore long been known to yield notoriously bad reliability measures (Berelson, 1952), pointing to the fact that coders also find it difficult to agree on the seemingly intuitive concept. Modern topic modeling algorithms, on the other hand, have their roots in information science, specifically in information retrieval. Here, topics are modeled to solve synonymy and polysemy problems in search requests and provide users with context-sensitive results (Dumais, Furnas, Landauer, Deerwester, & Harshman, 1988). The underlying algorithmic topic concept is rooted in psycholinguistic research on text comprehension (Kintsch, 1988; Kintsch, McNamara, Dennis, & Landauer, 2011), which does not explicate a definition for the topic, but vaguely describes it as "what is being talked/written about" (Brown & Yule, 1983, p. 73). In the following, the term *topic* accordingly refers to the inferred categories as specified in the topic model. As a fully automated analysis strategy, this enabled us to explore the general themes that authors write about.

Results

What Communication Scholars Write About

Our resulting topic model contained 145 topics for 15,172 articles. For reasons of presentability, we limit the number of topics that we discuss in the following to the 15 most frequent topics in our sample (i.e., the core topics of communication research). Consequently, the basis for the following results is the subset of the 10,017 articles that include these core topics with high probability. The topics cover a wide range of subjects, providing a strong representation of the interdisciplinary character of the field of communication research (see Table 1).

For each topic, we manually picked a title that represented its content at an abstract level beyond the automatically generated topic labels. In the following, we illustrate this process by way of example. With labels such as *teacher* and *classroom*, the topic that was covered most frequently in our sample displayed obvious references to education research. This assumption was confirmed when looking at articles that cover Topic 1: Recent abstracts mostly described school or university samples and research on learning, whereas early articles from the 1940s and 1950s focused on educating the public about political questions and the ways and forms in which propaganda was employed to "educate" the public. Automatically generated labels for Topic 6, such as *Facebook* and *e-mail*, on the other hand, pointed to a

focus on new media technologies. This is a topic that is closely connected to the digital age and is mostly present in publications from the 2000s onwards, but the theme was also present in abstracts from earlier decades, although to a much lesser extent. For example, several articles existed on innovative computational research methods from the 1960s and 1970s. Another popular topic in communication research is Topic 8, which we named *media violence and media effects*. Both its labels and the manual inspection indicate a focus on violence displayed in the media, such as in TV and video games, and a discussion of both positive (enjoyment) and negative effects (aggression) of media.

Table 1. Core Topics of Communication Research in Our Sample
(Correlated Topic Model, n = 10,017).

Topic	n	Title	Label (after stemming)
1	1,647	Education	teacher, instruct, instructor, classroom, teach, verbal, credibl, humor, skill, speech
2	1,449	Marketing and PR	advertis, brand, recal, creativ, placement, persuas, sale, memori, copi, repetit
3	1,106	Media use	emot, narr, viewer, persuas, stereotyp, immigr, fiction, drama, cultiv, enjoy
4	793	Comparative research and media stereotypes	white, black, racial, ethnic, crime, african, welfar, victim, polic, asian
5	750	Health	risk, intervent, efficaci, patient, literaci, drug, cancer, client, advic, medic
6	692	New media	mobil, phone, capit, privaci, facebook, ict, turnout, cell, email, lifestyl
7	661	Survey and interview research	mail, household, letter, nonrespons, incent, request, refus, sponsorship, district, donat
8	631	Media violence and media effects	game, video, violenc, aggress, violent, arous, player, avatar, enjoy, gamer
9	589	Religion	religi, protest, arab, religion, activist, toler, cohort, authoritarian, marriag, terror
10	551	Family and development	children, famili, parent, child, mother, grade, development, father, preschool, elementari
11	526	Youth at risk	sexual, adolesc, peer, youth, smoke, girl, gai, psa, pornographi, cigarett
12	509	Trust, privacy, and credibility	trust, disclosur, decept, cmc, nonverb, violat, anonym, modal, sender, credibl
13	425	Language	languag, speech, speaker, semant, gestur, marker, repertoir, token, sentenc, grammat
14	287	Relationships	partner, satisfact, uncertainti, tactic, coupl, attach, hurt, romant, intimaci, compens
15	223	Crisis and conflict	isra, photograph, death, bbc, israel, palestinian, hyperlink, photographi, photo, narrat

Note. We considered two topics maximum per abstract.

