The *Journal of Consumer Research* at 40: A Historical Analysis

XIN (SHANE) WANG
NEIL T. BENDLE
FENG MAI
JUNE COTTE

This article reviews 40 years of the *Journal of Consumer Research* (*JCR*). Using text mining, we uncover the key phrases associated with consumer research. We use a topic modeling procedure to uncover 16 topics that have been featured in the journal since its inception and to show the trends in topics over time. For example, we highlight the decline in family decision-making research and the flourishing of social identity and influence research since the journal's inception. A citation analysis shows which *JCR* articles have had the most impact and compares the topics in top-cited articles with all *JCR* journal articles. We show that methodological and consumer culture articles tend to be heavily cited. We conclude by investigating the scholars who have been the top contributors to the journal across the four decades of its existence. And to better understand which schools have contributed most to the knowledge of consumer research over this history, we provide an analysis of where these top-performing scholars were trained. Our approach shows that the *JCR* archives can be an excellent source of data for scholars trying to understand the complicated, challenging, and dynamic field of consumer research.

Keywords: topic modeling, *Journal of Consumer Research*, historical analysis, citation analysis

The late 1960s was a turbulent time. The optimism epitomized by the moon landings was threatened by generational clashes suggesting that traditions could not be relied on. At the same time, the academic field of marketing was undergoing its own revolution. Data analysis was dominating the *Journal of Marketing Research*, leaving the newer fields of consumer behavior and management science applications to marketing relatively neglected (Frank 1995). A coalition under the editorship of Ronald Frank created the *Journal of Consumer Research* (*JCR*) outside the aegis of the American Marketing Association (AMA). *JCR* was to be "the preeminent interdisciplinary journal in consumer behavior" (Frank 1995).

The first issue of *JCR* was published in June 1974. True to its interdisciplinary aim, behavioral economist George Katona (jointly appointed in psychology and economics at the University of Michigan) provided the new journal's debut article. The article examined methodological considerations around a key 1970s business topic: inflationary expectations (Katona 1974). Ever since the inaugural issue, *JCR* has striven to be at the leading edge of consumer research. But reaching 40 seems like the perfect time to review *JCR*'s past, figure out exactly what constitutes the field of consumer research today, and look forward to its future with a renewed self-awareness. In addition, this article considers *JCR*'s impact by examining the works that have been most cited and the authors that have become the top scholars in the journal in terms of

Xin (Shane) Wang is an assistant professor of marketing at the Ivey Business School, Western University (xwang@ivey.ca); Neil Bendle is the David G. Burgoyne Faculty Fellow and assistant professor of marketing at the Ivey Business School, Western University (nbendle@ivey.ca); Feng Mai is a doctoral student at Lindner College of Business, University of Cincinnati, Cincinnati, OH, 45221, USA (maifg@mail.uc.edu), and June Cotte is the Scott and Melissa Beattie Professor and an associate professor of marketing (jcotte@ivey.ca) at the Ivey Business School, Western University, 1255 Western Road, London, ON, Canada, N6G 0N1. Please direct correspondence to Xin (Shane) Wang. Thanks to the *JCR* editorial team for their helpful comments.

Advance Access publication April 15, 2015

 ● Vol. 42 ● 2015
DOI: 10.1093/jcr/ucv009

number of publications. We reveal the ebb and flow of consumer research topics over the years and how the topics covered in the most cited articles differ from the average article.

REVIEWING 40 YEARS OF THE *JOURNAL OF CONSUMER RESEARCH*

Our research covers the first 40 years of *JCR* from June 1974 to April 2014. The field of consumer research has blossomed in this time. From its emergence in the early 1970s, consumer research is now the most popular area of specialization among those seeking jobs in marketing departments. In 2014, 44% (and in 2013 56%) of PhD students who reported accepting an academic post in a marketing department classified themselves as consumer behavior researchers (Dingus, Mellema, and Langan 2013; Mellema, Mills, and Fox 2014). Furthermore, the AMA shows *JCR* as the journal contributing most to early career success for scholars in marketing departments. According to the AMA, in 2014, 14 (in 2013, 17) articles had already been published in *JCR* by marketing faculty applicants before they entered their first post in academia, far exceeding those for any other journal targeted by these scholars (Dingus et al. 2013; Mellema et al. 2014).

Growth in the number of consumer behavior researchers created pressure on the journal to publish more research that triggered the move from four to six issues a year for volume 34 (2007–8). Of course, the published articles only show a small part of the work because *JCR* has maintained an "overall historical rejection rate of approximately 90 percent" (*Journal of Consumer Research* 2014). In this burgeoning research area, the burden of editorship increased, leading to the pursuit of various forms of coeditorship over the years. As a result, many renowned researchers in the field have been involved in the editorial direction of the journal over the past 40 years.

The journal has maintained its quality, preserving its reputation as a top-tier journal focused on consumers and consumption. Throughout its history, *JCR* has sought to maintain its own identity. John Deighton (editor from 2005 to 2011) stated that, in contrast to the other top-tier journals that scholars in marketing departments traditionally target, "*JCR* is not part of the drive to the center" (Deighton 2007). The current editorial team describes the journal as having tripartite aims. Namely, *JCR* aims to "publish work on the broad spectrum of consumer behavior, publish work that contributes to conversations across a range of disciplines, and publish research that is impactful" (Dahl et al. 2014).

Of course, we are not the first to see the potential value in a historical review of a journal. In the field of marketing, recent years have seen the *Journal of Marketing Research* (Huber, Kamakura, and Mela 2014) and *Marketing Science* (Mela, Roos, and Deng 2013) use historical techniques to examine the progress of their journals. Our focus on consumer research is unique, but we also make some important methodological advances over the prior work. As we will explain, we go beyond key words supplied by authors to analyze the content of abstracts using advanced text-mining techniques while also using a new-to-the-field probabilistic topic modeling approach.

Textual Analysis and Understanding the *Journal of Consumer Research*

The archives of *JCR* are an incredibly rich source of data. We do not need to ask consumer researchers what they think makes for a great journal article because we can assess this empirically in several ways. First, we can explore what research has been published over the decades, showing which articles made it into the elite 10% of all submitted manuscripts. Second, we can see what topics have proven to be impactful, in the sense of being heavily cited by other researchers. (Citations are not the only way of assessing impact, but this is a popular method.) With a historical analysis we do not merely wish to see what was popular when it was published, but which topics proved resilient and important over the years. Third, we can determine the scholars most successful in their publishing efforts. We can also tie these scholars to their PhD-granting institution to see which schools have contributed most to *JCR* through training these successful scholars. Finally, *JCR*'s archives allow us to see the ebb and flow of topics in *JCR*, demonstrating how the discipline has evolved.

An interesting feature of *JCR* that sets it apart from other top-tier journals that marketing- and consumer-focused scholars might read is that *JCR* has not used key words until recently to describe articles. Thus the key word analysis techniques that are often highly insightful for historical reviews (Mela et al. 2013) cannot be used to examine the history of the *JCR*. JSTOR does generate its own key terms, but these lack the richness of description that our specialized text mining can achieve. We text-mine the article abstracts to isolate noun phrases, a small collection of words that share a common meaning. Our text mining, in extracting the noun phrases, allows us a more contextualized view of the topics that have proved important in *JCR*.

DATA

The data used in this analysis include all articles published in *JCR* from its inception in June 1974 through to April 2014. This covered 40 volumes in total and represented a total of 2031 articles. However, we omitted from the analysis articles that did not have abstracts. These include editorials, some research notes, and commentaries on research articles by other scholars. The remaining 1875 articles became our review data set. For each article our data come from the following two sources:

1. JSTOR Data for Research (dfr.jstor.org) provides us with article metadata such as Title, Abstract, Author(s), Volume, Issue, and Publication Date. In addition, we obtained word frequencies, key terms, and ngrams that can be utilized for conducting analysis of document-level data. (An ngram is *n* items of text that are contiguous.)
2. Citation data for each article, by year, was gathered from Thomson Reuters Web of Science. Note this is a more conservative measure than the Google Scholar count and so estimates lower levels of citations. Google Scholar uses a wide range of publications including theses and books, whereas Web of Science focuses only on a much smaller set of peer-reviewed content.

Over the last 40 years, the average *JCR* article in our analysis (i.e., the 1875 full articles with abstracts) was cited 66.1 times. There is considerable skew in the data because some articles have extremely high levels of citations relative to the average. The standard deviation is 108.3, the median number of citations is 34.0, and the modal number of citations is zero. (However, articles with zero citations are typically articles published recently, and so they may be expected to receive citations given time to influence the work of other scholars.) There were a total of 123,867 citations up to December 2014 when our data were captured.

