

Research Impact Metrics

A 50-Year Analysis of Education Research Article Feature Effects on Citation Counts

Royce Kimmons & Ross Larsen

Academic Writing

Research Impact

Bibliometrics

By analyzing 50 years of citation counts of 51,281 research articles across 86 education journals in conjunction with textual analysis of article titles and abstracts, we explore how a variety of article features, such as title length, use of a subtitle, reading difficulty, and open access status, have historically influenced the impact of education research articles. Results indicate that (a) shorter titles are more likely to be cited than long titles, (b) articles with subtitles (designated with a colon) are more likely to be cited, (c) articles with lengthy and more technical abstracts are more likely to be cited, and (d) open access status has no effect.

The guiding research question of this analysis was “What is the relationship between education research article features and citation counts?” Central to our asking this question is the notion that citation count as a measure of impact may be influenced by a variety of factors that may have little to do with a given study’s scientific or professional merit or that subtle decisions regarding an article’s title or abstract might influence its citability. To answer this question, we utilized hierarchical linear modeling (HLM) to analyze Scopus database metrics for top education research journals to determine the strengths of relationships between two independent citation variables, six independent article feature variables, and two covariates. In total, 51,281 articles from 86 journals were analyzed, inclusively representing the years 1969 to 2020 (see Table 1).

Table 1

List of Journals with Number of Articles and Year Ranges

Journal Title	Number of Articles	Year Range	
		Minimum	Maximum
Assessment & Evaluation in Higher Education	470	1981	2020
Australasian Journal of Educational Technology	783	2008	2020

Journal Title	Number of Articles	Year Range	
		Minimum	Maximum
Australian Journal of Teacher Education	357	2009	2014
British Journal of Educational Psychology	496	1969	2020
British Journal of Educational Technology	1,892	1973	2020
Cogent Education	744	2014	2020
Community College Journal of Research and Practice	345	1993	2020
Computer Assisted Language Learning	641	1990	2020
Computers & Education	3,204	1986	2020
Contemporary Educational Psychology	318	1980	2020
Counselor Education and Supervision	295	1969	2020
Education and Information Technologies	1,303	2000	2020
Education and Urban Society	306	1973	2020
Educational Administration Quarterly	553	1969	2020
Educational Assessment, Evaluation and Accountability	121	2013	2020
Educational Management Administration & Leadership	645	1974	2020
Educational Policy	399	1987	2020
Educational Psychologist	297	1970	2020
Educational Psychology	521	1981	2020
Educational Psychology Review	230	1990	2020
Educational Technology & Society	1,535	2000	2020
Educational Technology Research and Development	932	1993	2020
English Language Teaching	733	2012	2015
English Language Teaching Journal	1,060	1973	2020
European Journal of Teacher Education	861	1982	2020
Higher Education	446	1976	2020
Higher Education Research & Development	516	1982	2020
IEEE Transactions on Learning Technologies	458	2008	2020
Improving Schools	182	1998	2020
Innovations in Education and Teaching International	223	2015	2020
Innovative Higher Education	99	1996	2020
Interactive Learning Environments	1,043	1990	2020
International Journal of Artificial Intelligence in Education	283	2000	2020
International Journal of Doctoral Studies	63	2014	2020
International Journal of Educational Management	755	1987	2020
International Journal of Instruction	911	2013	2020