The Evolution of Communication Journals

Our first research question focused on the evolution of high-impact journals in the field of communication research. Figure 1 displays the output of all 19 journals in our sample over time and clearly shows a general upward trend in the number of journals and overall journal output. However, journals differ both in terms of how far their archives date back in time and in terms of their yearly output. *Public Opinion Quarterly* provides abstracts from as early as 1937, which makes it the oldest journal in our sample. With a mean of 49.45 articles ($SD = 12.78$), it is also one of the journals with the highest number of articles published per year, along with *New Media & Society* ($M = 47.06$ articles, $SD = 21.05$) and the *Journal of Computer-Mediated Communication* ($M = 42.00$ articles, $SD = 17.68$). With $M = 50.73$ articles ($SD = 10.46$), *Journalism Studies* tops the list. The overall yearly output per journal averages $M = 29.95$ articles ($SD = 12.61$), with a downward trend from $M = 42.17$ articles ($SD = 23.92$) for 1980 and $M = 30.75$ articles ($SD = 11.23$) for 1990 to a low for the year 1999 ($M = 22.00$ articles, $SD = 7.90$). A plausible explanation for the decreasing number of articles per year is that they simultaneously increased in length, leading to the publication of fewer but longer articles per journal issue. With the rise of the Internet, journals have obtained new possibilities for publication, a development we see mirrored in the recovery of yearly journal output to $M = 37.83$ articles ($SD = 17.92$) in 2014. Notably, some journals seem to make more use of the unlimited publication space online or follow other strategies that lead to a diversification of publication standards. The overall yearly output has nonetheless been steadily increasing, as many new journals have been launched over the past few decades. Research on CMC has played a big role in this development and is specifically targeted by some of the new outlets (e.g., *Journal of Computer-Mediated Communication* in 1995 and *New Media & Society* in 1999).
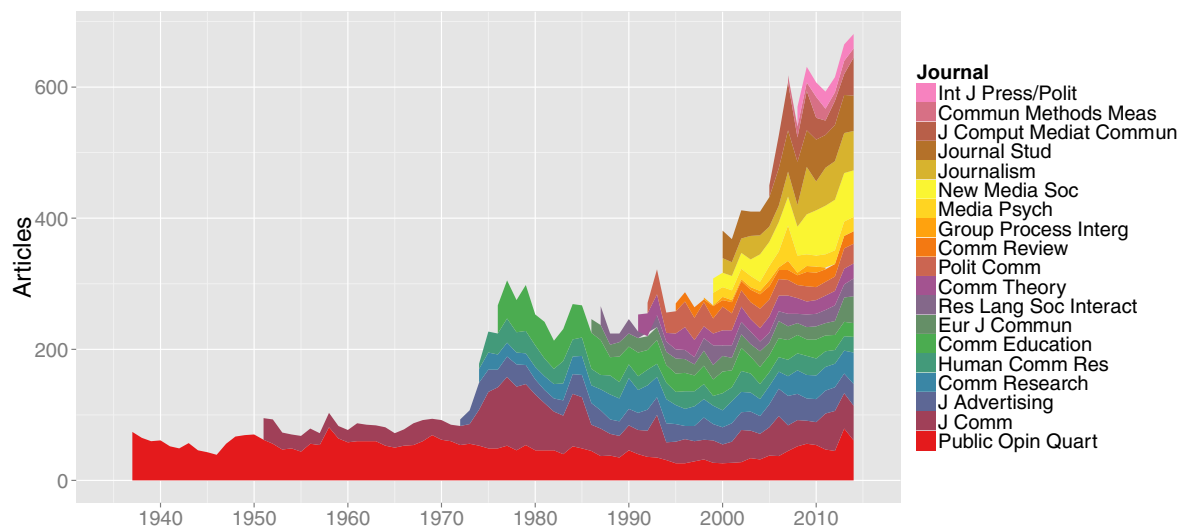

Figure 1. Yearly journal output over time
(including publications with missing abstract, $N = 17,858$).

As both the number of journals and the overall quantity of journal articles have increased over time, both H1.1a and H1.2a are supported. This development coincides with the rise of research on CMC beginning in the 1990s: In the 1970s, our sample included six journals that produced 1,107 articles. In the 1980s, eight journals published 1,548 articles. The number of journals and articles increases in the 1990s, the beginning of the Internet age, during which 14 journals ran a total of 2,341 articles. This is again outweighed by 16 journals and 4,727 articles in the 2000s; therefore, H1.1b and H1.2b are supported.

