Article

Communication Research
2022, Vol. 49(5) 599–626

DOI: 10.1177/0093650220944790
journals.sagepub.com/home/crx
SAGE

The Highs in Communication Research: Research Topics With High Supply, High Popularity, and High Prestige in High-Impact Journals

Chung-hong Chan[1] and Christiane Grill[1]

Abstract

More and more scholarly attention is paid to dissecting discipline of communication research under the microscope thereby aiming at revealing foci of scientific interest. The lion's share of research has hereby focused either on the supply side of research examining what topics scholars write about or at the popularity side of research shedding light on what scientific publications receive the most citations. Building up on this, we argue that these research strands are inadequate to the task of exhaustively identifying foci of scientific interest. Tailoring for the fragmented topical landscape of communication research, we propose an integrative combination of three metrics: supply, popularity, and prestige of research topics. By means of topic modeling, citation counts and citation networks, our study showcases how our approach is able to reveal the intellectual architecture of our discipline in order to identify relevant paths for further scientific inquiry.

Keywords

research topics, foci of scientific interest, topic modeling, citation counts, citation networks, high-impact journals

> "We all study *topics* today"—John Durham Peters (1993).

Starting point of any scientific debate, research project, or academic publication is a "research topic." A research topic is the subject or issue a researcher is interested in (Allen, 2017), that is, it is the central theme she is writing/talking about (Brown &

[1]University of Mannheim, Germany

Corresponding Author:
Chung-hong Chan, Mannheim Centre for European Social Research, University of Mannheim, MZES A5, 6 Bauteil A, Mannheim 68131, Germany.
Email: Chung-hong.chan@mzes.uni-mannheim.de

Yule, 1983, p. 73). While scholars may differ in how they approach a research topic—for example, by applying different methodological tools and different theories—the research problem they are looking at, that is, the focus of their scientific interest (Merton, 1938), is their common, organizing principle that unites them. As research topics structure the scientific debate evolving around particular research problems or puzzles, these topics rather than theories or methodological traditions shape a scientific field in an unprecedented way. This is particularly true for the field of communication research as the field has no "theoretical, ontological, epistemological, or thematic center" (Waisbord, 2019, p. 75) and consequently fragmentation into topical research foci is inevitable (Waisbord, 2019). The relevance and structuring function of topical foci for communication science is also shown by the fact that publication outlets as well as professional communication association—such as the *International Communication Association* (ICA), the *National Communication Association* (NCA) in the United States, or the *German Communication Association* (DGPuK)—organize themselves around topical subfields; for example, health communication; political communication; or LGBTQ studies.[1]

Dissecting one's own scientific field of research into "foci of scientific interest" (Merton, 1938, p. 397) has turned to a thriving research field itself. The endeavor of identifying research hubs not only arises out of natural human interest but it is also existential in order to advance knowledge in the entire field. This desire for "metaknowledge" (Evans & Foster, 2011, p. 721)—the knowledge about scientific knowledge—strives to discover structures and developments in the scientific knowledge production and to gain a deeper understanding of the respective field of research. It allows a bird's eye view on the rise and fall of research topics as well as on emerging research trends (Mane & Börner, 2004). Identifying research foci within a scientific discipline helps to determine abundance and scarcity as well as areas of strength and weaknesses of scholarly knowledge (Kamhawi & Weaver, 2003). Ultimately, scholars may use such knowledge on the intellectual architecture of their discipline to identify relevant paths for further inquiry. Informed by a rich knowledge about the nature of research foci, such paths can account for the possibilities of a dialogue across disciplines, address concerns about growing fragmentation while still using the possibilities for specialization. Only when scholarship knows the topical map and network of research foci, as well as their temporal dynamics, academics can engage in a collective and fruitful dialogue within and across their disciplinary boundaries to advance scholarly knowledge to ultimately address today's most pressing societal and political challenges.

How to identify these foci of scientific interest? Two strands of research have developed. On one hand, measuring the topical foci in academic journals has attracted a fair share of scholarly attention: These studies set out to analyze what topics scholars write about; in other words, what topics the scientific community is supplied with (e.g., Borah, 2017; Günther & Domahidi, 2017). On the other hand, revealing citation clusters has frequently stimulated scientific interest as well: These studies aimed at identifying what scientific publications receive the most citations; in other words, are popular within the community (e.g., Neumann & Guggenheim, 2011; Rauchfleisch, 2017). These two strands of research have previously been integrated to study selected

subfields of communication research (e.g., Rauchfleisch, 2017; Rauchfleisch & Schäfer, 2018), but not yet the entire field. We argue that these two research strands by themselves are inadequate to the task of exhaustively identifying foci of scientific interest. To paint a complete picture of the "chaotic" landscape of communication research (Waisbord, 2019, p. 117), we argue that one must consider three aspects: namely, the supply, the popularity, and additionally the prestige of research topics. This study thus addresses this research gap by being the first that uses such an integrative, threefold measure to map the foci of scientific interest for the field of communication research as represented by its high-impact journals. In doing so, this study offers an encompassing description of the network of topical foci in communication research, not only cross-sectionally, but also applying a temporal perspective across the last 18 years. Ultimately, this work provides a nuanced inventory of the field thereby stimulating understanding but also an informed debate among communication scholars.

Supply of Research Topics

In order to study how different research topics prevail or relinquish in the communication literature, researchers conducted so-called (topical) trend analyses.[2] First analyses on topical trends in communication science relied on manual coding of extant scientific publications and were realized for the fields of journalism (Cooper et al., 1993) and communication theories (e.g., Borah, 2017; Paul et al., 2000). Over the years, topical trend analyses via quantitative content analyses were also conducted for mass communication (Kamhawi & Weaver, 2003), media and communication effects (e.g., Neumann & Guggenheim, 2011), crisis communication (e.g., Ha & Boynton, 2014), new media (e.g., Tomasello et al., 2010), and public relations (e.g., Ye & Ki, 2012). The vast majority of topical trend analyses using quantitative content analyses have been realized for internet studies (e.g., Borah, 2017; Tomasello, 2001; Zhang & Leung, 2015).

Even though these trend analyses provided novel insights for evaluating the topical dynamics within these research fields, these studies were not without limitations: These analyses were very specific in their choice of journals as well as their foci of study (e.g., social networking sites, emerging communication technologies). Furthermore, these analyses mostly covered a comparable short time frame from 5 to 9 years.[3] Additionally, the applied methodological design—that is, manual content analysis—was not without drawbacks. From a technical viewpoint, manual content analysis becomes a very tedious and costly task once a large text corpus needs to be evaluated (Antons et al., 2016). It probably was the bottleneck preventing researchers from exhaustively analyzing the entire field. Most of the papers cited above studied only either a sample of randomly chosen articles or a subfield of the entire field of communication research. These trends with a limited scope might be useful for a subset of the audience, but not useful to evaluate the general trend of the whole field.

In order to counteract the bottleneck created by manual content analysis for assessing research trends of the entire field, more recent studies used a computational, that

is, natural language processing (NLP), approach. Specifically, the *Latent Dirichlet allocation* (LDA) and the *Correlated Topic Model* (CTM) approaches have been very successfully applied to a diverse set of research areas and questions (e.g., Blei et al., 2003; Blei & Lafferty, 2007; Griffiths & Steyvers, 2004; Roberts et al., 2014). The advantages of a topic modeling approach generally are that this approach allows exhaustive analyses of a large text corpus.[4] As to topical trends in communication science using topic modeling, the work by Günther and Domahidi (2017) is the most extensive one thus far. Günther and Domahidi (2017) studied a corpus of 15,172 abstracts published in 19 communication journals over a period of 80 years. In their analysis, they computationally confirmed an observation already made by D'Urso (2009): Computer-mediated communication (CMC) is the fastest growing discipline in communication research.[5]

While we agree that the supply of research topics is an important parameter for quantifying the foci of scholarly interest, we consider it just one part of a multidimensional, multifaceted concept. For instance, a surge in scientific publications on a specific topic, which does not subsequently attract the attention of the scientific community (i.e., does not generate citations), falls short of truly reflecting scientific interest. Consequently, we argue that a research topic can only become a focus of research if it fulfills three conditions: first, scholars prominently write or talk about the topic (=the research topic is in supply); second, scientific publications on the topic are widely cited (=the research topic is popular); and third, these publications are embedded in large citation networks (=the research topic is prestigious). Ultimately, the combined analysis of supply (i.e., topical patterns), popularity (i.e., citation frequency), and prestige (i.e., citation networks) provides a more integrative and complete picture of the foci of scientific interest.

Popularity of Research Topics

The most convenient yardstick to evaluate the scholarly attention a piece of research receives is by the number of its citations (Bornmann et al., 2008). Citation counts are thereby seen as attractive and unobtrusive raw data to measure such attention (Bornmann & Daniel, 2008; Bornmann et al., 2008). Within this line of argumentation, citation counts have frequently been considered a measure of a publication's "popularity" (e.g., Franceschet, 2010; Zhou et al., 2012).[6] Building on this, we conceptualize the variance in the number of citations for a given topic as its popularity. While this variance might be indicative for the variance in research quality, we refrain from extrapolating high quality from the citation count.