Journal Title	Number of Articles	Year Range	
		Minimum	Maximum
International Journal of Leadership in Education	344	2003	2020
Internet and Higher Education	540	2001	2020
Journal of College Student Development	142	2004	2020
Journal of College Student Retention: Research, Theory, & Practice	148	2010	2020
Journal of Computer Assisted Learning	953	1985	2020
Journal of Counseling & Development	415	1985	2020
Journal of Counseling Psychology	715	1975	2019
Journal of Diversity in Higher Education	129	2014	2020
Journal of Education Policy	421	1986	2020
Journal of Educational Administration	615	1969	2020
Journal of Educational Computing Research	490	2011	2020
Journal of Educational Psychology	497	1974	2020
Journal of Educational Research	1,864	1969	2020
Journal of Further and Higher Education	334	1977	2020
Journal of Higher Education Policy and Management	140	2009	2020
Journal of International Students	177	2015	2020
Journal of Mathematics Teacher Education	401	2005	2020
Journal of Psychoeducational Assessment	741	1983	2020
Journal of School Choice	203	2006	2020
Journal of School Psychology	213	1982	2020
Journal of Science Teacher Education	630	1989	2020
Journal of Studies in International Education	82	2004	2020
Journal of Teacher Education	908	1969	2020
Language Learning & Technology	361	2000	2020
Leadership and Policy in Schools	243	2014	2020
Learning and Individual Differences	751	2013	2020
Learning Disability Quarterly	252	1978	2020
Learning, Media and Technology	415	2005	2020
Management in Education	231	1987	2020
Phi Delta Kappan	695	1996	2020
Professional Development in Education	634	2009	2020
Psychology in the schools	736	1975	2020
Research in Higher Education	141	1974	2020
School Effectiveness and School Improvement	232	1990	2020

Journal Title	Number of Articles	Year Range	
		Minimum	Maximum
School Leadership & Management	190	2002	2020
School Psychology International	621	1979	2020
School Psychology Quarterly	55	2019	2020
Social Psychology of Education	306	2001	2020
Studies in Higher Education	736	1976	2020
Teachers and Teaching	793	1995	2020
Teaching and Teacher Education	2,484	1985	2020
Teaching in Higher Education	1,036	2005	2020
Technology, Pedagogy and Education	468	2003	2020
TechTrends	431	2000	2020
TESOL Quarterly	877	1981	2020
The International Review of Research in Open and Distributed Learning	1,054	2000	2020
The Journal of Higher Education	81	2015	2020
The Review of Higher Education	50	2015	2020
The Turkish Online Journal of Educational Technology	1,643	2008	2017
Theory Into Practice	643	1980	2020

Our independent citation variables consisted of two variations of the citation count metric provided by Scopus: (a) raw citations and (b) citations per year. Raw citations represented the total number of times that an article had been cited in its entire lifespan. As one might expect, these counts were somewhat influenced by publication date because it takes time for articles to be read and cited in subsequent publications, meaning that articles published earlier in a given year might exhibit a citation advantage over articles published later in the same year (see Figure 1). For this reason, we also recoded raw citation counts as citations per year by multiplying the citation count by 365 and dividing this value by the number of days that had elapsed since the article had been published (see Figure 2). This recoding helped control for elapsed time but also revealed a general positive relationship between year published and citations per year, suggesting that more recent articles were being cited at a higher rate than their predecessors. Uncertain of which of these two metrics would be the most reliable for accounting for complexities of time, we constructed separate models for each to see if results converged to tell a similar story.

Figure 1

Average Article Raw Citations by Year Published ($R^2 = 0.28$)