As the journals in our sample increased both in number and output volume, we now examine whether these changes were mirrored in the diversity of the content of their publications. To do this, we evaluated the results of our topic model at the journal level and determined each publication's thematic profile. Figure 2 displays these profiles as colorful barcodes, revealing changes over time (horizontal axis) and differences in thematic scope (vertical axis).
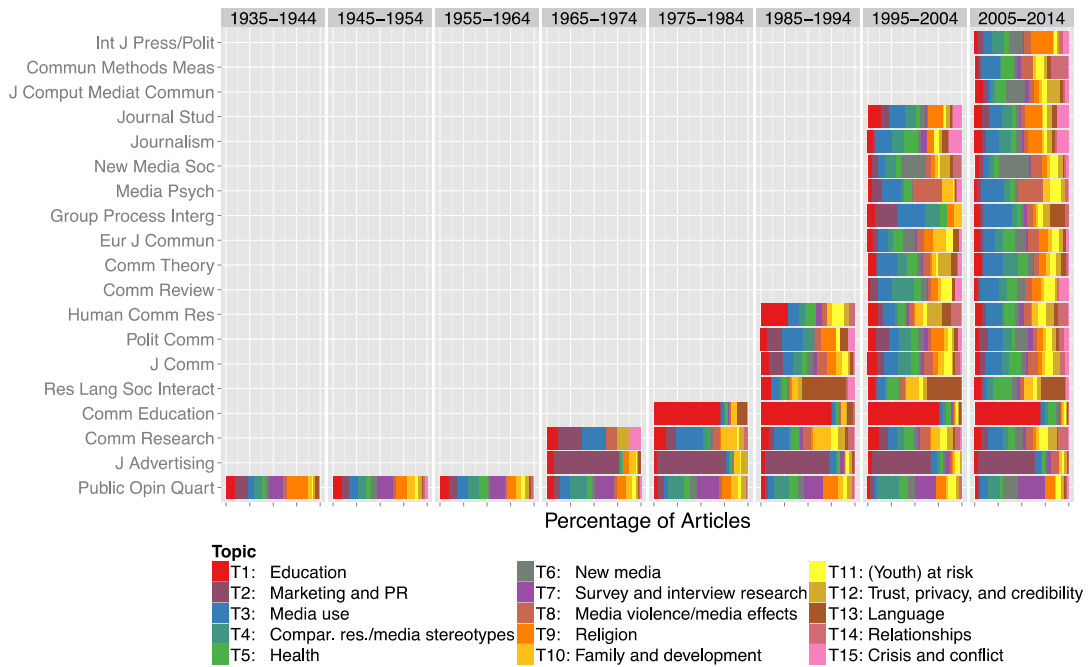

Figure 2. Topical variety per journal and decade
(core topics, n = 10,017, maximum of two topics per abstract).

Over time, most journals in our sample show consistent thematic profiles that are subject to only marginal shifts. The barcodes for *Public Opinion Quarterly*, for example, contain a similar range of colors, with the proportion of red (Topic 1 [education]) slightly fading over the decades. For the *Journal of Computer-Mediated Communication*, a slight focus on Topic 6 (new media) becomes apparent in the color-coding. Overall, however, results show neither an overall decline in thematic variety nor an overall increase in thematic differentiation over time. Journals generally feature both consistent and rather balanced thematic profiles that, with few exceptions (notably *Communication Education* and the *Journal of Advertising*), have

only nuanced differences from each other. Most journals in our sample are not focused on one specific topic, but rather tend to represent the core of the topical variety of the field itself. This pattern remains consistent over time. Consequently, we reject H1.3 and H1.4.

The Evolution of Topics in Communication Research

To answer our second research question, we now examine the topics for all abstracts in our sample. Looking at the development in absolute frequencies (see Figure 3), it is striking how little the spectrum of the most covered topics in our sample has been affected by time. Although the roots of most of today's topics go back to the 1930s, the core set has been steady especially since the mid-1970s, a period during which a growing number of journals has been launched. Although a discipline strongly characterized by the properties of the surrounding media landscape, communication research features a distinct set of core topics that are adaptable and permeable enough to be able to integrate new technologies. With no signs of increasing fragmentation of what communication scholars write about, H2.1a and H2.1b are rejected.
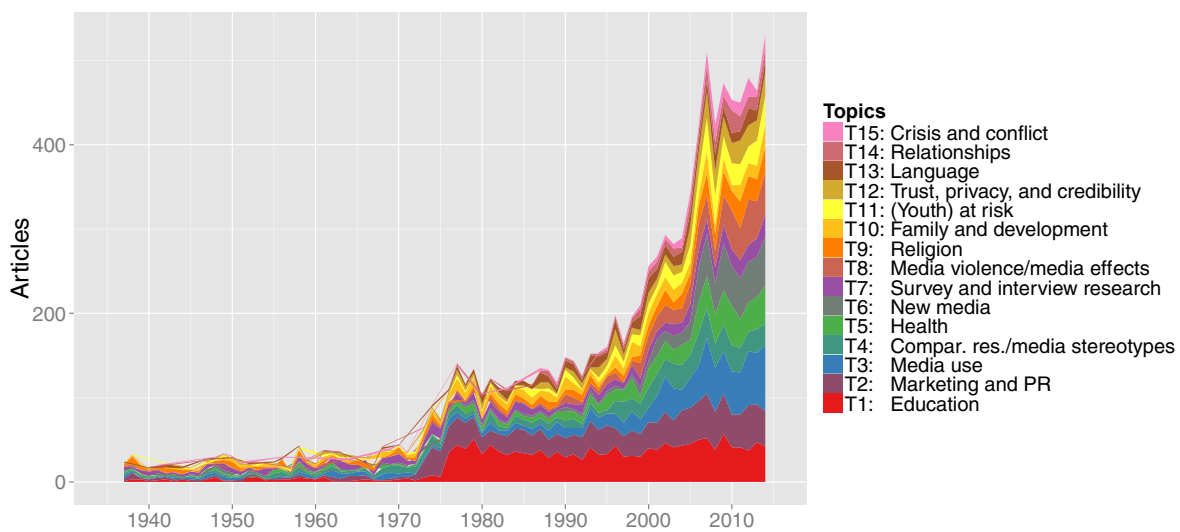

Figure 3. Topic attention over time (core topics, n = 10,017, maximum of two topics per abstract).