The first who used citation counts to measure the popularity of scientific publications were Gross and Gross (1927). Bibliometricians have applied this conceptualization to quantify the popularity of journals by developing the so-called "impact factor." Similarly, communication researchers have used the citation count to determine the popularity of entire journals (e.g., Levine, 2010) but also of specific communication theories (e.g., Neumann & Guggenheim, 2011). The same concept has also been extended to study the popularity of different research fields (e.g., Oppenheim, 1995).

This methodological approach can potentially be useful for the popularity analysis of topically organized fields of communication science. Freelon (2020) recently argued that one should evaluate the popularity of a communication science paper by comparing its citations with the citations of topically similar papers published by peers. However, this task has not yet been tackled and as Freelon emphasizes the process of constructing a topically organized citation metric is "painstaking but not impossible" (Freelon, 2020, p. 428). In this article, we will provide first insights into such a "painstaking" undertaking.

That being said, the conceptualization of a citation-based popularity measure is not without drawbacks. Citation-based metrics can easily be manipulated. One way of citation manipulation is self-citation—even coercively—to promote certain publications (Wilhite & Fong, 2012). Consequently, an adequate measure of a citation-based popularity of research topics should be adjusted for self-citations. As an add-on, a comparison between the popularity of research topics with and without adjusting for self-citations can reveal a topic-specific vulnerability to self-citations (Ioannidis et al., 2019).

Prestige of Research Topics

While the popularity of a research topic is measured using citation-based metrics assuming that citations reflect a static linear process, one can additionally measure the prestige of a topic by understanding the practice of citing others as a networking process (Xhignesse & Osgood, 1967). By using citations, one generates links to the work of others. On the aggregate, networks of citations and thus links among different pieces naturally arise. Communication science was envisioned by the founders of the field as an interdisciplinary subject (see discussion in Zhu & Fu, 2019). Thus, citations naturally link across subfields. At the same time, the emergence of topical subfields in communication science has been warned as "balkanizing" (Berger, 1991) or "hyperspecializing" (Waisbord, 2019) the field. How to counteract these developments and to intersect the subfields has turned into a landmark question for the future of the field (Herbst, 2008).[7] Whether and to what extent communication science is characterized by fragmentation or intersection of topics can be established by analyzing citation networks. One attempt in this direction has recently been made by Song et al. (2020) who studies the links between subfields via a network of topic correlations (i.e., whether or not two topics are similar). While their results suggest that the thesis on the increasing fragmentation of the field is exaggerated, the authors also admit that to truly measure a possible fragmentation of the field they should have also examined the citations across research topics.

For the field of communication research, numerous studies traced citation networks between journals to display the dialogue between outlets and research fields as well as to evaluate the journals' influences in specific fields of research (e.g., Park & Leydesdorff, 2009; Peng & Wang, 2013; Rauchfleisch, 2017; Reardon & Rogers, 1988; So, 1998). Furthermore, scholars drew on citation analyses to evaluate social procedures. By employing an author co-citation analysis, Lin and Kaid (2000) revealed

artificial barriers among political communication science scholars with different backgrounds as well as a strong intellectual fragmentation of the field. The most comprehensive analysis of citation networks so far is provided by K. Lee et al. (2016). They used a topic model to extract topics from a large corpus of abstracts. Combining the topic model with citation information, they mapped the knowledge structure of the field by network analysis.

In conclusion, to tackle the task of finding the foci of scientific interest, we can utilize the citation network of topics to determine how knowledge from one research topic integrates with other research topics. Instead of the linear nature of the citation-based popularity measure which assumes that all citations are the same, citation network analysis allows to consider the information of "who cites whom" and thus ultimately measures the prestige of a research topic (Franceschet, 2010). The concept involves a recursive definition: Prestige of a cited item increases with the number of citing items but also with the prestige of the citing items in turn (Pinski & Narin, 1976). For example, Paper A and Paper B both have been cited 5 times. Therefore, the popularity of both papers is the same. However, Paper A has been cited by five doctoral theses which have not been cited yet. Instead, Paper B has been cited by five papers from Deen Freelon, Jürgen Habermas, Patricia Moy, Zizi Papacharissi, and Jennifer Stromer-Galley. These papers in turn have been cited countless times. Consequently, Paper B has a higher prestige than Paper A because the papers citing Paper B are more prestigious than the papers citing Paper A. Foci of scientific interest which have high prestige are thus research topics that provide knowledge to different subfields due to their frequent citations while these citing subfields in turn are also knowledge providers to other subfields.

Centrality measurements from social network analysis, such as PageRank[8] (Bollen et al., 2006), have been used to quantify the prestige of journals or researchers. Previous attempts to quantify the prestige of communication research have thus far only focused on the prestige of journals, individual scholars (Griffin et al., 2016), and theories (Chung et al., 2013), but not on topics.

Research Goals and Questions

This article aims at detecting the research foci of the field of communication research using the above mentioned three dimensions: supply, popularity, and prestige. In a first step, this work looks at the supply side of research topics. To that end, we analyze the topical landscape of articles published in the top high-impact communication research journals. This initial step aims at outlining how the topical foci of scientific publications have developed over time. In this first analytical step, we are interested in answering the following two research questions:

Research Question 1 (RQ1): What topical foci have the top high-impact journals in communication research set?

Research Question 2 (RQ2): How have these topics within communication science research developed over time?

In a subsequent step, our study examines how often these topics have been cited by the scientific community, thus enabling an integrative analysis. This analytical step sets out to reveal the extent to which these topics have been popular in the scientific community. We thereby strive to answer the following research question:

Research Question 3 (RQ3): What research topics in communication research have increased or decreased in popularity over time?

In a third and last step, to truly analyze foci of scientific interest, we also look at the prestige of communication research topics. In other words, we are not only interested in how often these topics are cited but also by whom they are cited. The following research question will thus be answered:

Research Question 4 (RQ4): What research topics in communication research have the highest prestige?

By jointly measuring scholarly interest via the topic of the piece of research as well as via the number of citations it has received thus far and its citation network, this study is truly able to measure the foci of scientific interest. Ultimately, our integrative view on supply, popularity, and prestige allows to track larger patterns and trends of communication research's foci as represented by its top high-impact journals. Our insights thus strive to initiate a fruitful dialogue within and beyond the boundaries of our discipline to identify future paths of inquiry, to counteract possible weaknesses in our scholarly knowledge, and to help setting the discipline's upcoming priorities.

Methodological Design

Selection of Journals and Collection of Abstracts

For the purpose of our study we compiled a comprehensive list of journals to be analyzed. Specifically, we selected the top five communication journals in terms of their journal impact factor from each edition of the Journal Citation Reports (JCR) from 2000 to 2017. All unique journals—in total 23 journals—were included as potential candidates for our analysis. By applying the definition of "communication research" provided by the journal of *Communication Research* itself, we excluded five journals.[9] Eventually, 18 journals remained in our sample for further analysis and they are: *Communication Monographs* (CM); *Communication Research* (CR); *Communication Theory* (CT); *Cyberpsychology, Behavior, and Social Networking* (CBSN); *Discourse & Society* (DS); *Human Communication Research* (HCR); *Health Communication* (HC); *Information, Communication & Society* (ICS); *Journal of Health Communication* (JHC); *Journal of Communication* (JC); *Journal of Computer-Mediated Communication* (JCMC); *Media Psychology* (MP); *New Media & Society* (NMS); *Political Communication* (PC); *Public Opinion Quarterly* (POQ); *Public Understanding of Science* (PUS); *Research on Language and Social Interaction* (RLSI); and *Technical Communication* (TC).

For each of the included 18 journals, a set of custom-built web scrapers based on the Selenium framework was specially developed to collect information of all articles published from 2000 to 2017. Our scrapers controlled an ordinary web browser (Mozilla Firefox) and emulated ordinary user interactions with the journals' websites. The following information was collected for each scientific publication: title, authors, year of publication, and abstract. The topics of research were assessed by drawing on the abstracts of these papers. The reasons for not using full papers are twofold: First, an abstract is considered a summary of a study's key features, in other words, an abstract contains the core message of a research paper. Thus, the abstract encompasses the essential research topics a full article is addressing. While this approach might not give an exhaustive picture on methodological approaches or theoretical backgrounds, it certainly captures the key research topics of a paper. Second, we decided to ensure compatibility and comparability with extant studies on topical patterns in communication science, which all based their analyses on abstracts too (e.g., Günther & Domahidi, 2017; Peng & Wang, 2013; Tomasello et al., 2010). Almost all of the publications to be included in our analysis did have abstracts.[10] We excluded all publications from our sample, which did not fall into the publication categories of research papers, review articles, or theory papers. In other words, book reviews and editorials were excluded from the sample. In total, our final sample consisted of 12,990 journal articles.

Topic Model

All included abstracts were used to train a CTM. To that end, we used the R package *stm* (Roberts et al., 2014). The advantage of using CTM instead of LDA is that the former assumes topics can be correlated with each other, while the latter assumes that topics are independent which is rather unrealistic as an article can consist of multiple intertwined topics. Consequently, the CTM approach has successfully been implemented in previous studies (Günther & Domahidi, 2017).

In essence, a trained CTM model can calculate the topic-membership probabilities for a set of discovered topics based on a natural language text (a vector of $\theta_t = {}_i$, for i = first to nth topic). We hereby followed the recommendations by Maier and colleagues (2018) to ensure interpretability of our topic model. Basically, we preprocessed our corpus of abstracts by conducting a lemmatization of all words, deleting common stop words, numbers, and symbols, and then trimming terms that occur in less than 0.5% and more than 99% of the abstracts and then the processed corpus was subjected to CTM models with 30, 40, 50, and 60 topics. We qualitatively evaluated the four solutions following Maier and colleagues (2018)[11] and confirmed that the 40-topic solution has the best interpretability. We labeled each of the 40 topics independently. These topic labels are presented in Table 1.