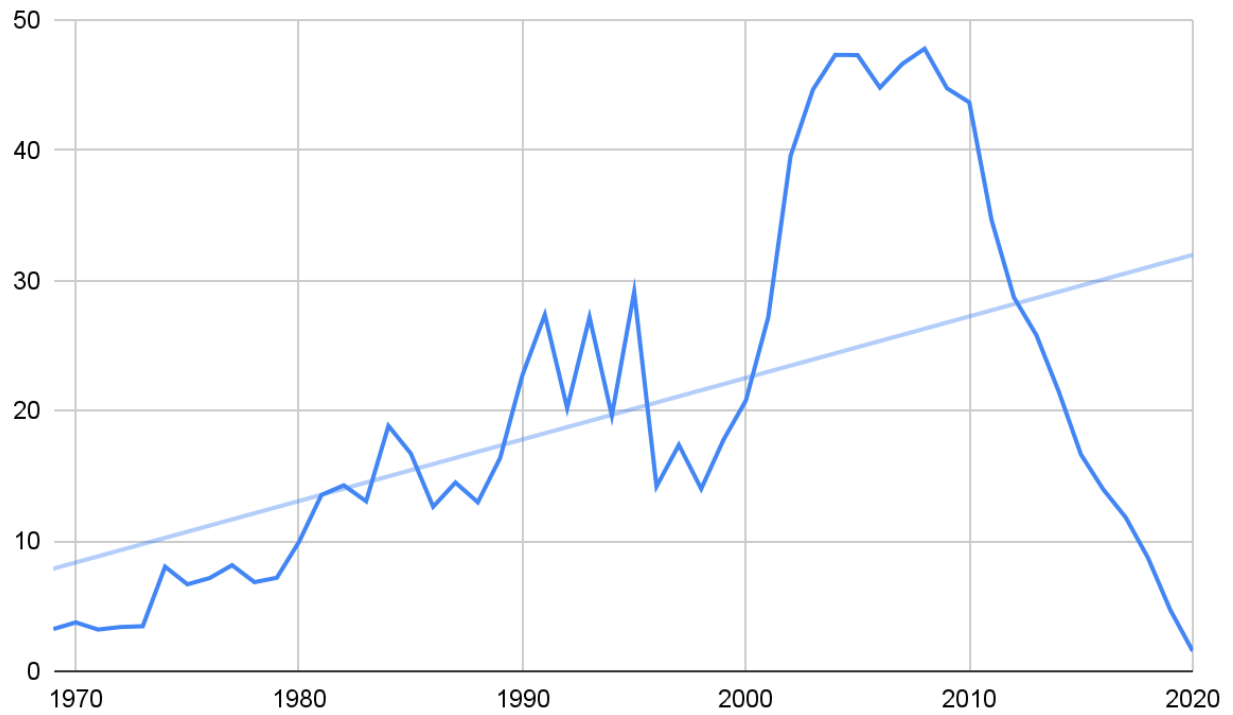
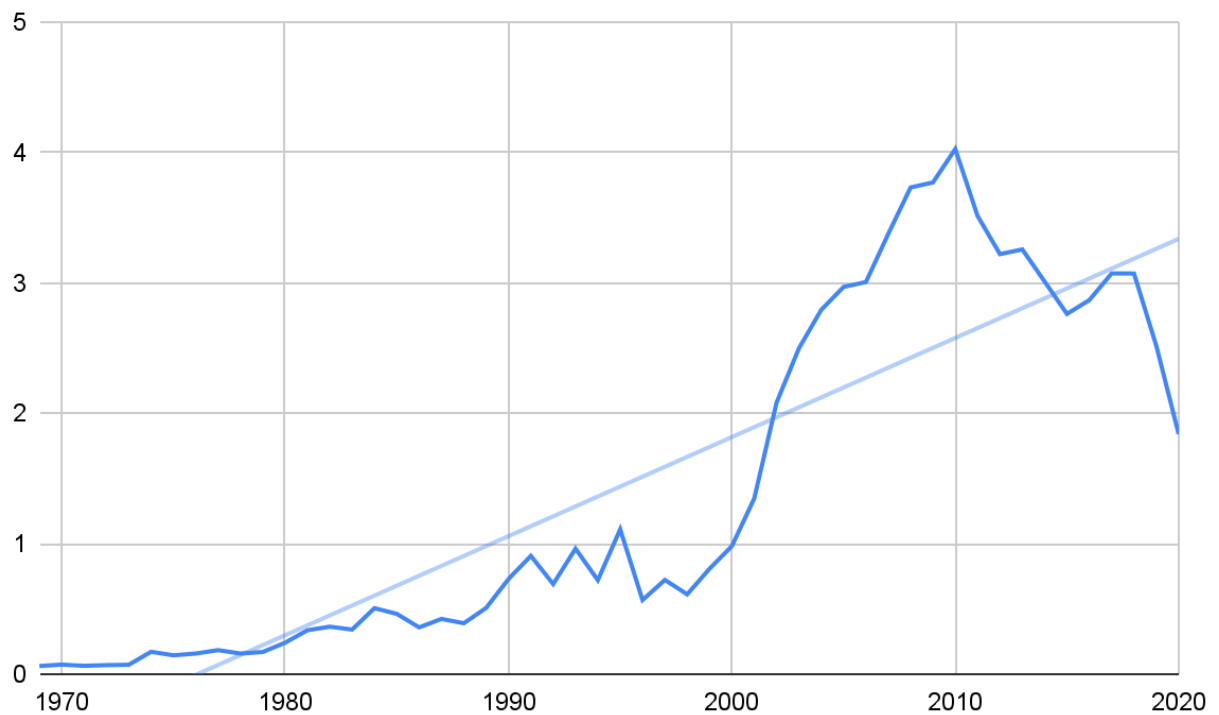