ANALYSIS AND RESULTS

Topic Popularity

Our research advances the field of historical analysis of journals focused on consumers and marketing by using text mining of the abstracts for key noun phrases and probabilistic topic modeling. Previous historical research in *Marketing Science* (Mela et al. 2013) focused on investigating the key words as given by authors. This is informative but risks authors offering key words that might appeal to the latest trends. The terminology in fashion may miss the connections between current approaches and the work of pioneering predecessors. We go beyond this by analyzing the text in the abstracts themselves, and so our text-mining approach uncovers the topics that the authors use without relying on authors to identify them.

Another historical analysis of the *Journal of Marketing Research* (Huber et al. 2014) used text-mining techniques, but we advance the field by using a latent Dirichlet allocation (LDA) approach. This approach uncovers the relationship between articles by uncovering which articles address similar topics. LDA allows us to reveal the hidden structure of the work. We can see what proportions of each abstract addresses which topics. Understanding the proportions of topics in each abstract also allows us to map the ebb and flow of topics over time.

An overarching purpose of this analysis is to retain the most important noun phrases while eliminating the common but uninformative ones that may be expected in *JCR* such as "research" or "consumers." We analyze the abstracts using an approach to historical analysis that is new to journals in this field. Specifically, we used Natural Language Toolkit, a language platform, and TextBlob, a library of tools, to extract noun phrases from the article's abstracts. For these noun phrases, we calculate the *term frequency-inverse document frequency* (tf-idf). The tf-idf reflects the most common terms in the abstract relative to how widely the term is generally used in the entire data set of *JCR* abstracts. This procedure gives us terms that were unexpectedly likely to be used in the abstract compared to their general usage in all the text that we analyze. This is more illuminating than simply viewing frequently used terms in the abstract because the most commonly used terms are typically uninformative common words such as "the" or "be." Our text mining of all the abstracts from the last 40 years produced a list of the top 15 noun phrases from each article.

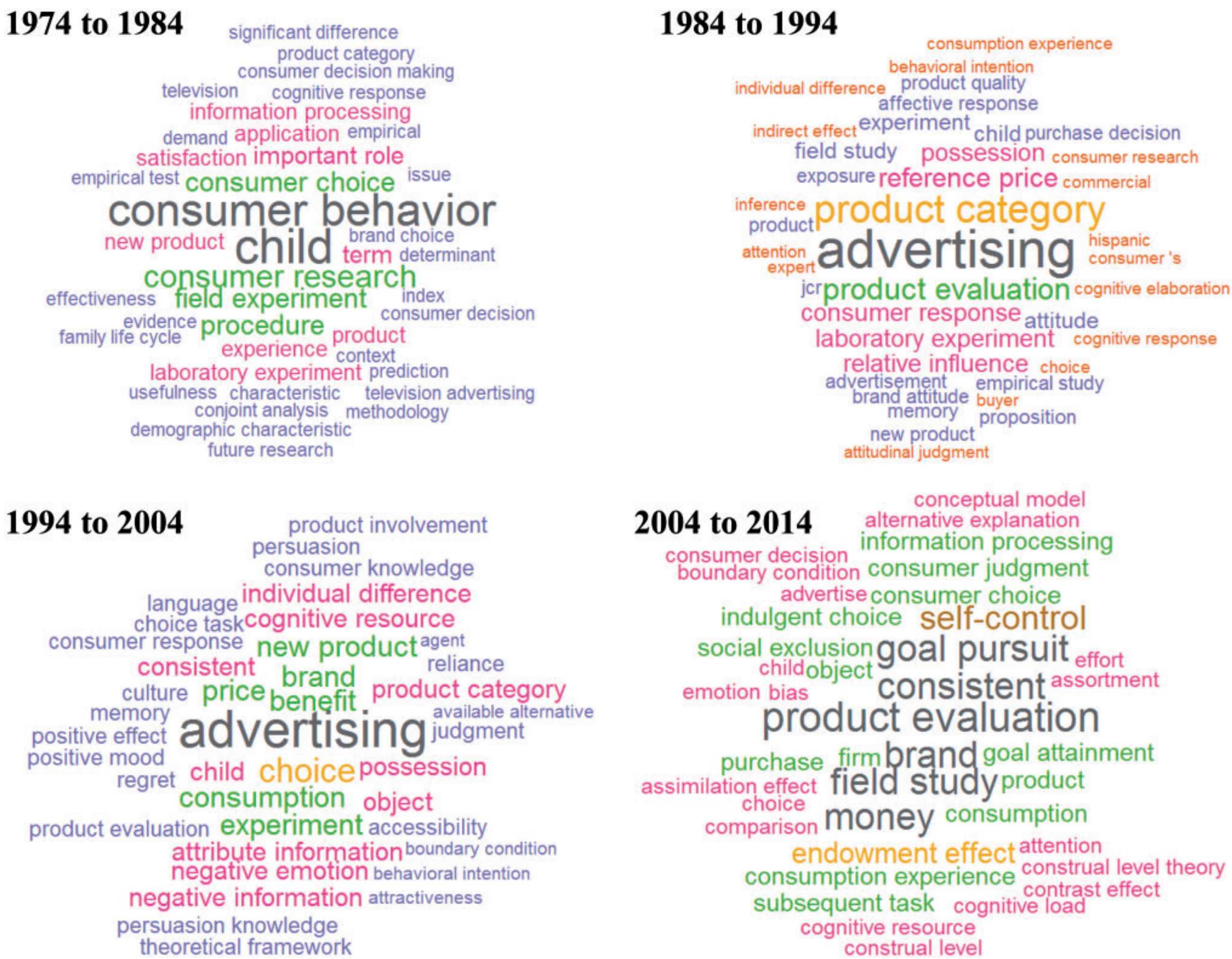
Changing Popularity Over Time

How the journal has changed over time is of great interest to consumer research historians. We therefore took the 40 most popular noun phrases for each 10-year period and created word clouds (figure 1). A word cloud allows noun phrases that have occurred with a greater frequency to have greater size and a more arresting color, allowing for easy identification of the most significant noun phrases over the years and a visual inspection of changes over time.

It is interesting to see the ebb and flow of areas of interest in consumer research. *Child* was a highly visible noun phrase in early research but has declined in relative prominence. That said, terms related to children do feature in all the decades, showing that issues around consumption by young people remain an important part of the journal's focus. Interestingly, field experiment was featured in the first 10 years of the journals. The recent prominence of field studies suggests we might be seeing a return to a technique that has been part of the repertoire of consumer researchers from the early days of the discipline. Interesting historical points arise. For example, one can see the emergence of conjoint analysis in the early years of the journal. Research on advertising rises to prominence in the period 1984–94, only to decline in recent years as product evaluation has gained relative focus. It is also possible to detect a growing theoretical focus in the journal over the years. Methodologically related terms (e.g., *procedure,*

FIGURE 1

WORD CLOUD OF THE 40 MOST COMMON NOUN PHRASES IN *JCR*, 1974–2014



evidence) seem to give way to more theoretical terms (e.g., *construal-level theory, goal pursuit*).

The overarching story seems to be one of consumer research widening and diverging in its focus while simultaneously narrowing the scope of individual inquiries. In the first decade a small number of quite general phrases—*consumer behavior* and *consumer research*—feature heavily, suggesting consumer research had yet to clearly divide into areas of specialization. The most recent decade of *JCR*, however, suggests a proliferation of areas of focus in the journal. A larger number of more specific terms—*field study, choice, brand*, and *product*—feature heavily. In addition, the word clouds depict a move from terms that are more layperson friendly (such as *satisfaction, child,* and *consumer choice*) toward terms that are more theoretical, academic, and, dare we say it, more jargon laden (such as *endowment effect, construal level, boundary condition,* and *conceptual model*). Together, we suggest these changes seem visually to suggest a move away from generalist inquiries of how consumers behave (e.g., evidence, prediction, procedure, empirical, methodology) to move specialist, largely psychologically based inquiries (e.g., goal attainment, contrast/assimilation effects, social exclusion, cognitive resource).

So topics have changed and diverged over the history of the journal. This is to be expected, of course, as interests wax and wane. However, this analysis cannot probe the importance and lasting impact of topics. We turn to this question now.

A DETAILED LOOK AT RESEARCH TOPICS IN *JCR*

Topic Modeling

We next analyze the topics that have featured in *JCR*'s history. To assess these from 40 years of abstracts we use a probabilistic topic modeling procedure, a LDA approach. (For a nontechnical explanation of such modeling, see Blei 2012. A technical appendix for our work is also available.) Using the data set of abstracts for all *JCR* articles, we uncover the hidden (latent) structure of the articles. We looked for three parts to the hidden structure: (1) We searched for a relatively small set of topics (how we determined the precise number is explained later). For example, a topic might be social identity and influence or persuasion. (2) Each article can be seen as a mixture of the topics that our model uncovers. For example, when we review an abstract, the model might judge it to have considered how emotions impact decisions about family budgeting. The article would be classified as a mixture of both the emotional decision-making and family decision-making topics, with the precise mix determined by how heavily each abstract is weighted toward each topic. (3) The model also creates an assignment of specific words from each featured topic to the article. For example, articles dealing with the topic methodological issues use the words *validity, structure,* and *alternative* with much greater frequency than articles that focus on other topics.