Regarding H2.2, there is some variation in the most popular topics per decade, which is most visible in the early years because of the lower yearly output. Until the 1970s, *Public Opinion Quarterly* was the only journal in our sample to represent communication research; for the following decades, our results show that topic attention is, among others, dependent on the changing composition of the sample: With the launch of the *Journal of Advertising* in 1972, for example, Topic 2 (marketing and PR) became the most popular topic in the subsample for that decade (78 in 1965–74, from 26 in 1955–64). Analogously, Topic 1 (education) increased in volume (33 in 1965–74 to 356 in 1975–84) after the introduction of the *Journal of Education*. Recent years have seen a tremendous growth in volume for articles covering Topic 6

(new media), with only 33 abstracts in 1985–94 to 97 abstracts in 1995–2004, and, with an almost fivefold increase, 457 abstracts in 2005–14. The two other topics that have gained most in popularity in the digital age are Topic 5 (health) and Topic 8 (media violence and media effects). Whereas research on Topic 5 did not stand out in the late 1970s/early 1980s, it has since about doubled every 10 years (72 in 1985–94, 184 in 1995–2004, 364 in 2005–14). Similarly, Topic 8 was hardly visible before 1975 (<10 per 10 years), but is the fifth most frequent topic in the 2000s and recently even received slightly more attention than Topic 5 (40 in 1985–94, 130 in 1995–2004, 370 in 2005–14). Consequently, H2.2a and H2.2b are supported.

To sum up our results, the range of topics covered in our sample has been fairly consistent over time, with results indicating that some topics have increased more in popularity in the digital age than others. With no indication for a disruption of the range of core topics covered by communication scholars that might be traced to the emergence of CMC, however, we reject H2.3.

For our next step, we conducted a keyword search (including articles' abstract, title, and keywords) for *Internet* and *social media*[5] to identify references to these media within the articles, as compared with traditional media such as newspapers, TV, or radio.[6] In the overall sample, the proportion of those articles was 12.44% ($n = 3,672$).

Figure 4 shows how much the Internet and social media have shaped communication research as represented by our sample in recent years. For 2014, 25.19% of all articles were matched in the keyword search for either Internet- ($n = 119$) or social media–related expressions ($n = 114$). Although the Internet has only been accessible to the general public since the 1990s, in sum, the Internet and social media are mentioned in more abstracts than any of the other media (TV: $n = 1,589$; Internet [without social media]: $n = 1,419$; newspaper: $n = 1,378$; personal communication: $n = 650$; book: $n = 510$; social media: $n = 477$; radio: $n = 368$; movie: $n = 258$). As the overall number of articles has also increased, it is not surprising that the research per medium has, overall, increased as well. In absolute numbers, the immense growth of Internet-related research is accordingly not accompanied by a displacement of other media in our sample. Looking at the development percentage-wise, however, it is significant how much importance Internet-related research has gained over the past two decades (for 2005–14: $n = 547$ for TV; $n = 1,143$ for Internet [without social media]; $n = 471$ for social media; see Table 2 for detailed information). We therefore maintain H2.4.

---

[5]Regular expression for Internet keywords: "internet|WWW|^web[^a-z]|[^a-z]online|digital|cyber[^n]|video\\sgam|computer\\sgam|console\\sgam|browser\\sgam|gamer|gaming|forum|fora|bulletin\\sboard|newsgroup|portal|message\\sboard|MUD[^a-z]|Usenet"; social media: "facebook|google\\+|^renren|^weibo|linkedin|^xing|researchgate|myspace|youtube|vimeo|last\\.fm|spotify|flickr|blog|tumblr|twitter|tweet|reddit|online\\sforum|discussion\\sforum|^chat|whatsapp|snapchat|wechat|weixin|social\\snetwork\\ssite|^SNS$|^OSN$|social\\smedia." In cases in which the social media RegEx was tested positive, "Internet" was recoded to a nonmatch for better differentiation.