Figure 2

Average Article Citations per Year by Year Published ($R^2 = 0.78$)



Independent article features included the following six variables:

- Title Character Count: The number of characters (i.e., numbers, letters, or punctuation) in the article's title (see Table 2 for descriptives).
- Title Colon: Whether the title included a colon, thereby suggesting the presence of a subtitle (0 = no colon [n = 27,921] and 1 = colon present [n = 23,336]).
- Abstract Reading Difficulty: The Flesch-Kincaid Reading Ease score for the article's abstract (0 = very difficult to read and 100 = very easy to read; see Table 2 for descriptives).
- Abstract Reading Time: The predicted number of seconds needed for the average adult to read the abstract as calculated on a range from 150 words per minute for a Reading Ease score of 0 to 300 words per minute for a score of 100 (see Table 2 for descriptives).
- Abstract Word Count: The number of words in the abstract (see Table 2 for descriptives).
- Open Access: Whether the article was marked as released under an open access agreement (0 = non-open access [n = 44,663] and 1 = open access [n = 6,618]).

Table 2

Descriptives of Continuous Variables

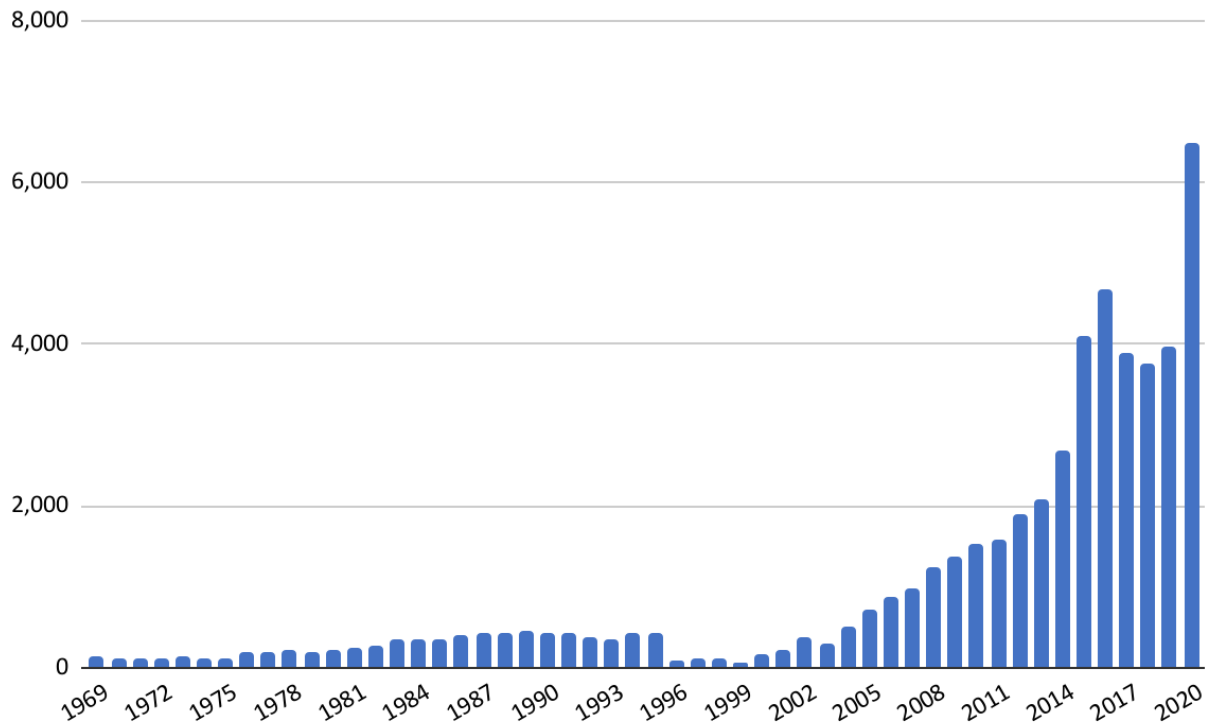
	Mean	SD	Min	Max
Title Character Count	92.107	30.990	6	255
Abstract Reading Ease	24.349	13.551	0	100
Abstract Reading Speed	52.755	20.434	1	459
Abstract Word Count	161.918	60.874	4	1,289