The idea behind the probabilistic approach is that we imagine that articles were generated randomly from a hidden structure (Griffiths and Steyvers 2004). (Imagine consumer research being written by monkeys drawing words randomly from urns, with the urns as the hidden structure.) Our probabilistic topic model uses the Dirichlet distribution to estimate the probability of any given hidden structure having been used to generate an abstract with the words we see. The model thus uncovers the hidden structure most likely to have generated the data that we observe. Superior models are those that generate data relatively similar to the actual data.

This approach utilizes the idea that each article is composed of a mixture of different topics, and each topic has certain words associated with it. For example, one topic has words such as *social, identity,* and *group* most closely associated with it. The topics are unnamed in the model, but after reviewing the words associated with each topic, we named each topic to ease the reader's understanding. For example, our topic most associated with the words *social, group,* and *identity*, we named *social identity and influence* given the words most heavily associated with it.

The number of topics judged to be represented in the hidden structure was found by minimizing perplexity. To do this our program splits the data into two subsets. The first subset is used to create a training model explaining the topic structure. The effectiveness of this model is then evaluated on the second subset of the data. We wish to minimize perplexity, which is the surprise that the training model registers. Surprise is the number of equally probable word choices in the evaluation subset of the data; this represents instances where our model is not confident of its prediction. The computed average of how surprised the model was by the words in the second half of the document is recorded, and the model with the lowest number of surprises was chosen.

Our model identified the 16 topics shown in table 1 with their most representative terms and the name we have assigned to each topic.

The model identifies words specifically associated with each topic. These are words with a much higher probability of occurring in articles related to that topic compared to their average chance of appearing across all the data. This allows us to ignore common words, for example, "the," "a," and "we," which are widely used within all abstracts and so not specifically related to any topic. We detailed the 20 most representative words for each topic and ranked them from most to least representative. Several words were representative of more than one topic, leading to only 235 unique words across the 16 topics. For example, evaluation is found in the memory, contextual effects, and satisfying customers topics. Figure 2 shows the words associated with each topic. We can see the weight of representativeness of the word in relation to the topic by the size of the horizontal bar. The larger the bar, the more we should expect to see the word occurring in work on that topic. Thus search is highly representative of the search topic, and we should expect to see it regularly in research on the topic. Variety is moderately representative of the search topic, whereas ambiguity is only mildly representative of the topic.

TABLE 1

JCR TOPICS, 1974–2014

Top Three Terms in Topic	Name Assigned to Topic
Validity, Structure, Alternative	Methodological Issues
Social, Identity, Group	Social Identity and Influence
Time, Cost, Resource	Resource Constraints
Price, Risk, Seller	Buying Process
Meaning, Social, Cultural	Consumer Culture
Response, Exposure, Advertisement	Advertising
Price, Reference, Purchase	Price and Price Associations
Child, Television, Age	Consumption by Children
Emotion, Experience, Emotional	Emotional Decision Making
Memory, Recall, Cue	Memory
Goal, Health, Food	Self-Control and Goals
Search, Set, Option	Search
Family, Household, Wife	Family Decision Making
Persuasion, Message, Processing	Persuasion
Context, Option, Preference	Contextual Effects
Satisfaction, Experience, Language	Satisfying Customers

FIGURE 2

THE REPRESENTATIVENESS OF TERMS WITHIN EACH TOPIC

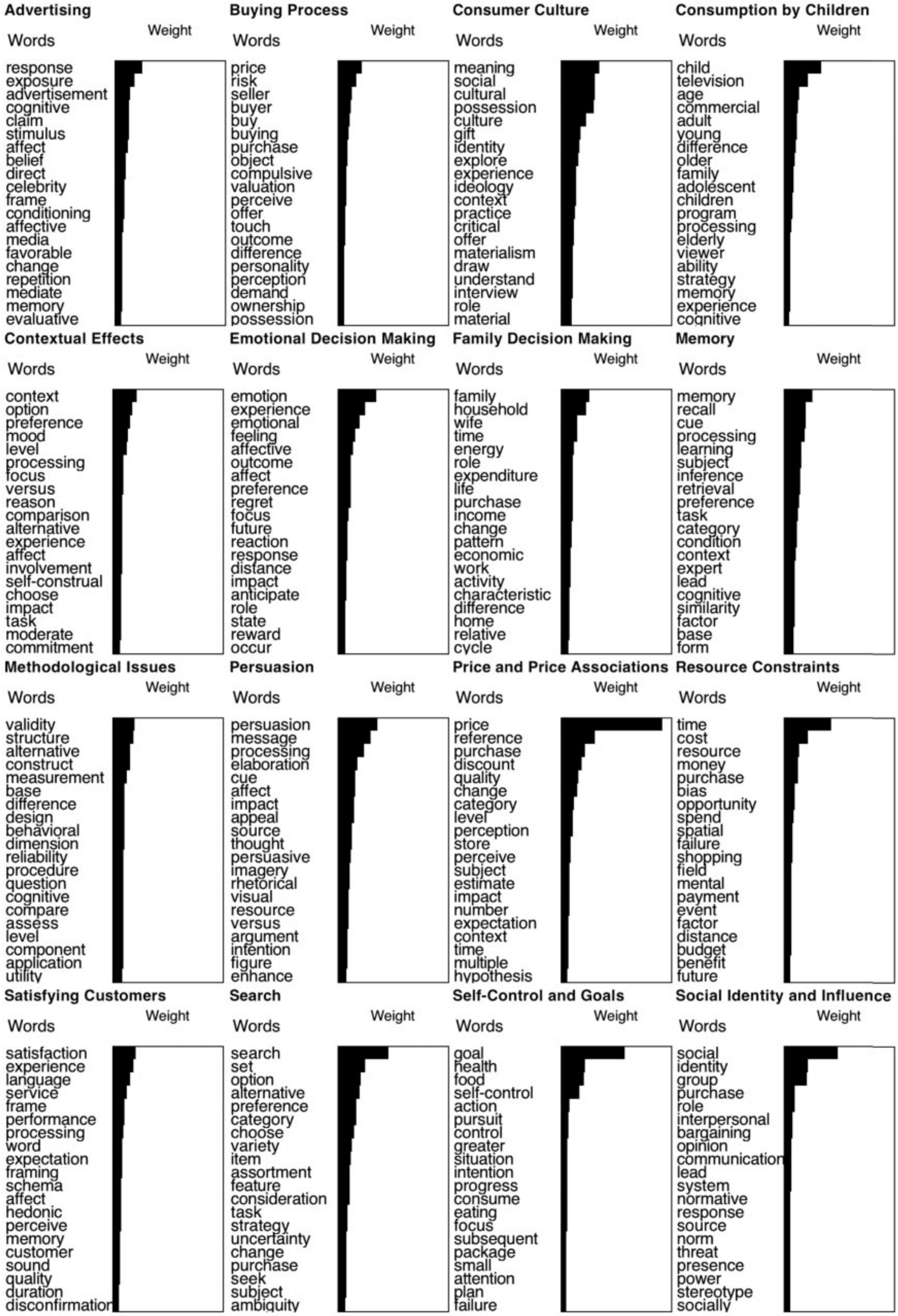
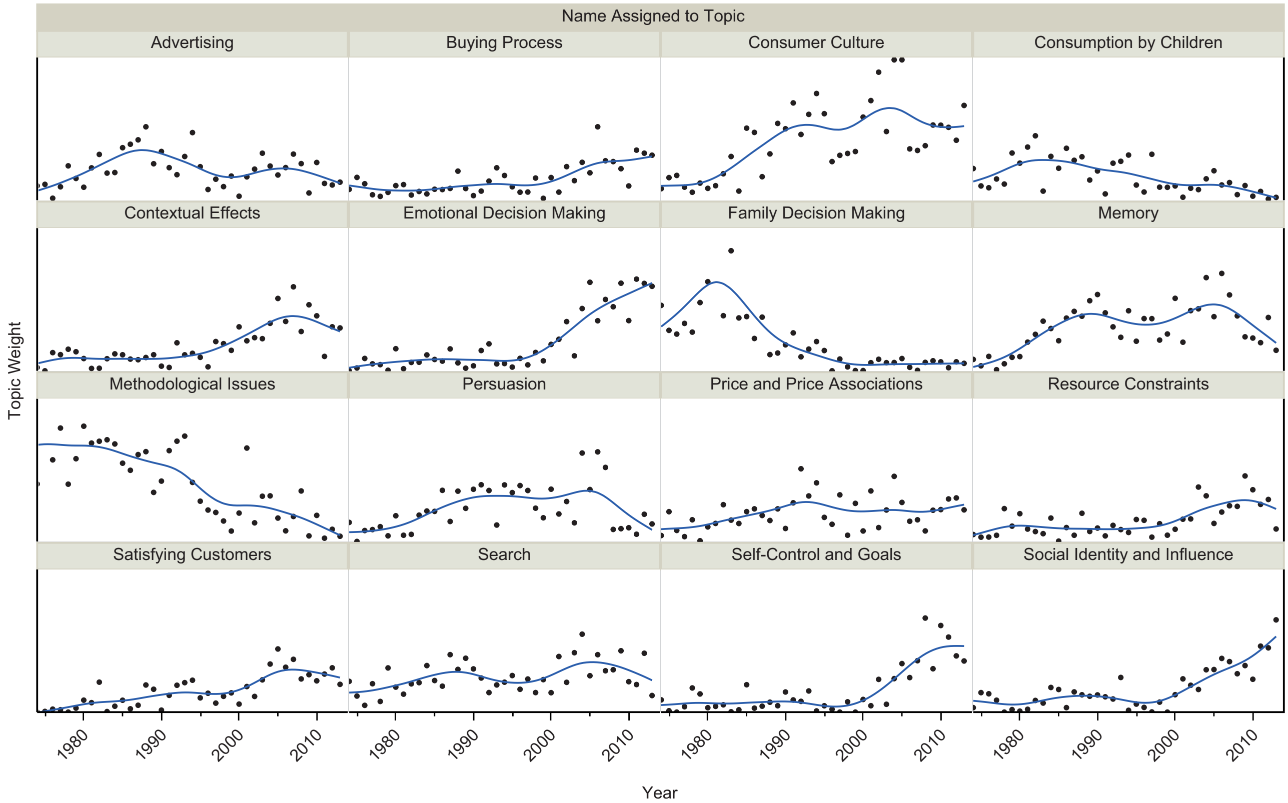