[6]Regular expression for personal communication: "face[^a-z]to[^a-z]face|inter[^a-z]personal|speech"; book: "book|monograph"; radio: "radio[^a-z]"; movie: "film|theatre|movie"; TV: "[^a-z]tv|television|broadcasting"; newspaper/magazine: "newspaper|magazine|print[^a-z]|gazette|freesheet."
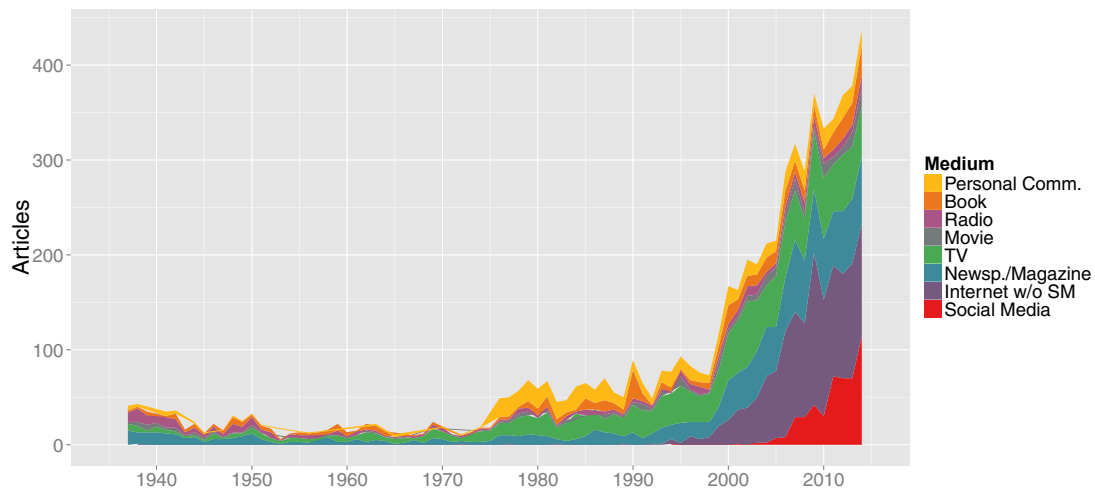
Figure 4. Media mentioned in abstract
*(*n* = 5,403, multiple media per abstract except for Internet and social media).*

Table 2. Media Mentioned in Abstract.

Medium, n (%)	1935–1944	1945–1954	1955–1964	1965–1974	1975–1984	1985–1994	1995–2004	2005–1914
Personal communication	18 (6.84)	7 (3.57)	7 (4.19)	6 (4.29)	171 (32.82)	134 (20.43)	132 (9.64)	175 (5.24)
Book	22 (8.37)	34 (17.35)	41 (24.55)	8 (5.71)	46 (8.83)	93 (14.18)	96 (7.01)	170 (5.09)
Radio	72 (27.38)	47 (23.98)	26 (15.57)	19 (13.57)	26 (4.99)	29 (4.42)	65 (4.75)	84 (2.52)
Movie	26 (9.89)	12 (6.12)	3 (1.8)	4 (2.86)	21 (4.03)	29 (4.42)	55 (4.02)	108 (3.24)
TV	33 (12.55)	40 (20.41)	46 (27.54)	67 (47.86)	178 (34.17)	241 (36.74)	437 (31.92)	547 (16.39)
Newspaper and magazine	91 (34.6)	56 (28.57)	44 (26.35)	35 (25.0)	77 (14.78)	121 (18.45)	315 (23.01)	639 (19.15)
Internet without social media	1 (0.38)	–	–	1 (0.71)	2 (0.38)	8 (1.22)	264 (19.28)	1,143 (34.25)
Social media	–	–	–	–	–	1 (0.15)	5 (0.37)	471 (14.11)

Note. *n* = 5,403, multiple media per abstract except for Internet and social media.

Discussion and Conclusion

In the present study, we aimed to deliver the first broad overview on the emergence and development of topics in communication research, with a focus on the impact of research on CMC. Our automated content analysis of 15,172 articles from 19 high-impact journals revealed surprising results: In general, we observed an increasing number of outlets, but no increasing topical fragmentation in our sample. The latter renders false the common beliefs about the development of the discipline (Schramm, 1983) and underlines the view that the research on CMC is integrated into preexisting research repertoires instead of radically transforming the field (Parks, 2009). Qualitative changes such as the emergence of new topics were not (yet) visible in our sample, but we did observe tremendous changes in the importance of certain topics (e.g., Topic 6 [new media]) over the past two decades. Moreover, we found that Internet and social media have become the most important media for investigation, parallel to (rather than displacing) classic media such as TV or newspapers. Thus, the present study unfolds a new perspective on the field: From early on, research in high-impact communication journals has at its core been characterized by a great topical variety that, for the most part, has changed only with regard to the quantity of the output.

The current study offers surprising insights into the development of the field's journals and topics, but there are several limitations to consider. Even though our sample is large, it does not consist of all potential outlets for communication scholars, which might lead to a bias. Although high-impact journals are some of the main outlets in the field, the relevance of other forms of publications is not to be underestimated. Monographs and collective volumes, for example, might serve to develop new and more complex theories that cannot be presented within the confines of a mostly standardized journal article. Moreover, some subfields may be represented more strongly in high-impact international journals than others, which might also lead to biased results of the most important topics. We also note that high-impact journals are defined as being of high impact at the time of data gathering, rather than eliciting journals with highest impact for each given year in the sample. Due to the diversity of languages and to avoid an even stronger bias for research published in English, we furthermore did not consider the journals of regional associations. Focusing on the local perspectives might be a fruitful research venue for future studies. At the same time, the analysis of topics in particular subfields, such as journalism studies or political communication, could reveal a more detailed view on the field's topical emphasis. In addition, our aim to provide a broad overview unfortunately brings up the challenge that it is impossible to represent every single journal accordingly. Although we chose certain examples to refer to specific developments in the field (e.g., for *Public Opinion Quarterly*), we could not equally highlight others (e.g., *The International Communication Gazette*). In this article, the focus was on the broad, overall themes. A meaningful follow-up on this work would be to zoom in on specific subsets, for example, a more detailed look into shorter time periods.