A year covariate was also included to better control for time-based effects on citation counts. Annual totals of articles revealed a general upward trend in article volume with a few notable exceptions between 1996 and 2003 (see Figure 3). The increase in article volume overall was likely due to more journals releasing online versions over time since the early 2000s (and thereby increasing the number of articles that could be published without the cost prohibitions of a paper-based medium), but it was unclear to us why a dip occurred in 1996. Nonetheless, we did not expect these variations in volume to impact results in a meaningful way but used year as a covariate to ensure that historical or other anomalies in the data would be accounted for. Furthermore, our models were constructed using M+ software, which preferred for these values to be normalized to small integers for greater ease in interpreting Betas and other values (e.g., 2012 = 2.012).

And finally, recognizing (a) that journals that have been publishing longer were being cited more on average than younger journals and (b) that journals that have been publishing longer had a lower percentage of open access articles, we also used the longevity of the journal as an additional covariate for our analysis. This further helped to control for journal characteristics outside the control of individual article authors that might be influencing citation counts, such as the perceived prestige of the journal in the field.

Figure 3

Distribution of Included Articles by Year



Results

Results indicated overall significant (but weak) effects on both raw citations ($R^2 = 0.022$, $p < .01$; see Table 3) and citations per year ($R^2 = 0.054$, $p < .001$; see Table 4). For raw citations, the model showed that articles would be cited more if their authors (a) shortened the title, (b) made the abstract more technical, (c) lengthened the abstract, and (d) included a colon in the title. For citations per year, the model showed that articles would be cited more if their authors (a) made the abstract more technical and (b) included a colon in the title. Furthermore, the size of the dataset allowed us to detect significant effects that had relatively small effect sizes, so the fact that reading time and open access status did not affect either result is also noteworthy.

Table 3

Article Feature Effects on Raw Citations

	Estimate	S.E.	Est./S.E.	Two-Tailed p Value
Model R-Square	0.022	0.008	2.624	0.009**
Title Character Count	-0.044	0.012	-3.71	0.000***
Title Colon	0.039	0.009	4.414	0.000***
Abstract Reading Ease	-0.117	0.02	-5.898	0.000***
Abstract Reading Time	-0.12	0.056	-2.15	0.032
Abstract Word Count	0.152	0.051	2.987	0.003**
Open Access	-0.023	0.018	-1.262	0.207
Year Covariate	-0.056	0.041	-1.373	0.17
Journal Longevity	0.075	0.039	1.926	0.054

Table 4*Article Feature Effects on Citations per Year*

	Estimate	S.E.	Est./S.E.	Two-Tailed <i>p</i> Value
Model R-Square	0.054	0.013	4.079	0.000***
Title Character Count	-0.023	0.013	-1.802	0.071
Title Colon	0.045	0.008	5.813	0.000***
Abstract Reading Ease	-0.062	0.02	-3.026	0.002**
Abstract Reading Time	0.027	0.061	0.447	0.655
Abstract Word Count	0.03	0.056	0.54	0.589
Open Access	0.001	0.021	0.051	0.959
Year Covariate	0.175	0.024	7.375	0.000***
Journal Longevity	0.164	0.048	3.41	0.001**

Discussion

Titles

Shorter titles were more likely to be cited than longer titles, but the inclusion of a colon (typically used in longer titles) also had a positive effect. This suggests to us that when writing titles, subtitles can be useful for improving citations but that authors should practice parsimony in the length of both the title and the subtitle. For articles without a colon in the title, there seems to be a Goldilocks zone of between 30 and 50 characters or 5 to 9 words for optimal length (see Figure 4). For articles with a colon, the Goldilocks zone appears to be slightly higher, between 40 and 70 characters or 7 to 12 words (see Figure 5).