FIGURE 3

TOPIC FOCUS OVER TIME IN *JCR*



The probabilistic topic model also gives us the topic distribution for all 1875 abstracts we reviewed, telling us how each article was classified. For instance, we can assign a topic mixture probability of 92% to Consumption by Children and 8% to Self-Control and Goals for Lan Nguyen Chaplin and Debbie Roedder John's 2007 article (volume 34, issue 4), "Growing Up in a Material World: Age Differences in Materialism in Children and Adolescents." We could theoretically create a crude draft of the abstract using our model. (Unfortunately it would lack connecting, common, and idiosyncratic words while also being in random order and thus completely lack meaning. Researchers do not need to worry yet about their jobs being taken over by algorithms.) To create this crude draft of the abstract we would draw words randomly from the Consumption by Children and Self-Control and Goals topics with a 92:8 mixture.

Consumer Research Topics Over Time

After establishing the topic structure, we considered topic usage in the journal's history. We did this by considering how the topics featuring in the journal have changed over time. Figure 3 maps the changes in topic focus since 1974, plotting time on the x-axis and the topic's relative weight on the y-axis. (The topic's weight is the product of the number of articles focused on each topic, multiplied by each article's precise mix of topics.) Several trends stand out. There has been a clear movement over time away from articles focused on methodological issues. The early years of *JCR* saw a considerable focus on methodology that steadily declined. After a peak in the 1980s, research on family decision making has declined, as has research on consumption by children. Advertising research peaked early in *JCR's* history, and while this research is still represented, the peak seems long past. Memory and persuasion rose in importance in the 1980s and 1990s but seem to show evidence of a small relative loss in recent years. Topics such as contextual effects, resource constraints, and satisfying customers have all risen in importance, although it is possible that the high point of these research streams has already been passed. One other notable trend has been the steady rise of research on consumer culture from the 1980s to the present day.

Where is the field of consumer research heading? We believe past data can help us project trends and hazard some predictions. Assuming that one is happy extrapolating, one might conjecture that social identity and influence research's recent rise may make it a healthy stream of research in coming years. This topic has shown evidence of a significant increase in research focus in the last 20 years. Similarly, the past few years have seen an increased focus on self-control and goals that remained at a relatively high level in 2014. But the last few years have seen some year-on-year declines in some topics, which may suggest that their high watermark has been reached. Despite a possible minor downturn one might presume, given that topics do not seem to fall dramatically off the research agenda, we still might expect to see considerable work in the area of self-control and goals in coming years. Emotional decision making has seen a considerable rise in the recent past with little evidence of any decline, suggesting that the role of emotions in the consumer decision process may remain an interesting topic for the coming years. Consumer culture research has experienced considerable growth since the 1980s and seems poised to flourish in the future.

ARTICLE IMPACT: CITATION ANALYSIS

An understanding of which individual *JCR* articles have made the greatest impact may be acquired by evaluating the level of citations that the article received. There are a number of ways of assessing the citations for an article: total citations, average citations per year, and standardized burstness. The standardized burstness measures the number of citations in the five years after an article is published, correcting for the number of citations for all articles published in that year. This compensates for the general trend toward increasing citations per year over time. We determined the top 30 articles using each of these three measures. (This gives a total of 47 articles, rather than 90, because many of the top-cited articles also had the greatest average number of citations per year, or a high burstness, and so they were featured on more than one list.)

First, we analyzed the total number of citations for each article (shown in the second column of table 2). This analysis shows us which articles have had the most academic impact over the entire period of the *JCR*. The strength of this approach is that it highlights the articles that have stood the test of time. This shows that Russell Belk's article "Possessions and the Extended Self," which we have highlighted in table 2, is the most highly cited *JCR* article. Not surprisingly, the most highly cited articles tend to be older articles because this measure rewards those articles that have had more time to accumulate citations.

(Note that a blank in a column for a given article in table 2 means that the article was not in the top 30 for that unique measure of impact. For example, although the article by Steenkamp and Baumgartner was ranked third for total citations, it is not in the top 30 on the measure of standardized burstness.)

We next analyze citations per year. This measure identifies the articles that have been relatively impactful over the years since their publication; that is, this somewhat controls for the age of the article and so is relatively likely to reward newer articles (except, of course, for the very newest articles, because scholars citing those would not likely have had their own work published at the time of our analysis). The third column of table 2 shows us the average citations per year with the rank in brackets. Using this measure we find Zhao, Lynch, and Chen's 2010 article on mediation analysis, which we have highlighted in the table, has been remarkably impactful in the short time since its publication.

Finally, as recent years have seen more activity in the field, including more articles in *JCR*, it could be argued that it is relatively easier to gain higher average citations for newer articles. We therefore measure the standardized burstness of the articles to give a fair comparison between older and newer articles. (We omitted articles published after 2010 because this burstness measure is skewed for articles published very recently.) The standardized burstness, the final column in table 2, is the number of citations in the five years after an article is published, correcting for the number of citations for all articles in that year. We find Zhao et al. (2010) does well on this measure as well, outperforming by some distance the second place article using this measure.

Topics, Citations, and Impact

The next question we considered was which topics are associated with the top-cited articles. To perform this analysis we returned to the top-cited articles identified in table 2. We identified the topics associated with all 47 of these articles, that is, all the articles that were in either the top 30 articles in terms of top cites, average cites per year, or standardized burstness. Using a simple average of the weighting of the 47 top articles, we created the topics that feature in the average top-cited article. This average topic focus of all the top-cited articles is represented by the vertical bars in figure 4 that plots the topics on the x-axis and how much the topics feature on the y-axis. We then compared this average of the top-cited articles to the average topics in *JCR* over the entire period, the dashed lines in the figure.

There are a couple of noticeable differences between the top-cited articles and all articles. Consumer culture research and methodological issues tend to over index, in the sense that these topics make up a much greater level of heavily cited work than the average of all *JCR* articles. This suggests that topics such as methodological issues may be highly impactful even if they are now less typical in the journal. The intuition is probably that these article