We argue that the journals selected are quite general and well suited to represent the field of communication research as a whole: If the study were to rely on an even bigger sample, it would necessarily also include journals with a more specific focus, making distortions of topic modeling results likely. From the results displayed in Figure 2, we see that, for two journals (*Communication Education* and the *Journal of Advertising*), their more narrow focus clearly shows up. Although the fields of

communication education and advertising are wide enough to be a legitimate part of the sample, and both journals also contain other topics, the results demonstrate that inclusion of a larger quantity of focused journals would mean that the results would tend to show only few topics per journal, which was not the goal of this research.

The last limitation concerning our sample lies in the fact that, unfortunately, not all journal archives were fully digitized at the time of data collection. This notably concerns early publications of the *Journal of Communication*, meaning that we have a relatively large number of missing abstracts in our data set for its early years (at the time of our data collection, the *Journal of Communication* archive was digitally available only from 1988 onwards).

The keyword search we conducted to identify the media being investigated also came at a price, as some media can be better described by keywords than others; results might therefore be biased and need to be interpreted carefully. All of our analyses are based on abstracts, which are understood to be a summary of the study's most important characteristics. We acknowledge that this might not always be the case. To ensure that the results are generalizable beyond our sample, it is advisable to replicate this study by analyzing other journals in communication research. Lastly, although our conclusion is seemingly common-sensual, we do not trace any causal direction, that is, whether gradual growth of a certain topic has led to creation of a certain journal or vice versa. Interviews with editors of the respective journals could shed light on this question.

Over the past 25 years, research on CMC has matured to a "primary area of communication studies" (D'Urso, 2009, p. 708). Our results show that CMC has indeed changed the field of communication research, not just in terms of quantity of journals and journal output. The field's main focus today is on researching the Internet and social media next to older media such as newspapers and TV. Although research on CMC has not changed scholars' topical focus in its core, the interest in certain topics has increased tremendously. What do the patterns uncovered in our sample mean for the landscape of communication research journals, the future of communication research, and especially for forthcoming research on CMC?

First, our results suggest that new research is promoted by both (a) organizational and (b) technological developments: Technological developments prompt organizational changes (e.g., the launch of a new journal), which in turn promote, and provide an outlet for, research on these technological developments. In our data, we see many examples of how decisions on an organizational level, such as the launch of new journals, have influenced the scientific output of our field. The case of the *Journal of Advertising* impressively illustrates the interconnectivity between output volume within the according topic (Topic 2 [marketing and PR]) and the availability of organizational structures for publication, from 6.48% of articles ($n = 26$) in 1955–64 to 21.03% ($n = 258$) in 1975–84; the *Journal of Advertising* was launched in 1972. Similarly, the launch of the *Journal of Computer-Mediated Communication* in 1995 was accompanied by a vast increase in publications on Topic 6 (new media). Next to academic publishers, the scientific organization itself offers important organizational structures for its members, in the case of the ICA and the European Communication Research and Education Association most importantly, by hosting an annual conference as a hub for scientific exchange and networking. Presentation slots are distributed

among the organizations' divisions, which means that their organizational structures both mirror and reinforce the research interests of the scientific community.

For the field of communication research, technological developments, such as the emergence of a new medium, have a particularly high relevance. In the specific case of research on CMC, the field has gained not only a new object of investigation, but also additional venues for the publication of research output. As we observed, research on CMC today is a "primary area of communication studies" (D'Urso, 2009, p. 708), in which we follow our fields' classic research questions, but introduce a special focus on CMC.

Second, we have learned from our analysis that although CMC has not necessarily disrupted the set of topics communication scholars write about, this does not mean that it has not had any substantial effect on our research. Although the core topics of our field seem surprisingly consistent, an in-depth analysis of specific subsets of the overall data set would be a fruitful follow-up project and would most likely reveal changes, both regarding a higher fluctuation of minor topics and regarding the way we approach them. With CMC, we have not only gained new media for communication but also fundamentally novel platforms for data access. The availability of digital trace data from millions of users and contents of huge databases might not have transformed our core research topics (yet), but it definitely has had a big impact on the prospects of our field. This is also because the same technological developments that made CMC possible have also provided new tools to analyze these huge amounts of data in almost real time (Günther & Quandt, 2016). Although communication science yet has to grasp the full meaning of "Big Data" (see, e.g., the special issue in the *Journal of Communication* from 2014, and in *Digital Journalism*, 2016), it has become increasingly obvious that computational methods are likely to turn out as a game-changer for communication research. In the light of these developments, scholars of CMC might develop new research questions and topics that were not possible before and thus might be a driving force in a new area of communication research. Our results show, however, that the integration of new research areas into the core of our field needs time. Considering the rapid increase of outlets and articles, it is crucial that these developments are accompanied by systematic reviews that extend our understanding of diverging research results and help us identify future research directions in communication research.