Figure 4*Distribution of Average Citations by Title Length for Articles without Colons*

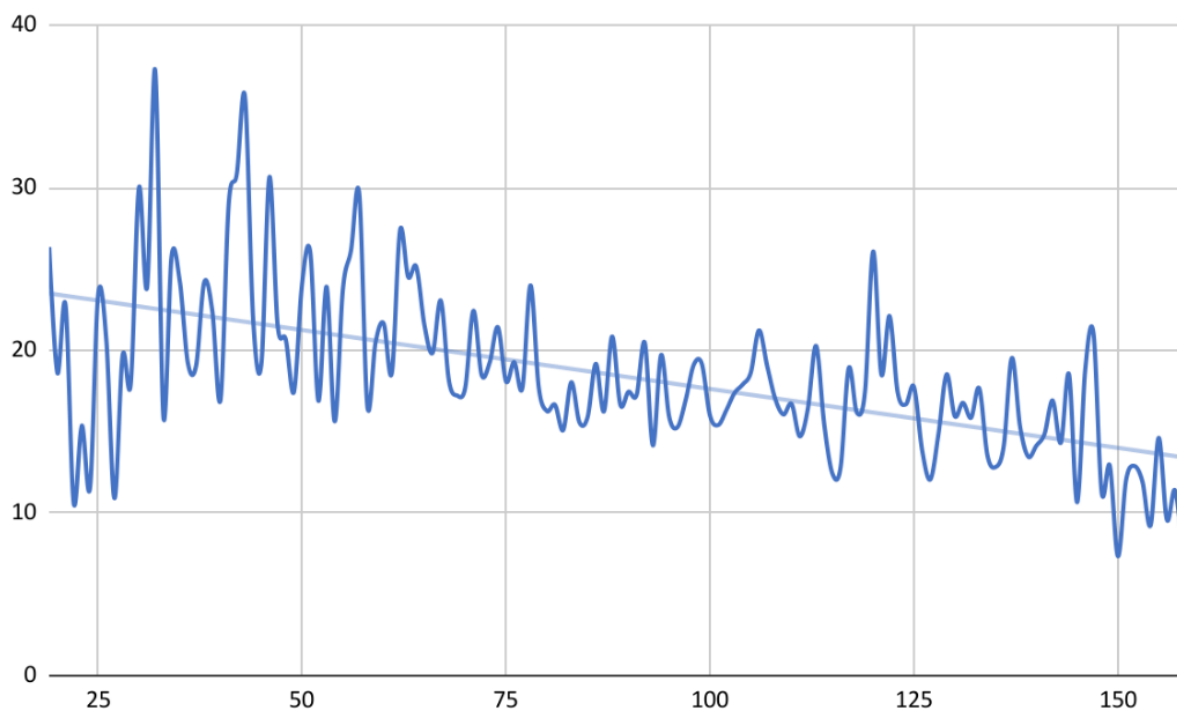
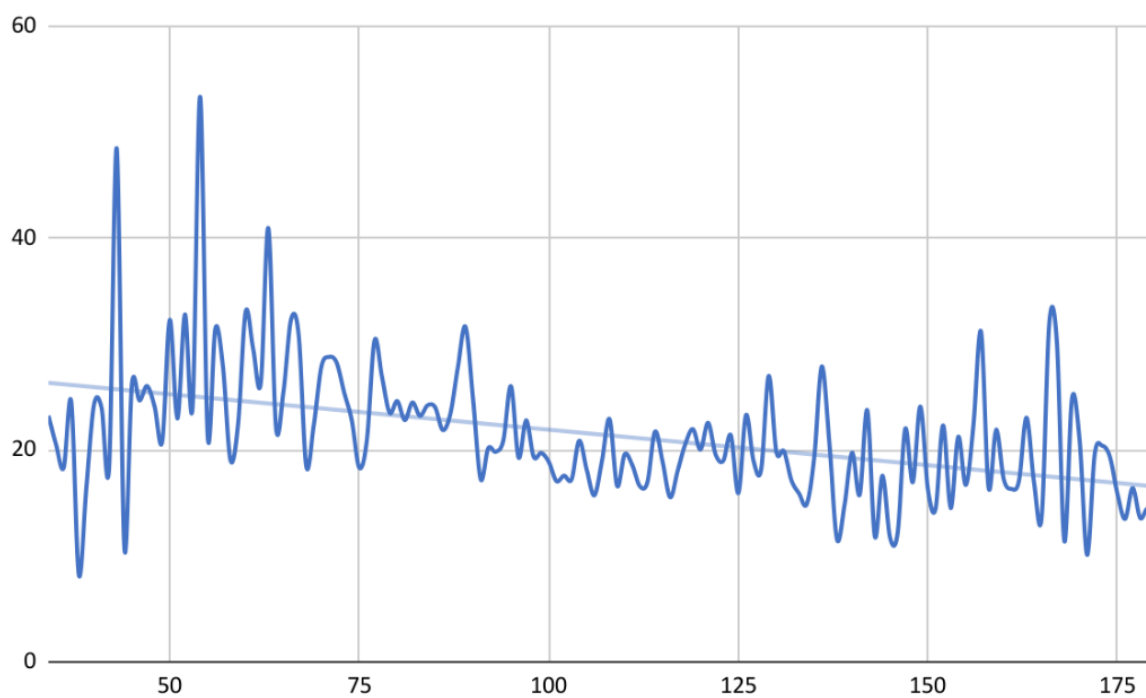