TABLE 2

MOST CITED ARTICLES IN *JCR*, 1974–2014

Authors, Title, and Reference	Total cites (rank)	Cites per year (rank)	Standard burstness (rank)
Belk, RW. Possessions and the Extended Self (1988), 15 (2)	1305 **(1)**	48.3 **(6)**	6.0 **(7)**
Petty, RE; Cacioppo, JT; Schumann, D. Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement (1983), 10 (2)	1087 **(2)**	34 **(13)**	6.6 **(6)**
Steenkamp, JBEM; Baumgartner, H. Assessing Measurement Invariance In Cross-National Consumer Research (1998), 25 (1)	1064 **(3)**	62.6 **(3)**	
Zaichkowsky, JL. Measuring the Involvement Construct (1985), 12 (3)	1058 **(4)**	35.3 **(11)**	
Alba, JW; Hutchinson, JW. Dimensions of Consumer Expertise (1987), 13 (4)	1032 **(5)**	36.9 **(9)**	6.0 **(9)**
Green, PE; Srinivasan, V. Conjoint Analysis in Consumer Research: Issues and Outlook (1978), 5 (2)	1001 **(6)**	27.1 **(19)**	9.2 **(3)**
Holbrook, MB; Hirschman, EC. The Experiential Aspects of Consumption: Consumer Fantasies, Feelings, and Fun (1982), 9 (2)	979 **(7)**	29.7 **(15)**	4.5 **(15)**
Sheppard, BH; Hartwick, J; Warshaw, PR. The Theory of Reasoned Action: A Meta-Analysis of Past Research with Recommendations for Modifications and Future Research (1988), 15 (3)	943 **(8)**	34.9 **(12)**	
Fournier, S. Consumers and Their Brands: Developing Relationship Theory in Consumer Research (1998), 24 (4)	932 **(9)**	54.8 **(4)**	4.2 **(19)**
Jarvis, CB; Mackenzie, SB; Podsakoff, PM. A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research (2003), 30 (2)	888 **(10)**	74 **(2)**	6.8 **(5)**
Muniz, AM; O'Guinn, TC. Brand Community (2001), 27 (4)	691 **(11)**	49.4 **(5)**	5.1 **(11)**
Bettman, JR; Luce, MF; Payne, JW. Constructive Consumer Choice Processes (1998), 25 (3)	651 **(12)**	38.3 **(8)**	
Babin, BJ; Darden, WR; Griffin, M. Work and or Fun: Measuring Hedonic and Utilitarian Shopping Value (1994), 20 (4)	630 **(13)**	30 **(14)**	
Oliver, RL. Cognitive, Affective, and Attribute Bases of the Satisfaction Response (1993), 20 (3)	596 **(14)**	27.1 **(18)**	
Friestad, M; Wright, P. The Persuasion Knowledge Model: How People Cope with Persuasion Attempts (1994), 21 (1)	565 **(15)**	26.9 **(21)**	3.7 **(28)**
Bolton, RN; Drew, JH. A Multistage Model of Customers' Assessments of Service Quality and Value (1991), 17 (4)	556 **(16)**	23.2 **(23)**	
Richins, ML; Dawson, S. A Consumer Values Orientation for Materialism and Its Measurement: Scale Development and Validation (1992), 19 (3)	556 **(16)**	24.2 **(22)**	**3.9 (25)**
Simonson, I. Choice Based on Reasons: The Case of Attraction and Compromise Effects (1989), 16 (2)	502 **(18)**	19.3 **(28)**	
Kassarjian, HH. Content-Analysis in Consumer Research (1977), 4 (1)	500 **(19)**		
Celsi, RL; Olson, JC. The Role of Involvement in Attention and Comprehension Processes (1988), 15 (2)	498 **(20)**	18.4 **(30)**	
McCracken, G. Culture and Consumption: A Theoretical Account of the Structure and Movement of the Cultural Meaning of Consumer-Goods (1986), 13 (1)	479 **(21)**		
Huber, J; Payne, JW; Puto, C. Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis (1982), 9 (1)	462 **(22)**		
Arnould, EJ; Thompson, CJ. Consumer Culture Theory (CCT): Twenty Years of Research (2005), 31 (4)	453 **(23)**	45.3 **(7)**	9.4 **(2)**
Brucks, M. The Effects of Product Class Knowledge on Information Search Behavior (1985), 12 (1)	442 **(24)**		
Arnould, EJ; Price, LL. River Magic: Extraordinary Experience and the Extended Service Encounter (1993), 20 (1)	442 **(24)**	20.1 **(26)**	
Calder, BJ; Phillips, LW; Tybout, AM. Designing Research for Application (1981), 8 (2)	437 **(26)**		5.7 **(10)**
Belk, RW; Wallendorf, M; Sherry, JF. The Sacred and the Profane in Consumer Behavior: Theodicy on The Odyssey (1989), 16 (1)	437 **(26)**		5.1 **(13)**
Shiv, B; Fedorikhin, A. Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making (1999), 26 (3)	436 **(28)**	27.3 **(17)**	
Zhao, Xinshu; Lynch, John G., Jr.; Chen, Qimei. Reconsidering Baron and Kenny: Myths and Truths About Mediation Analysis (2010), 37 (2)	430 **(29)**	86 **(1)**	21.8 **(1)**
Bettman, JR; Park, CW. Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision-Processes: A Protocol Analysis (1980), 7 (3)	418 **(30)**		4.6 **(14)**

TABLE 2 (*CONTINUED*)

Authors, Title, and Reference	Total cites (rank)	Cites per year (rank)	Standard burstness (rank)
Goldstein, Noah J.; Cialdini, Robert B.; Griskevicius, Vladas. A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation In Hotels (2008), 35 (3)		35.3 **(10)**	7.4 **(4)**
Watts, Duncan J.; Dodds, Peter Sheridan. Influentials, Networks, and Public Opinion Formation (2007), 34 (4)		29.5 **(16)**	5.1 **(12)**
Holt, DB. Why Do Brands Cause Trouble? A Dialectical Theory of Consumer Culture and Branding (2002), 29 (1)		27 **(20)**	4.1 **(22)**
Aaker, JL; Lee, AY. “I” Seek Pleasures and “We” Avoid Pains: The Role of Self-Regulatory Goals in Information Processing and Persuasion” (2001), 28 (1)		22.1 **(24)**	
Vohs, Kathleen D.; Faber, Ronald J. Spent Resources: Self-Regulatory Resource Availability Affects Impulse Buying (2007), 33 (4)		20.4 **(25)**	6.0 **(8)**
Berger, Jonah; Heath, Chip. Where Consumers Diverge from Others: Identity Signaling and Product Domains (2007), 34 (2)		19.6 **(27)**	4.4 **(16)**
Schouten, JW; McAlexander, JH. Subcultures of Consumption: An Ethnography of the New Bikers (1995), 22 (1)		18.9 **(29)**	
Bettman, JR; Kakkar, P. Effects of Information Presentation Format on Consumer Information Acquisition Strategies (1977), 3 (4)			4.2 **(18)**
Greenwald, AG; Leavitt, C. Audience Involvement in Advertising: 4 Levels (1984), 11 (1)			4.4 **(17)**
Johnson, EJ; Russo, JE. Product Familiarity and Learning New Information (1984), 11 (1)			4.1 **(20)**
Zajonc, RB; Markus, H. Affective and Cognitive-Factors in Preferences (1982), 9 (2)			3.7 **(27)**
Ryan, MT; Bonfield, EH. The Fishbein Extended Model and Consumer-Behavior (1975), 2 (2)			3.8 **(26)**
Sujan, M. Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgments (1985), 12 (1)			4.1 **(21)**
Muniz, AM; Schau, HJ. Religiosity in the Abandoned Apple Newton Brand Community (2005), 31 (4)			4.0 **(23)**
Firat, AF; Venkatesh, A. Liberatory Postmodernism and the Reenchantment of Consumption (1995), 22 (3)			4.0 **(24)**
Edell, JA; Staelin, R. The Information Processing of Pictures in Print Advertisements (1983) 10 (1)			3.6 (29)
Fitzsimons, GM; Chartrand, TL; Fitzsimons, GJ. Automatic Effects of Brand Exposure on Motivated Behavior: How Apple Makes You Think Different (2008), 35 (1)			3.5 (30)

are high risk; they are less likely to get published but highly impactful if they are published.