Topical Trends

Using the θ_t vector from the CTM model for each article, the historical trend of each topic can be traced by calculating the average $\theta_t = {}_i$ for the ith topic for papers

Table 1. Results of the CTM Model and Dominant Journal for Each Topic.

#	Topic name	Keywords	Dominant journal ($\overline{\theta_t}$)
13	Social media	twitter, facebook, blogs, snss, tweet, sns, blogging	JCMC (0.05)
28	Clinical communication	physician, literacy, diabetes, care, nurse, numeracy, provider	JHC (0.12)
2	Social network	capital, social, network, tie, collective, identity, movement	ICS (0.06)
23	Media effects	attitude, influence, variable, toward, predict, model, predictor	CR (0.12)
39	Health campaign 2	cancer, breast, vaccine, hpv, vaccination, risk, screen	JHC (0.08)
29	Interpersonal communication	self, disclosure, esteem, support, satisfaction, friendship, cope	HCR (0.06)
31	Health campaign 1	smoke, alcohol, drug, tobacco, drink, cigarette, food	JHC (0.07)
11	Persuasion	message, persuasive, organ, loss, donation, appeal, reactance	HC (0.07)
7	Health communication	hiv, aid, intervention, tailor, plan, program, mhealth	JHC (0.1)
16	Problematic ICT use	game, addiction, gaming, cyberbullying, gamers, player, gamble	CBSN (0.09)
40	Group-specific media and communication effects	adolescent, young, age, girl, peer, teen, adult	CBSN (0.03)
32	Journalism studies	news, medium, journalism, journalist, coverage, newspaper, journalistic	PC (0.12)
14	Political communication	election, candidate, vote, voter, presidential, partisan, poll	PC (0.18)
25	Media psychology	character, narrative, judgment, affective, involvement, identification, person	MP (0.14)
24	Online behaviors	online, privacy, credibility, offline, trust, shop, forum	JCMC (0.05)
30	Survey methodology	validity, item, scale, instrument, measure, reliability, correlation	POQ (0.05)
33	Science communication	science, scientific, scientist, religious, expert, religion, environmental	PUS (0.2)
37	Intercultural communication	immigrant, crisis, israeli, disaster, muslim, ethnic, immigration	DS (0.02)
19	Family	child, parent, mother, family, children, relational, parental	CM (0.1)
9	Information seeking	seek, information, source, group, member, decision, channel	JHC (0.04)
36	Deliberation	deliberation, government, deliberative, policy, opinion, democracy, governance	PUS (0.09)
8	Surveys	respondent, telephone, survey, nonresponse, mail, household, estimate	POQ (0.26)

(continued)

Table 1. (continued)

#	Topic name	Keywords	Dominant journal ($\overline{\theta_t}$)
18	Cultural studies/critique	space, critique, feminist, contemporary, copyright, meme, ethnographic	CT (0.15)
15	International communication	unite, international, chinese, country, european, state, china	PC (0.03)
20	Mobile communication	mobile, technology, ict, adoption, innovation, icts, diffusion	ICS (0.06)
38	Gender and sex	man, gay, sexual, deception, advice, sex, pornography	HCR (0.04)
5	Communication theory	research, literature, review, empirical, future, gap, researcher	CT (0.09)
6	Attitude change	exposure, viewer, selective, watch, crime, arousal, television	MP (0.14)
26	Discourse studies	discourse, war, metaphor, discursive, racist, racism, asylum	DS (0.3)
34	Language	speaker, interactional, sequence, talk, conversation, conversational, language	RLSI (0.52)
10	Quantitative analysis	code, question, analysis, interview, category, theme, quantitative	RLSI (0.06)
4	Technical communication	reader, book, format, suicide, word, comprehension, writer	TC (0.08)
35	Computer-mediated communication/face-to-face communication	cmc, cue, face, nonverbal, gender, computer, impression	HCR (0.05)
27	Team communication	team, organizational, collaboration, practitioner, communicative, practice, communication	CT (0.14)
22	Climate	climate, period, change, journal, long, decade, past	TC (0.06)
12	Education	student, teacher, learn, university, college, classroom, teach	TC (0.04)
21	ICT	web, website, search, internet, interactivity, wikipedia, google	JCMC (0.05)
1	Business communication	business, development, industry, company, market, creation, software	TC (0.08)
17	VR	virtual, immersive, therapy, pain, ptsd, real, rehabilitation	CBSN (0.13)
3	Communication in the workplace	employee, job, system, workplace, electronic, work, worker	TC (0.13)

Note. The dominant journal for each topic is selected by the highest among all journals. The θ_t for each journal is calculated by the average θ_t of all articles published in the respective journal. CTM = correlated topic model; CM = *Communication Monographs*; CR = *Communication Research*; CT = *Communication Theory*; CBSN = *Cyberpsychology, Behavior, and Social Networking*; DS = *Discourse & Society*; HCR = *Human Communication Research*; HC = *Health Communication*; ICS = *Information, Communication & Society*; ICT = Informational and Communication Technology; JHC = *Journal of Health Communication*; JC = *Journal of Communication*; JCMC = *Journal of Computer-Mediated Communication*; MP = *Media Psychology*; NMS = *New Media & Society*; PC = *Political Communication*; POQ = *Public Opinion Quarterly*; PUS = *Public Understanding of Science*; RLSI = *Research on Language and Social Interaction*; TC = *Technical Communication; VR = Virtual Reality.*

published in each year (Griffiths & Steyvers, 2004). This average topic-membership probability of a topic can be interpreted as the popularity of that topic for the given year. The growth as well as the decline of a given topic over time can be studied by a simple correlation between the year and the average topic-membership (Griffiths & Steyvers, 2004).

Collection of Citation Data

To obtain the citation counts of each publication included in our sample, we collected two sets of citation data. The first set is derived from Google Scholar, in which we launched auto-searches for the title and the authors' names during the month of July 2018. Additionally, another, and possibly more nuanced set was derived from searching for articles in Web of Science (WoS). In this set of data, we also have information on the papers cited by each paper included in our sample. Based on the distinction between out-field and in-field citations (Zhu & Fu, 2019), we calculated three different citation counts for every publication: the global citation count (GCC), the in-field citation count (ICC), and the in-field external citation count (IECC). The GCC of an article is the number of citations the article has received by articles from all other journals indexed by Google Scholar since the time of publication. The ICC of an article is the number of times the article has been cited by all included articles from the 18 high-impact communication research journals. The IECC excludes all self-citations from the ICC. We define a self-citation as a citation referring to another article with overlapping authorship of at least one person (Carley et al., 2013). In sum, the three citation counts have slightly different meanings. The GCC is the popularity of an article within and across the discipline boundary. The ICC is the popularity of an article among authors having published in the selected high-impact journals. The IECC is the popularity of an article among authors having published in the selected high-impact journals excluding one's own research group.

Due to their different meanings, we benefit from jointly looking at these three measures in order to provide a comprehensive perspective on the popularity of a research topic. The IECC is a better quantification of popularity than ICC within the discipline boundary because it controls for inflating the popularity of one's research (and in turn of one's subfield) by self-citation. However, the metric also tends to overadjust for the popularity of subfields of very collaborative scholars. In terms of completeness, the GCC does have the best coverage due to limitation of inclusiveness in the WoS data.[12]

Bayesian Multilevel Zero-Inflated Negative Binomial Regression Analysis

To study the popularity of research topics, a Bayesian multilevel zero-inflated negative binomial regression model was constructed.[13] Our dependent variable was the number of citations per year of the *i*th paper in its specific journal *j* (μ_{ij}/t_i).[14] As the independent fixed-effect (Level-1) variable, the θ_t for each article was used to represent the topics of the article. As the value of one element of θ_t can be perfectly predicted by the rest of the other 39, the problem of a dummy variable trap arises.

Therefore, we decided to drop the topic of health communication (see Table 1, Topic 7) as publications on this topic had the lowest average number of citations. The journal (J_i), in which the paper was published, was entered as a random-effect (Level-2) variable into the regression as citation counts from papers published in the same journal are more likely to be correlated. Therefore, the regression equation is:

$$\log\left(\frac{\mu_{ij}}{t_i}\right) = \beta_0 + \left(u_{0j}J_i\right) + \beta_1\theta_{t1i} + \beta_2\theta_{t2i} + \beta_3\theta_{t3i} + \cdots + \beta_{39}\theta_{t39i} + \varepsilon_{ij}$$

We used the *brms* implementation of the Bayesian mixed-effect zero-inflated negative binomial regression for R (Bürkner, 2017). We interpret the regression coefficients for each θ_t as the magnitude of topic influence on the citation count of an article adjusted for the year of the publication and the journal. For example, the regression coefficient β_1 for θ_{t1i} of 1.7 indicates that one unit increase in θ_{t1} (e.g., from 0 to 1) translates into an averaged 1.7 difference in log expected citation per year adjusted for other variables.