Figure 5

Distribution of Average Citations by Title Length for Articles with Colons



Abstracts

Contrary to our assumption, reading ease had a negative effect on citations. This was surprising because we assumed that if an abstract was more readable and less esoteric that people would be more likely to cite it. The opposite result, however, suggests that more technical abstracts yield greater citations. This might be the result of greater specificity provided in abstracts, or it might be due to certain topics or methodologies that rely upon long words with many syllables being cited more often, such as studies that rely upon advanced statistical procedures like “hierarchical linear modeling.” It could also mean that articles are often cited based on the content of their abstracts and that leaner abstracts do not provide other authors with enough information to warrant a citation. We do not take this result to mean that authors should attempt to make their abstracts intentionally difficult to decipher, but it does suggest that including technical language and detail in abstracts might be beneficial. Couple this with the positive effect that abstract length had on raw citations and the lack of effect that reading time had on citations, and the takeaway seems to be that more detail in abstracts is a good thing.

Open Access

Contrary to previous studies seeking to understand open access effects on citation counts, we did not detect an open access bump. At least two possible explanations exist for this discrepancy: time and context. Regarding time, many studies exploring the open access topic have restricted their analyses to relatively short timeframes, suggesting that there may be an initial open-access bump to citations but that this advantage might fade over time. In addition, the context of most studies in this realm has focused on the natural sciences, and it may be that education or the social sciences more broadly exhibit different citation patterns than other fields.

Conclusion

Results from our analysis reveal that some education research article features have significant (though relatively small) effects on citation counts. Notably, articles are most likely to be cited if (a) their titles include a semi-colon-designated subtitle, (b) their titles are 7 to 12 words in length, (c) their abstracts are longer, and (d) their abstracts include technical language.





Royce Kimmons

Brigham Young University

Royce Kimmons is an Associate Professor of Instructional Psychology and Technology at Brigham Young University where he studies digital participation divides specifically in the realms of social media, open education, and classroom technology use. He is also the founder of EdTechBooks.org. More information about his work may be found at <http://roycekimmons.com>, and you may also dialogue with him on Twitter [@roycekimmons](https://twitter.com/roycekimmons).



Ross Larsen

Brigham Young University

Dr. Ross Larsen is an Associate Professor of Instructional Psychology & Technology at Brigham Young University.



This content is provided to you freely by EdTech Books.

Access it online or download it at https://edtechbooks.org/50_years/impact_metrics.