TOP *JCR* SCHOLARS AND WHERE THEY COME FROM

The Most Productive *JCR* Scholars

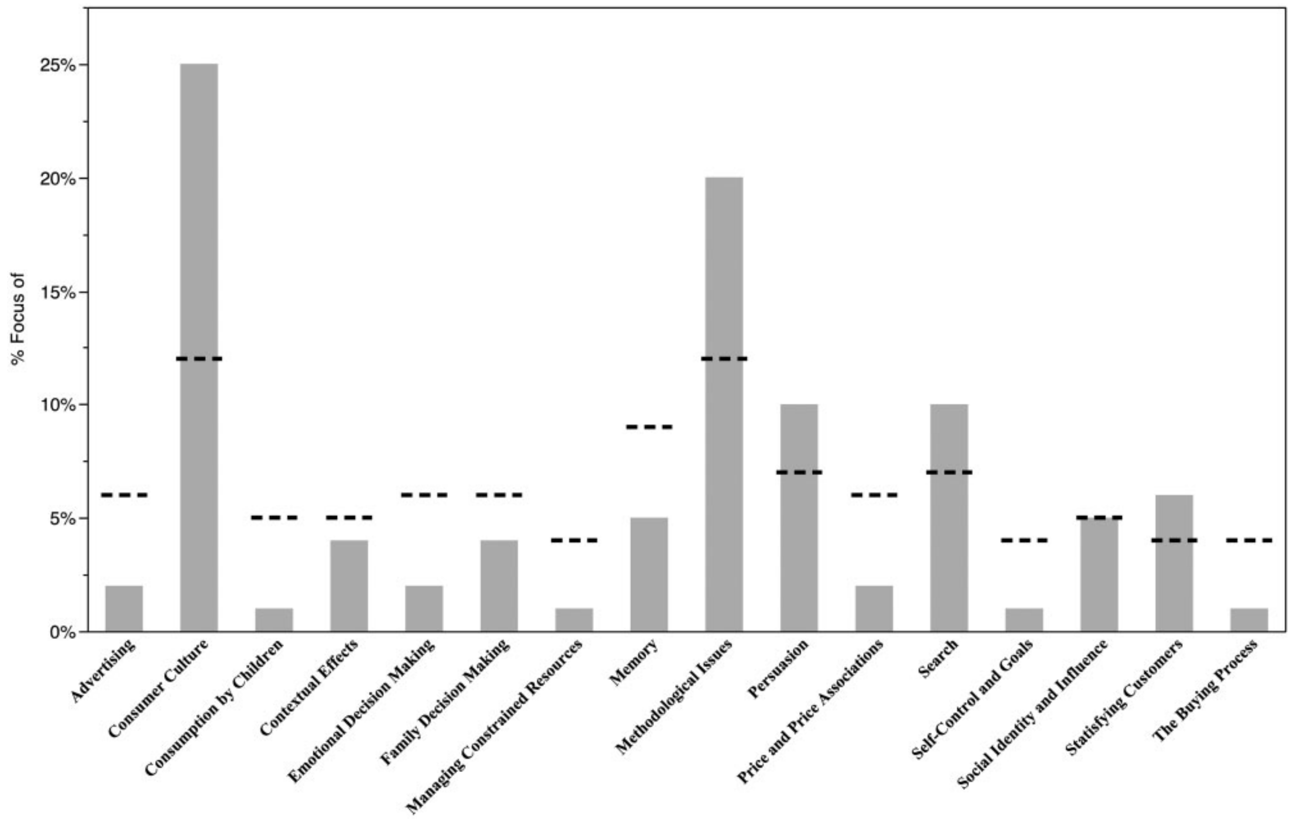
Who are the most published authors in *JCR*? We considered all the articles published in the 40-year period to April 2014 and found a total of 1784 scholars who have published in the journal. The median and modal publication rate was 1, with an average rate of 2.2 given the skew caused by the most productive scholars. We then determined the top scholars, which we defined as those scholars with 10 or more published articles in the 40-year period. This gave 45 scholars, or 2.5% of those who have published in *JCR*. Table 3 shows the number of articles published by the top researchers. The most successful scholar, in the sense of articles published, is Chris Janiszewski, with an astonishing 29 articles, closely followed by Russell Belk with 27.

When looking at the table, note that we considered only full articles with abstracts, which excluded some scholars from the list. In the earlier days of *JCR*, contributions often included notes without abstracts. Some highly successful scholars would be added to the list if we were to count research notes in our data set, including Gerald Gorn, Marvin Goldberg, and Alice Tybout. Furthermore, our data cover the first 40 years of the journal. Some scholars have reached 10 articles after the April 2014 issue that marked the end of the 40 years. Any analysis that continued to the end of 2014 would include such currently active scholars as Angela Lee, Jonah Berger, and Juliano Laran.

If we concentrate on the last decade of *JCR* (actually from the March 2004 to the April 2014 issue) to see who has been the most productive, we can see 896 scholars have been published in the journal. We consider the top scholars those who have published seven or more articles, which totals 26 scholars, or 2.9%, of those publishing. Many of the same scholars appear as in the list from the

FIGURE 4

TOPICS OF TOP-CITED ARTICLES (BARS) VERSUS ALL ARTICLES (DASHED LINES)



entire history, but some more recently emerged talents have been added. We have shown these new authors in boldface type in table 3A.

Training Future Consumer Researchers

In addition to identifying the impact of individual scholars, we were also interested in institutional excellence represented in *JCR*. Specifically, we asked, What schools have made the greatest contribution to the field of consumer research? One approach is to consider the schools where researchers work when completing their research. This is technically difficult to derive from journal data, given that researchers often move over their careers. Furthermore, historians can address this question from already provided outside rankings, for example, the University of Texas at Dallas ranking of research productivity. Therefore, we decided to determine those schools that have made the greatest contribution to the field not as the schools where scholars work, but the schools that have trained the top scholars of the future. To do this, we examined where those researchers on our top scholars list received their PhD training. Figure 5 plots the average number of articles each scholar on this list produced on the x-axis and the number of top scholars produced by a given school on the y-axis.

Many of the most prestigious universities also train much of the top talent. This is not surprising; prestigious schools often have relative large PhD programs and are able to attract top talent both in terms of the students they admit and the faculty advising the students. The epitome of this is Northwestern University, which produced the highest number of scholars on the list (seven). The top scholars from Northwestern also produced an impressive average number of articles between them. (Note this is not the average for Northwestern alumni, but the average for Northwestern alumni who are also top scholars.) From this data we cannot say conclusively that Northwestern has trained scholars who have produced the most articles in *JCR,* as it is possible that another school may have contributed more articles by having a greater number of productive scholars who missed out on our top scholar list. But using our methodology, the closest competition to

TABLE 3

TOP AUTHORS OVER THE HISTORY OF *JCR*

Scholar	Articles	Scholar, cont.	Articles, cont.
Chris Janiszewski	29	Vicki G. Morwitz	12
Russell W. Belk	27	Stephen M. Nowlis	12
Morris B. Holbrook	21	C. Whan Park	12
Craig J. Thompson	20	Eric J. Arnould	11
James R. Bettman	18	Ravi Dhar	11
Joan Meyers-Levy	18	Elizabeth C. Hirschman	11
William O. Bearden	17	Joel Huber	11
Barbara E. Kahn	17	Frank R. Kardes	11
John G. Lynch Jr.	17	Donald R. Lehmann	11
Alexander Chernev	15	Michel Tuan Pham	11
Laura A. Peracchio	15	Itamar Simonson	11
Baba Shiv	15	Robert S. Wyer Jr.	11
Brian Sternthal	15	Paul E. Green	10
Joseph W. Alba	14	Stephen J. Hoch	10
Gita V. Johar	14	Jacob Jacoby	10
Deborah Roedder John	14	Aradhna Krishna	10
Jennifer J. Argo	13	Jaideep Sengupta	10
Gavan J. Fitzsimons	13	John F. Sherry Jr.	10
Durairaj Maheswaran	13	Terence A. Shimp	10
Jennifer L. Aaker	12	Melanie Wallendorf	10
Darren W. Dahl	12	Peter Wright	10
Valerie S. Folkes	12	Rui (Juliet) Zhu	10
David Glen Mick	12		

TABLE 3A

TOP AUTHORS OVER THE LAST DECADE OF *JCR*

Scholar	Articles	Scholar, cont.	Articles, cont.
Chris Janiszewski	17	Stephen M. Nowlis	8
Jennifer J. Argo	12	Jaideep Sengupta	8
Alexander Chernev	12	Jennifer L. Aaker	7
Darren W. Dahl	11	Russell W. Belk	7
Baba Shiv	11	**Aparna A. Labroo**	7
Craig J. Thompson	11	**Angela Y. Lee**	7
Gavan J. Fitzsimons	10	**Naomi Mandel**	7
Rui (Juliet) Zhu	10	**Ann L. McGill**	7
Jonah Berger	9	Vicki G. Morwitz	7
Gita V. Johar	9	Laura A. Peracchio	7
Aradhna Krishna	9	Michel Tuan Pham	7
Juliano Laran	9	**Kathleen D. Vohs**	7
Anirban Mukhopadhyay	8	Robert S. Wyer Jr.	7

Northwestern is Columbia, which has the joint second greatest number of top scholars (3) and a high average number of articles per top scholar (13.7). A number of traditionally strong schools each have an impressive showing including Illinois, Duke, and Pennsylvania. Minnesota and Tennessee also stand out, with these schools each benefiting greatly from an individual outlier, Russell Belk and Craig Thompson, respectively.

A total of 26 schools feature as having produced top scholars. Schools in the United States feature heavily as the training grounds of the top scholars, with Canada the only other country featured, by virtue of the University of British Columbia and the University of Manitoba.

If we examine top scholars from only the last decade (those that published seven or more articles since early 2004), we find many of the same schools featured as shown in figure 5. Other schools that have trained scholars who have been relatively more productive in the last decade are Chicago, Cornell, Dartmouth, and another non-US school, Toronto.

CONCLUSION

The *Journal of Consumer Research* seeks to be the premier interdisciplinary journal for consumer researchers and the field of consumer research. In this article, we hope we have provided an increased understanding of this field as the journal looks back and assesses its progress at this important anniversary. Our research has uncovered the topics that have proved important in the field over the past 40 years. Using text mining and probabilistic topic modeling (LDA), we were able to determine the breadth of topics that *JCR* articles consider. We are also able to determine what makes for an impactful article and what topics have risen or declined during the 40 years of *JCR*.