References

Anseel, F., Duyck, W., De Baene, W., & Brysbaert, M. (2004). Journal impact factors and self-citations: Implications for psychology journals. *American Psychologist, 59*(1), 49–51.

Berelson, B. (1952). *Content analysis in communication research*. Glencoe, IL: Free Press.

Berelson, B. (1959). The state of communication research. *Public Opinion Quarterly, 23*(1), 1–2. doi:10.1086/266840

Blei, D. M., & Lafferty, J. D. (2006). Correlated topic models. *Advances in Neural Information Processing Systems, 18*, 147–154.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022.

Block, S., & Newman, D. (2011). What, where, when, and sometimes why: Data mining two decades of women's history abstracts. *Journal of Women's History, 23*(1), 81–109. doi:10.1353/jowh.2011.0001

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge, UK: Cambridge University Press.

Chang, T., & Tai, Z. (2005). Mass communication research and the invisible college revisited: The changing landscape and emerging fronts in journalism-related studies. *Journalism & Mass Communication Quarterly, 82*(3), 672–694.

Cherry, C. (1957). *On human communication: A review, a survey, and a criticism*. Cambridge, MA: MIT Press.

Cho, C. H., & Khang, H. K. (2006). The state of Internet-related research in communications, marketing, and advertising: 1994–2003. *Journal of Advertising, 35*(3), 143–163. doi:10.2753/JOA0091-3367350309

Craig, R. T. (1993). Why are there so many communication theories? *Journal of Communication, 43*(3), 26–33. doi:10.1111/j.1460-2466.1993.tb01273.x

De Solla Price, D. J. (1965). The science of science. *Bulletin of the Atomic Scientists, 21*(8), 2–8.

Dumais, S. T., Furnas, G. W., Landauer, T. K., Deerwester, S., & Harshman, R. (1988). Using latent semantic analysis to improve access to textual information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 281–285). New York, NY: Association for Computing Machinery.

D'Urso, S. C. (2009). The past, present, and future of human communication and technology research: An introduction. *Journal of Computer-Mediated Communication, 14*(3), 708–713. doi:10.1111/j.1083-6101.2009.01459.x

Edelstein, A. S. (1993). Thinking about the criterion variable in agenda-setting research. *Journal of Communication, 43*(2), 85–99.

Evans, J. A., & Foster, J. G. (2011). Metaknowledge. *Science, 331*, 721–725. doi:10.1126/science.1201765

Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA, 295*(1), 90–93.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences, 101*(Suppl. 1), 5228–5235.

Grimmer, J. (2015). We are all social scientists now: How big data, machine learning, and causal inference work together. *PS: Political Science & Politics, 48*(1), 80–83. doi:10.1017/S1049096514001784

Günther, E., & Quandt, T. (2016). Word counts and topic models: An overview of automated text analysis methods for digital journalism research. *Digital Journalism, 4*(1), 75–88.

Hall, D., Jurafsky, D., & Manning, C. D. (2008). Studying the history of ideas using topic models. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp. 363–371). Stroudsburg, PA: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1613763

Katz, E. (1983). The return of the humanities and sociology. *Journal of Communication, 33*(3), 51–52. doi:10.1111/j.1460-2466.1983.tb02405.x

Kim, S. T., & Weaver, D. (2002). Communication research about the Internet: A thematic meta-analysis. *New Media & Society, 4*(4), 518–538.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182.

Kintsch, W., McNamara, D. S., Dennis, S., & Landauer, T. K. (2011). LSA and meaning: Theory and application. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 467–479). New York, NY: Routledge.

Koivisto, J., & Thomas, P. (2008). *Mapping communication and media research: Paradigms, institutions, challenges* (Research Reports 11/2008). Helsinki, Finland: Communication Research Center, University of Helsinki.

Kuhn, T. S. (2012). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press. (Original work published 1962)

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., . . . Wolff, S. (2009). A brief history of the Internet. *ACM SIGCOMM Computer Communication Review, 39*(5), 22–31.

Lenhart, A., Purcell, K., Smith, A., & Zickuhr, K. (2010). *Social media and mobile Internet use among teens and young adults: Millennials*. Retrieved from http://eric.ed.gov/?id=ED525056

Löblich, M., & Scheu, A. M. (2011). Writing the history of communication studies: A sociology of science approach. *Communication Theory, 21*(1), 1–22. doi:10.1111/j.1468-2885.2010.01373.x

Luhmann, N. (1990). *Die Wissenschaft der Gesellschaft* [The science of society]. Frankfurt am Main, Germany: Suhrkamp.

Manning, C. D., Raghavan, P., & Schütze, H. (2009). *Introduction to information retrieval*. Cambridge, UK: Cambridge University Press.

Merton, R. K. (1938). Science, technology and society in seventeenth-century England. *Osiris, 4*, 360–632.

Meyen, M. (2012). The founding parents of communication: 57 interviews with ICA fellows—An introduction. *International Journal of Communication, 6*, 1451–1459.