Citation Network Analysis

All articles for which we had information on citing articles were included for the citation network analysis. For this analysis, each article was then assigned the best topic by selecting the topic with the highest θ_t (Rauchfleisch & Schäfer, 2018) and all self-citations were excluded as in the case of IECC. The aim of this analysis is to study how papers are cited across topics and subsequently to extract the prestige of topics based on PageRank. A node in the citation network represents all papers of a particular topic as assigned by our CTM model. In other words, a node represents one of the 40 topics we extracted with our CTM model. All information assigned to a node is derived from the universe of papers that has been identified to belong to this topic. A weighted, directed edge is the size-adjusted frequency of citations that papers on this particular topic receive from papers belonging to different topics. As the raw frequency of citations is size-dependent—that is, a topic with more papers should have been cited more frequently than other topics with less papers—we adjusted the raw frequency of citations by the widely adopted method of Bollen et al. (2006).[15] This weighted, directed network maps the relationships between research topics based on their citing-cited relationships. This network was then used to extract the PageRank of different research topics.

Results

The distribution of the 12,990 publications included in our sample by year and journal is presented in Figure 1. While in the year 2000, 460 articles were published in high-impact communication research journals, in the year 2017 the number increased to 1,031 articles. Overall, the figure reveals an increasing trend of communication research publications. The number of journal articles being published per year has constantly risen since 2000; reaching its peak in 2016 with 1,100 articles.

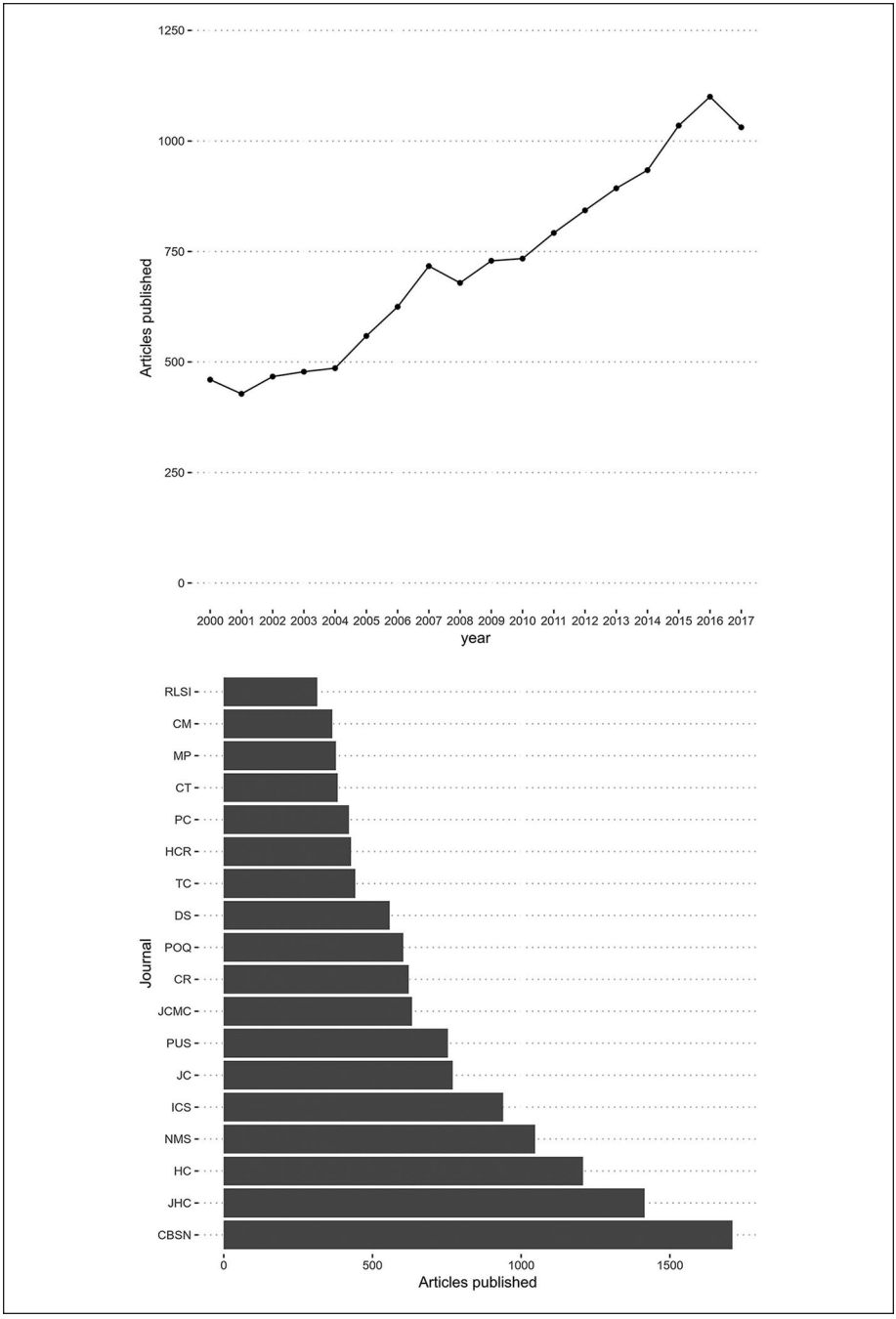

Figure 1. Distribution of papers per year (top) and per journal (bottom).

Figure 1 displays the journals focusing on new media as well as on health communication contributed the majority of publications included in our analysis: CBSN with 1,710, NMS with 1,047, and ICS with 939 publications on new media; and JHC with 1,415 and HC with 1,208 articles on health communication. Supplemental Appendix I depicts the supply of scientific publications by year for each individual journal. The same journals that drive the increase in supply of scientific publications have also raised their publication opportunities. Predominantly since the mid-nineties, these five journals (for new media: CBSN, NMS, and IC; and for health communication: HC and JHC) have been publishing more articles than ever before. In contrast, other journals such as POQ and PUS have only slightly increased possibilities for scholars to publish their research pieces, while other journals (e.g., CT, RLSI, TC) have marginally given scholars an increased opportunity to publish their work.

Supply of Research Topics

Regarding the topical patterns in communication research, a summary of our CTM topic model is presented in Table 1 together with the most dominant journal for the respective topic. Additional analysis on the best fitted papers for each topic is available in Supplemental Appendix II. As the majority of papers in our analysis was published in journals focusing on new media and health communication, our topic model—not surprisingly—identified many subtopics concerning health and online behaviors, for example, social networks, mobile communication, and health campaign. However, communication science research has also focused on other research topics since 2000 such as political communication, journalism, or media psychology. The most dominant topics in our sample can also be linked to a higher concentration in specific journals. For instance, research on problematic ICT use or virtual reality is by far mostly published in CBSN. When publishing research on health campaigns, JHC is the outlet of choice among the high-impact journals.

The correlations between the average topic-membership for each topic and the year of publication are also depicted in Figure 2. As the correlation coefficients show, the number of publications on social media, clinical communication, and studies on media effects and health campaign has particularly increased over time. The growth of research on online communication (e.g., social media) and health communication (e.g., health campaign) can hereby be explained by the growth of publishing opportunities provided by the journals ICS, NMS, HC, and JHC.

Popularity of Research Topics

Regarding the popularity of communication science research, our analysis demonstrates that citations are concentrated on only a handful of papers; a possible symptom of the "rich-gets-richer" phenomenon. Using the "80/20" Pareto's principle, the top 20% of scientific papers with the highest GCC, ICC, and IECC account for 70.5%, 81.3%, and 88.7% of all citations, respectively.[16]

The regression coefficients for each topic from the previously described Bayesian mixed-effect zero-inflated negative binomial regression model are shown in Figure 3

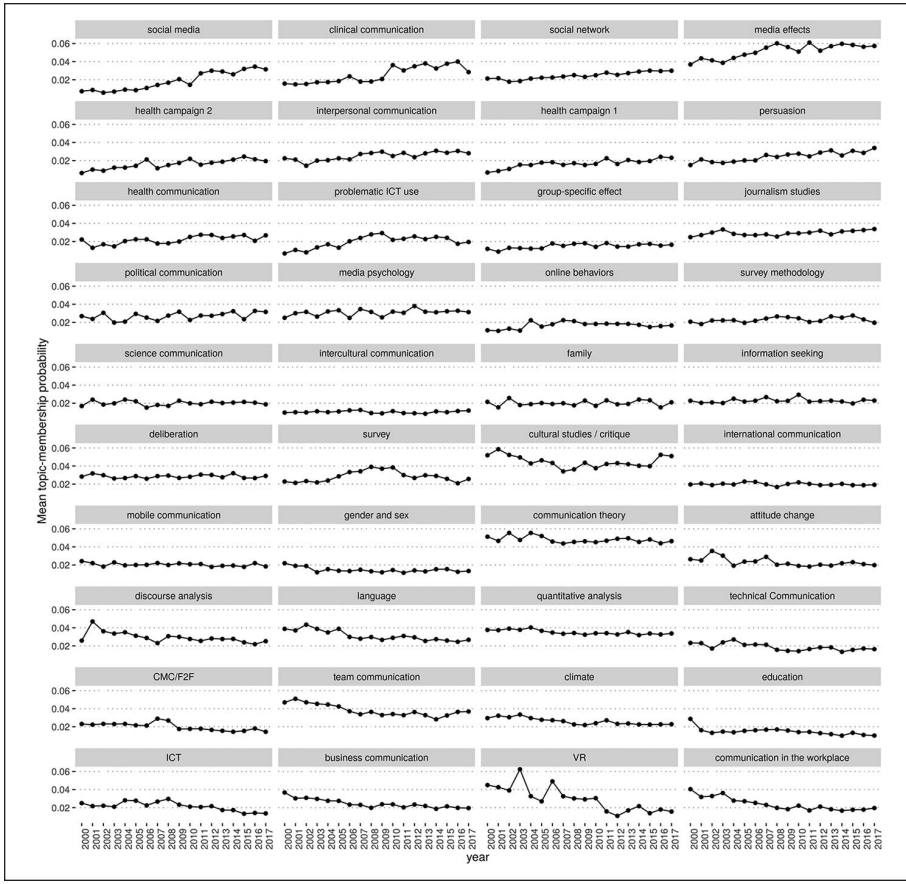

Figure 2. Topical trends in high-impact journals from 2000 to 2017.
Note. From left-to-right, top-to-bottom, the graphs are arranged by decreasing correlation. CMC/F2F = computer-mediated communication/face-to-face communication; ICT = Informational and Communication Technology; VR = Virtual Reality.