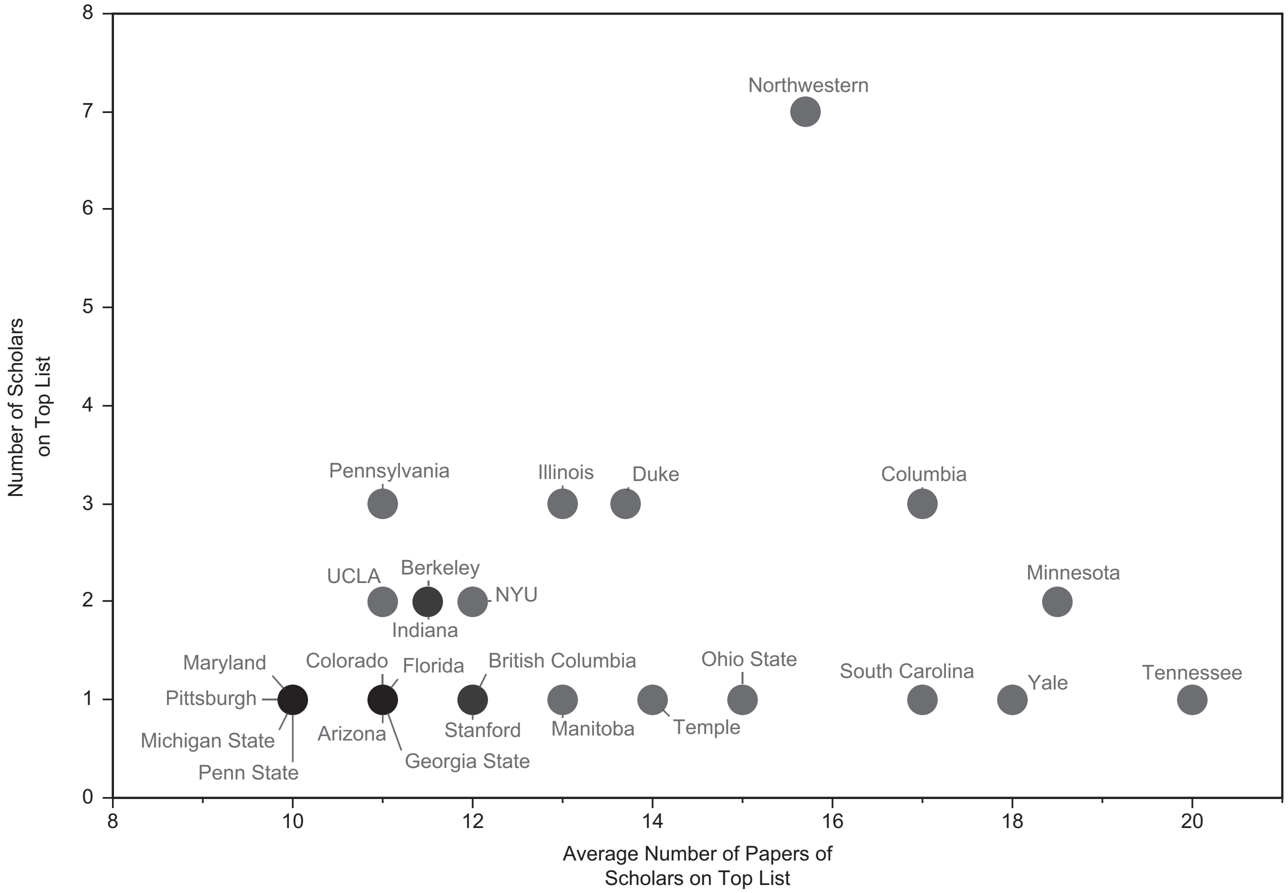
JCR's research focus in an individual article appears to have narrowed, and yet the entire field has diverged, considering more varied ideas and theories in the last couple of decades, and particularly in the last 10 years. *JCR* research has become more theoretical and more academic in tone, with research on self-control and goals, emotional decision making, and consumer culture issues having become popular topics in recent years. We noted the interesting trends of proliferation of research topics in recent years and yet the return to some styles of research that predominated in the early days of the journal, such as field studies of consumer behavior.

We have considered the top scholars in the field and noted the large skew in productivity. Most scholars only achieve one *JCR* article in their entire careers, but some achieve many more. Indeed four scholars have achieved 20 articles or more. Given that the field of consumer research would become stagnant without a pipeline of well-trained and productive scholars, we considered which schools have been most critical in training these top scholars. We found that some schools made an especially large impact on the field, with Northwestern having produced 7 of the top 45 scholars, or 16% of the total.

With the world's research available to anyone armed with a search term or two, anecdotal evidence suggests that full cover-to-cover readership of any one academic journal may be declining. With a narrow and specialized search, researchers may gain focus at the expense of breadth. Or perhaps that is always the case as a field reaches maturity and begins to diverge into specialties. At this important juncture of *JCR*'s history, we suggest to readers that the journal itself, taken as a whole, provides a rich source of insight for those who wish to understand the messy, complicated, and wonderful field of consumer research.

FIGURE 5

SCHOOLS THAT TRAINED *JCR* TOP SCHOLARS



DATA COLLECTION INFORMATION

The first and third authors downloaded the abstracts of all articles published in *JCR* from its inception in June 1974 through to April 2014. For each article our data come from the following two sources: (1) JSTOR Data for Research (dfr.jstor.org) provided us with article metadata such as Title, Abstract, Author(s), Volume, Issue, and Publication Date. In addition, we obtained word frequencies, key terms, and ngrams. (2) Thomson Reuters Web of Science Citation helped us gather data for each article by year. All authors then analyzed the data using a variety of programs including SAS JMP and Excel.

REFERENCES

Blei, David M. (2012), “Probabilistic Topic Models,” *Communications of the ACM,* 55 (4), 77–84.

Dahl, Darren, Eileen Fischer, Gita V. Johar, and Vicki Morwitz (2014), “From the Editors-Elect: Meaningful Consumer Research,” *Journal of Consumer Research,* 41 (1), iii–v.

Deighton, John (2007), “From the Editor: The Territory of Consumer Research: Walking the Fences,” *Journal of Consumer Research,* 34 (3), 279–82.

Dingus, Rebecca, Hillary N. Mellema, and Ryan J. Langan (2013), “2013 Who Went Where? Survey Results,” *American Marketing Association* report. http://docsig.org/who-went-where/. Accessed April 30, 2015.

Frank, Ronald E. (1995), “Notes on the Journal of Consumer Research: The Unexpected Challenges of a Start-up,” in *Advances in Consumer Research,* Vol. 22, ed. Frank R. Kardes and Mita Sujan, Provo, UT: *Association for Consumer Research*, 486–87.

Griffiths, Tom and Mark Steyvers (2004), “Finding Scientific Topics,” *Proceedings of the National Academy of Sciences,* 101 (Suppl 1), 5228–235.

Huber, Joel, Wagner Kamakura, and Carl F. Mela (2014), “A Topical History of JMR,” *Journal of Marketing Research*, 51 (1), 84–91.

Journal of Consumer Research (2014), “Instructions to Reviewers,” retrieved August 31, 2014, http://www.ejcr.org/instr-revs.htm.

Katona, George (1974), “Psychology and Consumer Economics,” *Journal of Consumer Research*, 1 (1), 1–8.

Mela, Carl F., Jason Roos, and Yiting Deng (2013), “Invited Paper—A Keyword History of Marketing Science,” *Marketing Science,* 32 (1), 8–18.

Mellema, Hillary N., Paul Mills, and Alexa Fox (2014), “Who Went Where 2014? Survey Results,” *American Marketing Association report*. http://docsig.org/who-went-where/. Accessed April 30, 2015.

Web Appendix

The *Journal of Consumer Research* at Forty: A Historical Analysis

XIN (SHANE) WANG
NEIL BENDLE
FENG MAI
JUNE COTTE

Topic Models

Topic modeling is used to automatically discover the index of ideas contained in the documents and identify which documents are about the same kinds of ideas (Blei and Lafferty 2009). The statistical topic model we use, latent Dirichlet allocation (LDA) (Blei, Ng, and Jordan 2003), assumes that the procedure of producing a JCR abstract can be decomposed into a number of simple probabilistic steps. The statistical inference based on hierarchical Bayesian analysis can then uncover the semantic structures in the texts and discover patterns of word usage. LDA has been shown to be a powerful technique for gaining insight into the content of academic documents (Griffiths and Steyvers 2004).

There are many advantages provided by the LDA model compared to more traditional text analytic models such as naïve Bayes classification and Latent Semantic Analysis (LSA). For example, LDA is built upon a rigorous foundation of Bayesian statistical inference and therefore has more principled model fitting and selection procedures. It provides "soft" classification for documents and therefore allows each document to be a multi-membership mixture of different topics. LDA also extends the ideas of probabilistic latent semantic analysis (PLSA) (Blei and Lafferty 2009) and can automatically learn contexts of word usage without recourse to a dictionary or thesaurus (Hofmann 2001).