Neuman, W. R., & Guggenheim, L. (2011). The evolution of media effects theory: A six-stage model of cumulative research. *Communication Theory, 21*(2), 169–196. doi:10.1111/j.1468-2885.2011.01381.x

Parks, M. (2009). What will we study when the Internet disappears? *Journal of Computer-Mediated Communication, 14*(3), 724–729. doi:10.1111/j.1083-6101.2009.01462.x

Pasadeos, Y., & Renfro, B. (1992). A bibliometric analysis of public relations research. *Journal of Public Relations Research, 4*(3), 167–187.

Pasadeos, Y., Renfro, R. B., & Hanily, M. L. (1999). Influential authors and works of the public relations scholarly literature: A network of recent research. *Journal of Public Relations Research, 11*(1), 29–52. doi:10.1207/s1532754xjprr1101_02

Peng, T. Q., & Wang, Z. Z. (2013). Network closure, brokerage, and structural influence of journals: A longitudinal study of journal citation network in Internet research (2000–2010). *Scientometrics, 97*, 675–693. doi:10.1007/s11192-013-1012-x

Peng, T. Q., Zhang, L., Zhong, Z. J., & Zhu, J. J. H. (2013). Mapping the landscape of Internet studies: Text mining of social science journal articles 2000–2009. *New Media & Society, 15*(5), 644–664. doi:10.1177/1461444812462846

Perrin, A., & Duggan, M. (2015). *Americans' Internet access: 2000–2015*. Retrieved from http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015

Rauchfleisch, A. (2017). The public sphere as an essentially contested concept: A co-citation analysis of the last 20 years of public sphere research. *Communication and the Public, 2*(1), 3–18. doi:10.1177/2057047317691054

Rice, R. E. (2005). New media/Internet research topics of the Association of Internet Researchers. *Information Society, 21*(4), 285–299. doi:10.1080/01972240500189232

Rice, R. E., Borgman, C. L., & Reeves, B. (1988). Citation networks of communication journals, 1977–1985: Cliques and positions, citations made and citations received. *Human Communication Research, 15*(2), 256–283. doi:10.1111/j.1468-2958.1988.tb00184.x

Rice, R. E., Chapin, J., Pressman, R., Park, S., & Funkhouser, E. (1996). What's in a name? Bibliometric analysis of 40 years of the *Journal of Broadcasting (& Electronic Media)*. *Journal of Broadcasting & Electronic Media, 40*, 511–539. doi:10.1080/08838159609364373

Rice, R. E., & Fuller, R. P. (2013). Theoretical perspectives in the study of communication and the Internet. In W. Dutton (Ed.), *Oxford handbook of Internet studies* (pp. 353–377). Oxford, UK: Oxford University Press.

Rogers, E. M. (1997). *History of communication study: A biographical approach*. New York, NY: Free Press.

Rosengren, K. E. (1983). Communication research: One paradigm, or four? *Journal of Communication, 33*(3), 185–207. doi:10.1111/j.1460-2466.1983.tb02420.x

Scharkow, M. (2013). Thematic content analysis using supervised machine learning: An empirical evaluation using German online news. *Quality & Quantity, 47*(2), 761–773. doi:10.1007/s11135-011-9545-7

Schramm, W. (1983). The unique perspective of communication: A retrospective view. *Journal of Communication, 33*(3), 6–17. doi:10.1111/j.1460-2466.1983.tb02401.x

Simonson, P., & Park, D. W. (2016). *The international history of communication study*. New York, NY: Routledge.

Thussu, D. K. (2009). *Internationalizing media studies*. New York, NY: Routledge.

Tomasello, T. K., Lee, Y., & Baer, A. P. (2010). "New media" research publication trends and outlets in communication, 1990–2006. *New Media & Society, 12*(4), 531–548. doi:10.1177/1461444809342762

Valkenburg, P. M., & Peter, J. (2013). Communication research paradigms: Five challenges for the future of media-effects research. *International Journal of Communication, 7*, 197–215.

Vorderer, P., & Kohring, M. (2013). Permanently online: A challenge for media and communication research. *International Journal of Communication, 7*, 188–196.

Wallach, H. (2016). Computational social science: Toward a collaborative future. In R. M. Alvarez (Ed.), *Computational social science: Discovery and prediction* (pp. 307–316). Cambridge, UK: Cambridge University Press.

Wang, X. S., Bendle, N., Mai, F., & Cotte, J. (2015). The *Journal of Consumer Research* at forty: A historical analysis. *Journal of Consumer Research, 42*(1), 5–18. doi:10.1093/jcr/ucv009

Wiedemann, T., & Meyen, M. (2016). Internationalization through Americanization: The expansion of the International Communication Association's leadership to the world. *International Journal of Communication, 10*, 307–316.

Zamith, R., & Lewis, S. C. (2015). Content analysis and the algorithmic coder: What computational social science means for traditional modes of media analysis. *ANNALS of the American Academy of Political and Social Science, 659*(1), 307–318. doi:10.1177/0002716215570576