(the complete models are available in Supplemental Appendices III–V). As the size of the regression coefficients signals, the top five research topics published in high-impact communication research journals that have received the highest GCC scores since the year 2000 are journalism studies, social media, communication theory, social networks, and problematic ICT usage. However, when looking at the regression coefficients of ICC, journalism studies is no longer the most cited topic. Instead, papers about social media are cited more frequently by other papers published in the 18 high-impact communication journals. The discrepancy between the regression coefficients for GCC and ICC can be interpreted as the difference in popularity in communication research between papers cited by all authors of the universe and those by authors having published in the 18 high-impact communication journals. For example, authors having published in high-impact journals cited significantly fewer papers on

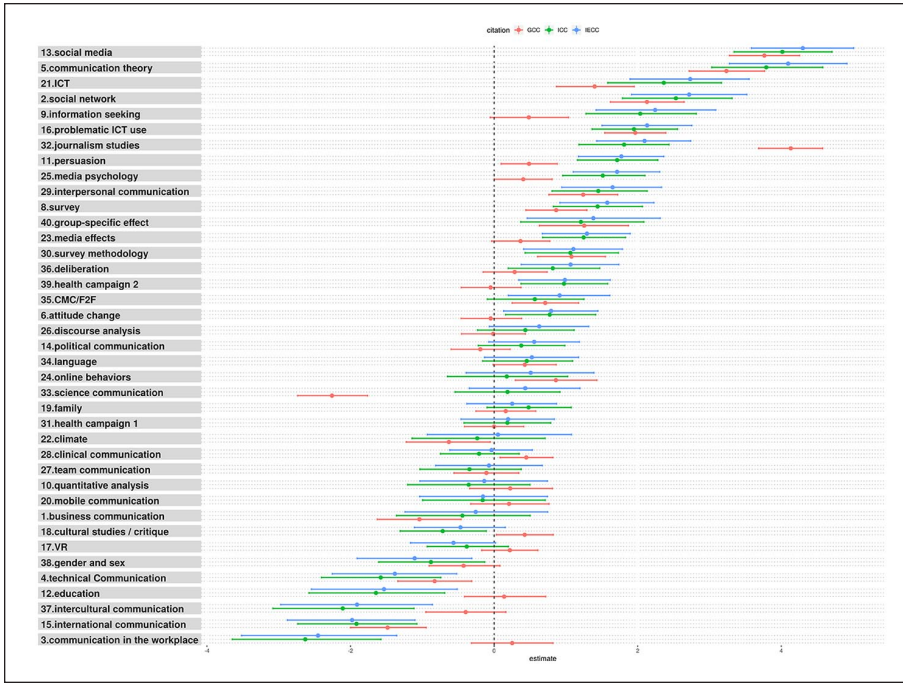

Figure 3. Results of the Bayesian zero-inflated negative binomial regression.
Note. Regression coefficients and their associated 95% credible intervals are shown. Regression coefficients are sorted by their IECC scores in descending order. GCC = global citation count; ICC = in-field citation count; IECC = in-field external citation count; CMC/F2F = computer-mediated communication/face-to-face communication; ICT = Informational and Communication Technology; VR = Virtual Reality.

journalism studies, cultural studies, education, communication in the workplace, and intercultural communication than all authors. However, authors having published in high-impact journals cited significantly more papers on media psychology, persuasion, science communication, and information seeking than all other authors. Regarding IECC, social media is the most frequently cited topic by papers published in the 18 high-impact communication journals; even when excluding self-citation. There is no significant discrepancy between the results of ICC and IECC indicating that self-citations do not influence our estimations of the topical popularity.

Prestige of Research Topics

We calculated the prestige of each topic using the weighted topic citation network (Figure 4). The top five topics with the highest prestige are studies on media effects, social media, persuasion, media psychology, and political communication. In Figure 4, these highly prestigious topics occupy the central position in the backbone structure as determined by the Fruchterman-Reingold layout algorithm. Some clusters of related

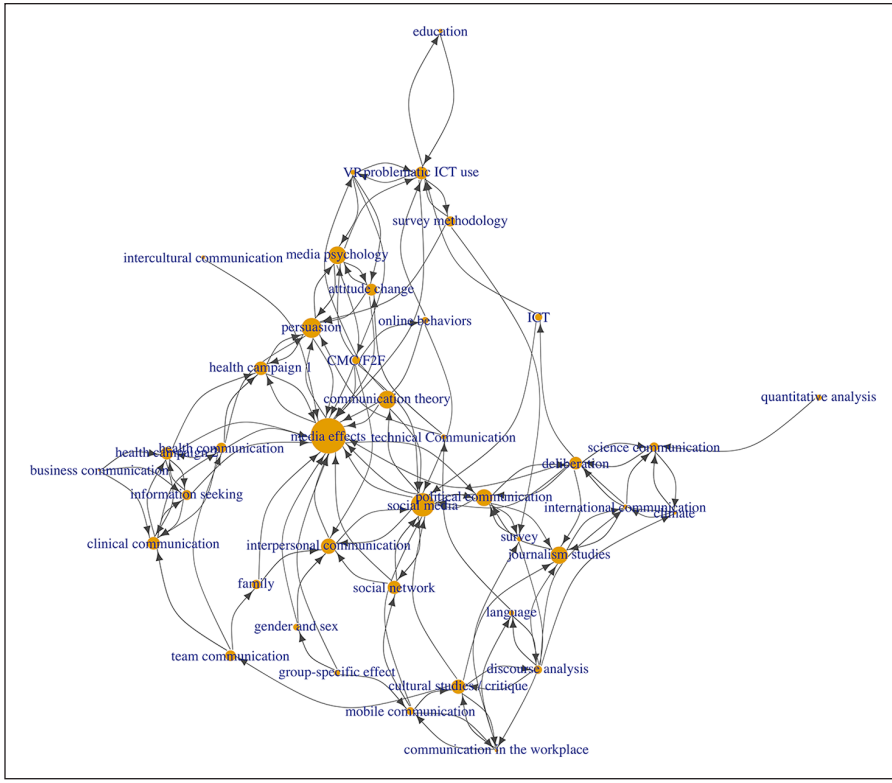

Figure 4. The backbone structure of the topic citation network.
Note. A node represents all articles belonging to a specific topic. A node's size reflects the prestige of a topic. A weighted edge represents the group-size-adjusted number of citations this topic receives from papers belonging to a different topic, that is, a different node. For the purpose of a clearer visualization, only the important edges are extracted to reveal the backbone structure of the network. This was done by filtering all edges using the disparity filter with alpha value equal to .001 and then extracting the giant component. The top 10 topics with the highest PageRank are media effects, social media, persuasion, communication theory, media psychology, journalism studies, political communication, interpersonal communication, cultural studies/critique, and health campaign. An interactive version of this figure is available at: https://doi.org/10.17605/osf.io/uc359. We encourage readers to interact with that version of the figure to experience how the concept of prestige works. CMC/F2F = computer-mediated communication/face-to-face communication; ICT = Informational and Communication Technology; VR = Virtual Reality.

communication research topics can be observed. For instance, the topics on the left are all related to health communication.

What Are the Foci of Scientific Interest?