There are several statistical assumptions inherent in the LDA model. The first major assumption of the LDA is the "bag of words" model, which means that the words appearing in the abstracts are assumed to be exchangeable. Therefore, when applying statistical topic models such as LDA, we represent each abstract as a vector of word counts and neglect the order of the words. Although the order of the words is important for human readers to comprehend a document, Blei et al. (2003) have argued, using De Finetti (1977)'s exchangeability theorem, that this simple representation can result in computationally efficient methods while preserving semantic themes in the collection of documents.

Given the “bag of words” assumption, the LDA model further assumes that: 1) words contained in each abstract are generated from a mixture of topics; 2) each topic has a probability distribution over a fixed word vocabulary; 3) the topics are shared by all of the JCR abstracts, but the topic proportions differ across abstracts. Formally, LDA can be described using a generative process. It assumes that there are K different topics (the parameter K can be chosen using model selection techniques) and the vocabulary size is V. Each topic is associated with a Dirichlet distribution over all words in the vocabulary with parameters $\boldsymbol{\beta}$. For all topics $k \in 1 \dots K$ the process first draws a vocabulary mixture $\boldsymbol{\phi}_k$ for the topic from Dirichlet ($\boldsymbol{\beta}$). Then, each JCR abstract $m \in 1 \dots M$ is assumed to be produced from the following generative process:

1. Sample length of the abstract N_m from a Poisson distribution with parameter ξ.
2. Sample topic proportions $\boldsymbol{\theta}_m$ from a Dirichlet distribution with parameters $\boldsymbol{\alpha}$.
3. For each of the $n \in 1 \dots N$ words in m:
    a. Sample a topic assignment $z_{m,n}$ from Multinomial ($\boldsymbol{\theta}_m$), where $z_{m,n}$ is a topic index between $1 \dots K$.
    b. Choose a word $w_{m,n}$ from Multinomial ($\boldsymbol{\phi}_{z_{m,n}}$).

The objectives of topic modeling can be viewed as reversing the above generative process using Bayesian inference (Blei 2012). We wish to infer the topic mixture of each abstract $\boldsymbol{\theta}_m$, and the word distributions of each topic $\boldsymbol{\phi}_k$. The former parameters indicate which topic(s) are covered in a given abstract, while the latter parameters tell us the representative words for each topic. Researchers have developed approximate inference algorithms such as Gibbs sampling (Steyvers and Griffiths 2006) and variational methods (Blei et al. 2003; Teh, Newman, and Welling 2006) as exact inference is intractable for the model. Heinrich (2005) presents a detailed discussion of various parameter estimation methods for LDA.

Implementation

We downloaded the abstracts of the 1875 JCR articles from JSTOR Data for Research (dfr.jstor.org). We completed sentence and word segmentation, part-of-speech (POS) tagging, and word lemmatization using Stanford CoreNLP (Manning et al., 2014).

We used the Stanford Topic Modeling Toolbox (ver. 0.4.0) developed and distributed by the Stanford Natural Language Processing Group (Ramage et al. 2009) for the inference task. To account for the power-law of word usage, and avoid the domination of common words across topics (McCallum, Mimno, and Wallach 2009), we excluded common stopwords. In addition, we excluded the top 40 most frequent words, outside of the stopwords list, that are most common in the JCR abstracts. These words include methodology related words such as *examine*, *method* and *study*, and words that are less topic-specific but are widely used in many topics such as *behavior*, *brand*, and *people*. We take these preprocessing steps to reduce noise and improve the interpretability of the resulting topics. A robustness check was performed by fitting the model without excluding these words and the analysis resulted in qualitative similar topics.

We used the collapsed variational Bayes (CVB0) approximation outlined in Asuncion et al. (2009) to approximate the posterior distribution. Two hyperparameters α and β in the LDA model control the smoothing for document-topic distributions and topic-term distributions respectively. A smaller β generates more fine-grained topics and a smaller α tends to assign fewer topics to a document. We used $\beta = 0.01$ and $\alpha = 50/K$ following the recommendation of Griffiths and Steyvers (2004)[1]. The optimal number of topics K was chosen to minimize perplexity, a widely-used performance metric that gives useful characterization of the predictive quality of a language model and correlates with other measures well (Asuncion et al. 2009). More specifically, we trained LDA models with $K = 2$ to 24 on half of the data with 2,000 iterations of CVB0 algorithm and evaluated the perplexity on the other half of the data. We chose $K = 16$ because it offered the minimal perplexity on the test set (Figure 1). The final results presented in the paper were produced using 10,000 iterations of the CVB0 algorithm with $K = 16$ on all the data.

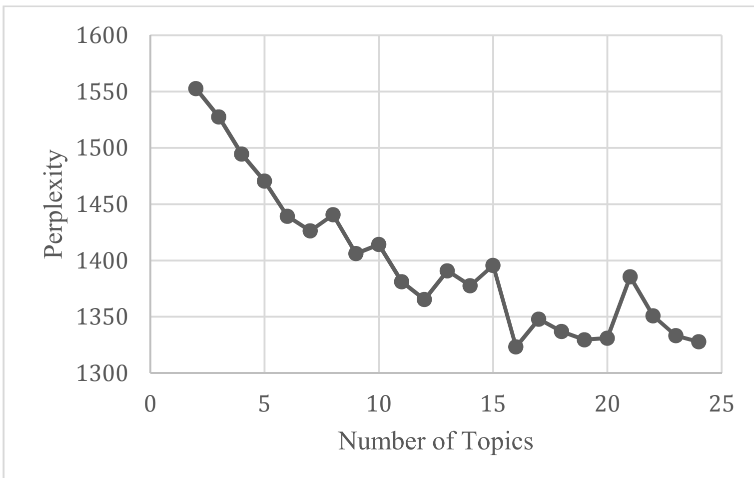


Figure 1. Perplexity as a function of number of topics

Lastly, we estimated the consumer research topics distribution over time using the Hall model (Hall, Jurafsky, and Manning. 2008), in which *post hoc* calculations are performed based on the observed probability of topics over the years. The Hall model is non-restrictive on the overall trend of the topics and thus offers the flexibility needed for our exploratory analysis. The empirical probability that an arbitrary abstract d written in year y was about topic z is:

$$\hat{p}(z|y) = \sum_{d:t_d=y} \hat{p}(z|d)p(d|y)$$

where $\hat{p}(z|d)$ is the estimated document-topic distribution, and $p(d|y)$ is the proportion of the articles written in year y.

[1] α and β are parameter vectors for Dirichlet distribution. We follow the convention of using symmetric Dirichlet priors, i.e., each entry is the same scalar.

References

Asuncion, Arthur, Max Welling, Padhraic Smyth, , and Yee Whye Teh (2009). *On Smoothing and Inference for Topic Models.* Paper presented at the Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence.

Blei, David M. (2012). Probabilistic Topic Models. *Communications of the ACM, 55*(4), 77-84.

Blei, David M, and John D. Lafferty (2009) Topic models. *Text Mining: Classification, Clustering, and Applications, 10*, 71.

Blei, David M, Andrew Y. Ng, and Michael I. Jordan (2003). Latent Dirichlet Allocation. *the Journal of Machine Learning Research, 3*, 993-1022.

De Finetti, Bruno. (1977). *Theory of Probability, volume I.*

Griffiths, Thomas L, and Mark Steyvers (2004). Finding Scientific Topics. *Proceedings of the National Academy of Sciences of the United States of America, 101* (Suppl 1), 5228-235.

Hall, David, Daniel Jurafsky, and Christopher D. Manning (2008). Studying the History of Ideas Using Topic Models. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 363-371.

Heinrich, Gregor. (2005). Parameter Estimation for Text Analysis.

Hofmann, Thomas. (2001). Unsupervised Learning by Probabilistic Latent Semantic Analysis. *Machine Learning, 42*(1-2), 177-196.

Manning, Christopher D, Surdeanu, Mihai, Bauer, John, Finkel, Jenny, Bethard, Steven J, & McClosky, David. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics System Demonstrations*, 55-60.

McCallum, Andrew, David M. Mimno, and Hanna M. Wallach (2009). *Rethinking LDA: Why Priors Matter.* Paper presented at the Advances in Neural Information Processing Systems.

Ramage, Daniel, Evan Rosen, Jason Chuang, Christopher D. Manning, and Daniel A. McFarland, (2009). *Topic Modeling for the Social Sciences.* Paper presented at the NIPS 2009 Workshop on Applications for Topic Models: Text and Beyond.

Steyvers, Mark, and Tom Griffiths (2006). Probabilistic Topic Models. In T. Landauer, D. McNamara, S. Dennis & W. Kintsch (Eds.), *Latent Semantic Analysis: A road to Meaning*: Lawrence Erlbaum.

Teh, Yee W, David Newman, and Max Welling (2006). *A Collapsed Variational Bayesian Inference Algorithm for Latent Dirichlet Allocation.* Paper Presented at the Advances in Neural information Processing Systems.