By using the three metrics of supply (correlation with year), popularity (regression using IECC), and prestige (PageRank), we ranked the 40 topics. By applying a threshold for "high" as >30th place, we partitioned the 40 topics into eight groups (Figure 5). Overall,

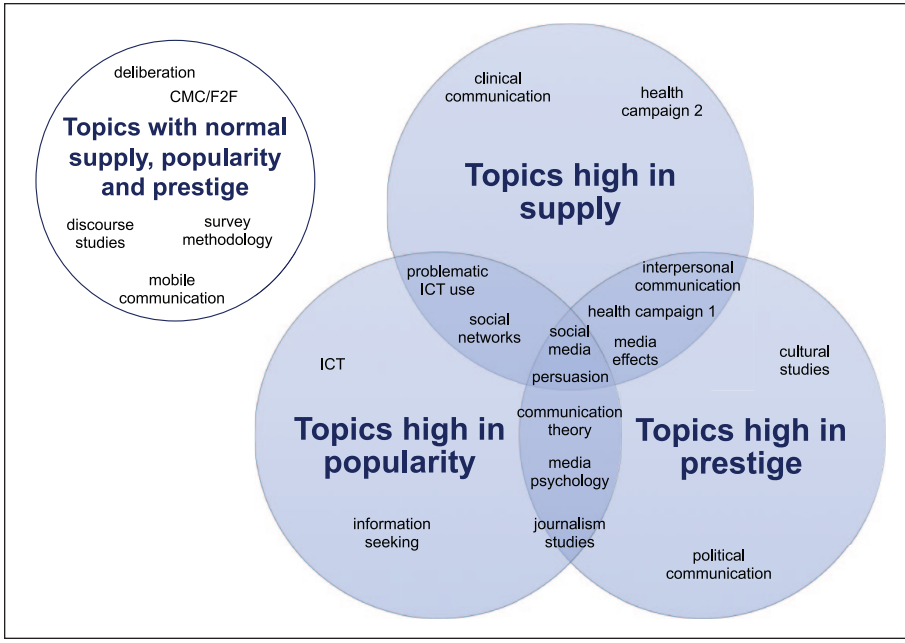

Figure 5. A Venn diagram of topics with high supply, high popularity, and high prestige.
Note. A topic is considered high in one of the three categories, that is, indicated by being inside one of the blue circles, when its metric ranks higher than the 30th place. Only some selected topics are shown under the rubric "Topics with normal supply, popularity, and prestige." CMC/F2F = computer-mediated communication/face-to-face communication; ICT = Informational and Communication Technology.

there exist 17 topics scoring high in at least one of our three metrics. Among these 17, only two topics are high in supply, popularity, and prestige: Those are the research topics of social media and persuasion. Research on social media is the highest among all with the highest supply, and second highest popularity and prestige scores.

Discussion

This study set out to provide an integrative picture on the topical foci in communication research as represented by its top high-impact journals from 2000 to 2017. By using computational methods, this study gained novel insights into the topics of communication science research that are not only high in supply but are also high in popularity and prestige. By employing this cocktail approach of combining supply, popularity, and prestige, the contribution of our work is manifold. First, our study provided insights not only into the topical trends in current communication research but also evaluated their sustainability over time. By complementing these insights with our results on the most popular and most prestigious topics in communication research, we can map the intellectual architecture of our discipline as represented by

its top high-impact journals. Having done so, our findings reveal abundance and scarcity as well as areas of strength and weaknesses of scholarly knowledge on communication research. Informed by this knowledge on our discipline's current foci of research, our study paves the way for an intra- and inter-disciplinary exchange on how to meet the challenges of simultaneously deepening but also broadening research foci, which research foci to prioritize in the future, and how to support today's most pressing societal and political challenges with empirical facts.

As we discussed earlier, this is not the first time for scholars to search for foci of scientific interest. Günther and Domahidi (2017) have studied 80 years of scientific literature to quantify the rise and fall of research topics in communication research. However, their attempt only considers the supply side. While their approach has merits, it does not truly reveal the foci of scientific interest in communication research. First, the increase in CMC-research supply can be attributed to the disproportionate increase in publishing opportunities triggered by the journals NMS and ICS. We can also observe such parallel development for health communication research. Not only the number of research publications has increased over time, also the publishing opportunities provided by the journals HC and JHC have risen. Therefore, tracking *only* the supply of research topics as a method to seek for foci of scientific interest, as in Günther and Domahidi (2017), could be misleading due to variations in publication opportunities. Second, tracking only the supply of research topics is scratching only the surface. In this study, we have incorporated two other dimensions—popularity and prestige—to enrich previous findings such as the ones by Günther and Domahidi (2017). We believe that additionally incorporating the metric of prestige into the mix is particularly important for the analysis of communication research's foci of scientific interest. Some communication scholars worry about the continuous fragmentation of the field (Ang et al., 2019).[17] Therefore, research topics that can have the potential to intersect knowledge from different subfields should also be considered valuable and thus be rightfully deemed as foci of scientific interest. Previous studies revealing the most prestige theories, communication journals, and scholars (Chung et al., 2013; Griffin et al., 2016) only contribute little to the search of foci of scientific interest. This is for instance reflected in the fact that thematically open communication journals, such as JC and CR, were found to be the top two most prestigious journals (Griffin et al., 2016). While this information might be helpful for researchers to select a journal for submitting their papers to, it says almost nothing about foci of scientific interest: These two journals publish papers on a vast variety of topics. While the discovery of framing theory as the most prestigious theory in the field (Chung et al., 2013) is more informative, it still is a broad observation, particularly in light of the fact that the field of communication research is largely organized by research topics rather than theories.

The immediate benefit of our cocktail approach becomes clear when one looks at the intersections between topics with high supply, high popularity, and high prestige (Figure 4). When one only considers the supply side as Günther and Domahidi (2017) did, topics with high supply but normal popularity and normal prestige—such as clinical communication—would be considered foci of scientific interest. Although these

topics are frequently written about, they receive comparatively few citations and are in relatively peripheral position in the citation network (Figure 4). On the contrary, traditional subfields, such as media psychology and journalism studies, are popular and prestigious but not in high supply. These topics would have been missed when one only looks at the supply side.

In sum, we have identified only two topics that rank in the top 10 in the three dimensions: persuasion and social media. Both research topics are timely and with great relevance to our increasing digitized and mediated environment. The ability for these research topics to be the integration force of the field has previously been theorized. Many subtopics of persuasion research—such as social influence and motivated reasoning—are central topics across a variety of fields such as media psychology, political communication, and health communication (Carpenter, 2019). Persuasion increasingly occurs now through the internet (Y. J. Kim & Hollingshead, 2015). Similarly, O'Sullivan (1999) has theorized that innovation in communication technologies, such as social media, can bridge the so-called "great divide" of mass and interpersonal communication of the field. O'Sullivan and Carr (2018) therefore introduce the concept of "masspersonal" communication through digital communication tools and mark the end of the so-called "great divide."

This timely focus of communication research on these two topics should be welcomed as it highlights the societal need to understand the new digitalized world and provides a force to unite the field intellectually. However, many scholars have expressed their concerns about this digital turn of communication research to the field as a whole. For instance, Waisbord (2019) reminds us that communication researchers are not the only issue owner of digital communication research. Other fields such as computer science probably have a bigger voice in that area of research. Scholars in these fields have the skill and the data to unleash knowledge from vast amount of social media data. Unfortunately, the Facebook-Cambridge Analytica data scandal has pushed social media companies to close their application program interfaces (APIs). Communication researchers are still debating about the repercussions of this sudden closure (Bruns, 2019; Puschmann, 2019). Supposing these easily available data were still available, the field would still need to establish a systematic theory building strategy surrounding the analysis of social media data (Parks, 2014). Rather than just being a trend—or worse, a fad—communication research needs to address the above-mentioned concerns to turn our foci of scientific interest into a sustainable, integrative, and genuine force of knowledge advancement.

Some observations from this study deserve further elaboration. Regarding the research topic of journalism studies, it is a classic subfield of communication research and has a relatively stable supply in high-impact journals. The respective studies in this topic show high popularity and high prestige. However, we observe varying interest among the respective authors of these studies and authors having published in high-impact journals by comparing the regression coefficients derived from GCC and ICC/IECC (Figure 3). As our study is based on articles published in the top five high-impact journals between 2000 and 2017, the major journalism studies journals

with a comparatively low impact factor, such as *Digital Journalism* or *Journalism Studies*, were not included in our sample. Consequently, citations by papers published in these or similar journals are not reflected in our ICC and IECC scores. The discrepancies on one hand show the limitation of our ICC and IECC scores, on the other hand allow us to study the difference in popularity of journalism studies papers among authors published in all publication outlets and those published in high-impact journals. Scholars, high-impact journals' editors included, should not ignore journalism studies, because this prestigious research topic is still highly popular *outside* the high-impact journals' circle.

Second, our study drew on abstracts rather than full papers to measure topical patterns in communication research. Although we have discussed the appropriateness of this approach for the extraction of research topics from research papers, the limitations this approach entails must not be overlooked. In particular, the methodological characteristics of the collected scientific publications are not reflected by our CTM model. Extraction of methodological characteristics and then studying their supply, popularity, and prestige are not part of our research goals and research questions. Scholars who are interested in analyzing these features could follow the Subject-Method Topic Network Analysis approach by K. Lee et al. (2016).

Due to the long publication cycle, the results from this study might be instantly outdated when this article hits the journal, similar to all other previous trend analyses published in academic journals. In contrast to other trend analysis paper, the unique and timeless feature of this study specifically relies on our integrated methodological approach which is tailored for the analysis of our field.[18] In order to grab the complete and up-to-date picture, communication research community should push for a collaborative effort to study the entirety of our discipline, which includes the study of all journal articles published in all communication journals in full text as well as books and proceedings. It also eliminates the fragmentation of findings due to the differences in inclusion criteria, for example, the findings from this study are not completely comparable with Günther and Domahidi (2017) because our study includes more recent publications. We are convinced that in the near future the communication research community will be able to extend our data set and conduct such analysis more regularly.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Chung-hong Chan https://orcid.org/0000-0002-6232-7530

Supplemental Material

Supplemental material for this article is available online.

Notes

1. However, there also exist some exceptions; like the computational methods division of the *International Communication Association* (ICA) or communication journals dedicated to specific methodological approaches (e.g., *Computational Communication Research*).
2. Some communication researchers also refer to this type of analysis as a "thematic meta-analysis" (e.g., Kamhawi & Weaver, 2003; S. T. Kim & Weaver, 2002). We opt not to use this term as it may create confusion. A "meta-analysis" is a research synthesis: Statistical procedures are applied to estimate effect sizes (Johnson et al., 2008). A mere analysis of themes in the research literature should thus not be labeled a meta-analysis.
3. The study of Borah (2017) is an exception as it encompassed a 16-year period.
4. When assigning a research topic to a specific journal article with manual content analysis, there exist two additional pitfalls: First, there is a greater risk of classification errors, that is, of wrongly assigning a topic to a journal article due to human errors such as coder fatigue or coder drift (Neuendorf, 2016, p. 170) when the text corpus is large. Second, a-priori-determined categories are by no means exhaustive, and new, just emerging topics are highly unlikely to be included in the category scheme (H. Lee & Kang, 2018). An additional benefit of using computational approach is that it reduces the risk of topic classification errors due to human coder errors. In other words, the likelihood of assigning an inaccurate topic to a document can be decreased by using topic modeling instead of coding the topics manually based on a category scheme (Jacobi et al., 2016). Furthermore, while manual coding relies on predetermined categories, a computational approach has the advantage of a flexible and adaptable categorization, which thus increases the likelihood of discovering new and emerging topics (Roberts et al., 2014).
5. Although not for the analysis of topical trends, the same topic extraction approach has been adopted by Song et al. (2020) to study the sub-disciplinary linkages of the field.
6. A similar but potentially misleading label for this citation-based popularity measure is "impact." However, over the past, referring a publication's "impact" from its number of citations has been deemed scientifically inaccurate: A citation count does not provide any information on the scientific impact of a paper as the reasons for citing a paper are manifold and not necessarily a sign of acknowledgment (e.g., Case & Higgins, 2000).
7. This difficult process is described by Waisbord (2019) as if one was "hoping for the Beatles to reunite" as there is little incentive to do so and it is marred by technical difficulties. Nonetheless, one of the flagship journals of the field has recently published a special issue about speaking across multiple subfields (https://academic.oup.com/joc/issue/70/3).
8. PageRank was invented by Larry Page and Sergey Brin to rank the importance of web pages resulting from a search engine query. It uses the same principle of prestige to rank the importance of web pages: A web page becomes important when it not only receives many hyperlinks, but these hyperlinks also originate from other important web pages.
9. We excluded *Augmentative and Alternative Communication, International Journal of Advertising, International Journal of Language & Communication Disorders*, and *Journal of Advertising and Public Culture*. This methodological decision is also supported by the fact that these five journals do not mention the word "communication" in their aims and scopes at all.

10. A small number of research papers published in *Political Communication* during the year 2000 did not contain abstracts. These papers were excluded from the analysis.
11. As per the suggestions by Maier et al. (2018), two authors of this paper independently read the topic keywords and random articles from each topic with high θ_t. Then, we selected the best solution which the topics are not too granular (many subtopics that should be merged into one topic) or too broad (many topics have been merged into one big topic). We have tried experimentally to test the solution with 70 topics. The topics are too granular to be interpretable.
12. Web of Science (WoS) data does not have complete coverage of all papers published in *Information, Communication & Society* (ICS; since 2009), *Journal of Computer-Mediated Communication* (JCMC; since 2005), and *New Media & Society* (NMS; since 2001). Also, the matching of the two sets of data is based on DOIs. Thus, papers without DOIs—for example, papers from TC—could not be matched.
13. The two major reasons for using a Bayesian model instead of a frequentist mixed-effect model are its flexibility in handling the very complicated multi-level structures and the possibility of incorporating prior knowledge into the model (McElreath, 2015). As we have not incorporated prior information into our models and the multi-level structure in our analysis is relatively simple, our results should be very similar, if not the same, as the frequentist counterpart. Practically speaking, the *brms* implementation of Bayesian models is more efficient than the built-in implementations of frequentist mixed-effect models (e.g., nlme, lme4) in R. Also, standard errors from these built-in frequentist implementations are quite difficult to interpret; for example, no standard errors are provided for random effect coefficients.
14. Any count model is modeling the log transformed value of count. We entered the count of citations (μ_{ij}) as the dependent variable and the year of publication (log-transformed of 2018—year t_i) as an offset value in the regression model. An offset value is a value in the right-hand side of the regression equation that does not have a regression coefficient. By moving the offset value to the left-hand side of the regression, we effectively modeled the rate of citation per year as $\log(\mu_{ij}) - \log(t_i) = \log(\mu_{ij} / t_i)$.
15. This method adjusts the raw frequency of citations of a topic j (t_j) by a topic i (t_i) using the frequency of citations of the t_j by all topics. This size-adjusted frequency is called propagation proportion (Bollen et al., 2006). Therefore, the sum of all propagation proportions of a t_j is always 1.
16. Alternatively, we can demonstrate this by fitting a power law distribution to our data and then calculate the *p*-value based on a bootstrap version of the Kolmogorov-Smirnov test (Clauset et al., 2009). Unfortunately, this procedure is not perfectly suitable for our data because a power law distribution does not allow zeros while a large portion of our citation counts is zeros (GCC = 1.6%; ICC = 41.6%; IECC = 45.7%).
17. The addition of prestige rather just relying on citation-based popularity can also counteract some disadvantages of our popularity measure we have suggested previously (e.g., easily manipulable). Once again, we would like to stress that papers with more citations do not imply better quality or higher impact because raw frequency of citations can be influenced by other factors such as the Matthew effect, that is, papers cited more by more researchers tend to be cited even more often. Thus, we frame measures based on citation count barely as a measurement of popularity.
18. To adhere to the open science principle, we share the data and software used in this study at: https://doi.org/10.17605/osf.io/uc359

References

Allen, M. (2017). *The sage encyclopedia of communication research methods*. SAGE. https://doi.org/10.4135/9781483381411

Ang, P. H., Knobloch-Westerwick, S., Aguaded, I., Muñoz-Uribe, J. F., Wasserman, H., & Athique, A. (2019). Intellectual balkanization or globalization: The future of communication research publishing. *Journalism & Mass Communication Quarterly*, *96*(4), 963–979. https://doi.org/10.1177/1077699019878461

Antons, D., Kleer, R., & Salge, T. O. (2016). Mapping the topic landscape of JPIM, 1984-2013: In search of hidden structures and development trajectories. *Journal of Product Innovation Management*, *33*(6), 726–749. https://doi.org/10.1111/jpim.12300

Berger, C. R. (1991). Chautauqua: Why are there so few communication theories? Communication theories and other curios. *Communication Monographs*, *58*, 101–113. https://doi.org/10.1080/03637759109376216

Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, *1*(1), 17–35. https://doi.org/10.1214/07-AOAS114

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(1), 993–1022. https://dl.acm.org/doi/10.5555/944919.944937

Bollen, J., Rodriquez, M. A., & Van de Sompel, H. (2006). Journal status. *Scientometrics*, *69*(3), 669–687. https://doi.org/10.1007/s11192-006-0176-z

Borah, P. (2017). Emerging communication technology research: Theoretical and methodological variables in the last 16 years and future directions. *New Media & Society*, *19*(4), 616–636. https://doi.org/10.1177/1461444815621512

Bornmann, L., & Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, *64*(1), 45–80. https://doi.org/10.1108/00220410810844150

Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H. D. (2008). Citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, *8*(1), 93–102. https://doi.org/10.3354/esep00084

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

Bruns, A. (2019). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, *22*(11), 1544–1566. https://doi.org/10.1080/1369118X.2019.1637447

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Carley, S., Porter, A. L., & Youtie, J. (2013). Toward a more precise definition of self-citation. *Scientometrics*, *94*(2), 777–780. https://doi.org/10.1007/s11192-012-0745-2

Carpenter, C. J. (2019). Cognitive dissonance, ego-involvement, and motivated reasoning. *Annals of the International Communication Association*, *43*, 1–23. https://doi.org/10.1080/23808985.2018.1564881

Case, D. O., & Higgins, G. M. (2000). How can we investigate citation behavior? A study of reasons for citing literature in communication. *Journal of the American Society for Information Science*, *51*(7), 635–645. https://doi.org/10.1002/(SICI)1097-4571(2000)51:7%3C635::AID-ASI6%3E3.0.CO;2-H

Chung, C. J., Barnett, G. A., Kim, K., & Lackaff, D. (2013). An analysis on communication theory and discipline. *Scientometrics*, *95*(3), 985–1002. https://doi.org/10.1007/s11192-012-0869-4

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703. https://doi.org/10.1137/070710111

Cooper, R., Potter, W. J., & Dupagne, M. (1993). A status report on methods used in mass communication research. *The Journalism Educator*, *48*(4), 54–61. https://doi.org/10.1177/107769589304800408

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. https://doi.org/10.1111/j.1083-6101.2009.01459.x

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

Franceschet, M. (2010). The difference between popularity and prestige in the sciences and in the social sciences: A bibliometric analysis. *Journal of Informetrics*, *4*(1), 55–63. https://doi.org/10.1016/j.joi.2009.08.001

Freelon, D. (2020). On writing in communication and media studies: Two brief points on publication impact. *International Journal of Communication*, *14*, 427–429. https://ijoc.org/index.php/ijoc/article/view/13829

Griffin, D. J., Bolkan, S., Holmgren, J. L., & Tutzauer, F. (2016). Central journals and authors in communication using a publication network. *Scientometrics*, *106*(1), 91–104. https://doi.org/10.1007/s11192-015-1774-4

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, *101*, 5228–5235. https://doi.org/10.1073/pnas.0307752101

Gross, P. L., & Gross, E. M. (1927). College libraries and chemical education. *Science*, *66*(1713), 385–389. https://doi.org/10.1126/science.66.1713.385

Günther, E., & Domahidi, E. (2017). What communication scholars write about: An analysis of 80 years of research in high-impact journals. *International Journal of Communication*, *11*, 3051–3071. https://ijoc.org/index.php/ijoc/article/view/6989

Ha, J. H., & Boynton, L. (2014). Has crisis communication been studied using an interdisciplinary approach? A 20-year content analysis of communication journals. *International Journal of Strategic Communication*, *8*(1), 29–44. https://doi.org/10.1080/1553118x.2013.850694

Herbst, S. (2008). Disciplines, intersections, and the future of communication research. *Journal of Communication*, *58*(4), 603–614. https://doi.org/10.1111/j.1460-2466.2008.00402.x

Ioannidis, J. P., Baas, J., Klavans, R., & Boyack, K. W. (2019). A standardized citation metrics author database annotated for scientific field. *PLoS Biology*, *17*(8), Article 3000384. https://doi.org/10.1371/journal.pbio.3000384

Jacobi, C., Van Atteveldt, W., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, *4*(1), 89–106. https://doi.org/10.1080/21670811.2015.1093271

Johnson, B. T., Scott-Sheldon, L. A. J., Snyder, L. B., Noar, S. M., & Huedo-Medina, T. B. (2008). Contemporary approaches to meta-analysis in communication research. In A. F. Hayes, M. D. Slater & L. B. Snyder (Eds.), *The sage sourcebook of advanced data analysis methods for communication research* (pp. 311–347). SAGE.

Kamhawi, R., & Weaver, D. (2003). Mass communication research trends from 1980 to 1999. *Journalism & Mass Communication Quarterly*, *80*(1), 7–27. https://doi.org/10.1177/107769900308000102

Kim, S. T., & Weaver, D. (2002). Communication research about the Internet: A thematic meta-analysis. *New Media & Society*, *4*(4), 518–538. https://doi.org/10.1177/146144402321466796

Kim, Y. J., & Hollingshead, A. B. (2015). Online social influence: Past, present, and future. *Annals of the International Communication Association*, *39*(1), 163–192. https://doi.org/10.1080/23808985.2015.11679175

Lee, H., & Kang, P. (2018). Identifying core topics in technology and innovation management studies: A topic model approach. *The Journal of Technology Transfer*, *43*(5), 1291–1317. https://doi.org/10.1007/s10961-017-9561-4

Lee, K., Jung, H., & Song, M. (2016). Subject-method topic network analysis in communication studies. *Scientometrics*, *109*(3), 1761–1787. https://doi.org/10.1007/s11192-016-2135-7

Levine, T. R. (2010). Rankings and trends in citation patterns of communication journals. *Communication Education*, *59*(1), 41–51. https://doi.org/10.1080/03634520903296825

Lin, Y., & Kaid, L. L. (2000). Fragmentation of the intellectual structure of political communication study: Some empirical evidence. *Scientometrics*, *47*(1), 143–164. https://doi.org/10.1108/14691931211276124

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H., & Silke, A. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, *12*(2–3), 93–118. https://doi.org/10.1080/19312458.2018.1430754

Mane, K. K., & Börner, K. (2004). Mapping topics and topic bursts in PNAS. *Proceedings of the National Academy of Sciences*, *101*(Suppl. 1), 5287–5290. https://doi.org/10.1073/pnas.0307626100

McElreath, R. (2015). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

Merton, R. K. (1938). Science, technology and society in seventeenth-century England. *Osiris*, *4*, 360–632. https://doi.org/10.1086/368484

Neuendorf, K. A. (2016). *The content analysis guidebook*. SAGE.

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

Oppenheim, C. (1995). The correlation between citation counts and the 1992 research assessment exercise ratings for British library and information science university departments. *Journal of Documentation*, *51*(1), 18–27. https://doi.org/10.1108/eb026940

O'Sullivan, P. B. (1999). Bridging the mass-interpersonal divide synthesis scholarship in HCR. *Human Communication Research*, *25*(4), 569–588. https://doi.org/10.1111/j.1468-2958.1999.tb00462.x

O'Sullivan, P. B., & Carr, C. T. (2018). Masspersonal communication: A model bridging the mass-interpersonal divide. *New Media & Society*, *20*(3), 1161–1180. https://doi.org/10.1177/1461444816686104

Park, H., & Leydesdorff, L. (2009). Knowledge linkage structures in communication studies using citation analysis among communication journals. *Scientometrics*, *81*(1), 157–175. https://doi.org/10.1007/s11192-009-2119-y

Parks, M. R. (2014). Big data in communication research: Its contents and discontents. *Journal of Communication*, *64*(2), 355–360. https://doi.org/10.1111/jcom.12090

Paul, B., Salwen, M. B., & Dupagne, M. (2000). The third-person effect: A meta-analysis of the perceptual hypothesis. *Mass Communication & Society*, *3*(1), 57–85. https://doi.org/10.1207/s15327825mcs0301_04

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*(3), 675–693. https://doi.org/10.1007/s11192-013-1012-x

Peters, J. D. (1993). Genealogical notes on "the field." *Journal of Communication*, *43*(4), 132–139. https://doi.org/10.1111/j.1460-2466.1993.tb01313.x

Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing & Management*, *12*(5), 297–312. https://doi.org/10.1016/0306-4573(76)90048-0

Puschmann, C. (2019). An end to the wild west of social media research: A response to Axel Bruns. *Information, Communication & Society*, *22*(11), 1582–1589. https://doi.org/10.1080/1369118x.2019.1646300

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. https://doi.org/10.1177/2057047317691054

Rauchfleisch, A., & Schäfer, M. S. (2018). Structure and development of science communication research: Co-citation analysis of a developing field. *Journal of Science Communication*, *17*(3), Article A07. https://doi.org/10.22323/2.17030207

Reardon, K. K., & Rogers, E. M. (1988). Interpersonal versus mass media communication a false dichotomy. *Human Communication Research*, *15*(2), 284–303. https://doi.org/10.1111/j.1468-2958.1988.tb00185.x

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103

So, C. (1998). Citation ranking versus expert judgment in evaluating communication scholars: Effects of research specialty size and individual prominence. *Scientometrics*, *41*(3), 325–333. https://doi.org/10.1007/bf02459049

Song, H., Eberle, J.-M., & Eisele, O. (2020). Less fragmented than we thought? Toward clarification of a subdisciplinary linkage in communication science, 2010-2019. *Journal of Communication*, *70*(3), 310–314. https://doi.org/10.1093/joc/jqaa009

Tomasello, T. K. (2001). The status of Internet-based research in five leading communication journals, 1994-1999. *Journalism & Mass Communication Quarterly*, *78*(4), 659–674. https://doi.org/10.1177/107769900107800403

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. https://doi.org/10.1177/1461444809342762

Waisbord, S. (2019). *Communication: A post-discipline*. Polity.

Wilhite, A. W., & Fong, E. A. (2012). Coercive citation in academic publishing. *Science*, *335*(6068), 542–543. https://doi.org/10.1126/science.1212540

Xhignesse, L. V., & Osgood, C. E. (1967). Bibliographical citation characteristics of the psychological journal network in 1950 and in 1960. *American Psychologist*, *22*(9), 778–791. https://doi.org/10.1037/h0024961

Ye, L., & Ki, E. J. (2012). The status of online public relations research: An analysis of published articles in 1992-2009. *Journal of Public Relations Research*, *24*(5), 409–434. https://doi.org/10.1080/1062726x.2012.723277

Zhang, Y., & Leung, L. (2015). A review of social networking service (SNS) research in communication journals from 2006 to 2011. *New Media & Society*, *17*(7), 1007–1024. http://dx.doi.org/10.1177/1461444813520477

Zhou, Y. B., Lü, L., & Li, M. (2012). Quantifying the influence of scientists and their publications: Distinguishing between prestige and popularity. *New Journal of Physics*, *14*(3), Article 033033. https://doi.org/10.1088/1367-2630/14/3/033033

Zhu, Y., & Fu, K. W. (2019). The relationship between interdisciplinarity and journal impact factor in the field of communication during 1997–2016. *Journal of Communication*, *69*(3), 273–297. https://doi.org/10.1093/joc/jqz012

Author Biographies

Chung-hong Chan (PhD, University of Hong Kong 2017; ORCID: 0000-0002-6232-7530) is a research associate at the Mannheim Centre for European Social Research (MZES). He is a trained epidemiologist but with strong interest in media studies. His research interests include online opinion polarization, online radicalization, and platform interventions. His research has been published in the *Journal of Computer-Mediated Communication, Policy & Internet*, and *IEEE Internet Computing*, among others.

Christiane Grill (PhD, University of Vienna 2016; ORCID: 0000-0003-1334-4657) is a research associate in the project "The Conversations of Democracy. Citizens' Everyday Communication in the Deliberative System" at the Mannheim Centre for European Social Research (MZES) at the University of Mannheim (Germany). Her primary research focuses on interpersonal communication, deliberation, as well as political offline and online communication. Her work has been published in *International Journal of Public Opinion Research, International Journal of Public Opinion Research, Journalism & Mass Communication Quarterly*, and *Communications—The European Journal of Communication*.