7-12-2021

The readability of information systems research over three decades: An analysis of the Senior Scholars' Basket of Journals from 1990 to 2020

Lorenz Graf-Vlachy
TU Dortmund University, mail@graf-vlachy.com

The readability of information systems research over three decades: An analysis of the Senior Scholars' Basket of Journals from 1990 to 2020

Completed Research Paper

Lorenz Graf-Vlachy
TU Dortmund University & ESCP Business School
Otto-Hahn-Str. 4, 44227 Dortmund, Germany
lorenz.graf-vlachy@tu-dortmund.de

Abstract

If information systems research is to be widely disseminated, then it is important that its scientific texts are easily accessible. This study analyzes one aspect of accessibility, namely readability. Specifically, this study analyzes 5,467 abstracts of the eight information systems journals included in the Senior Scholars' Basket of Journals between 1990 and 2020. It operationalizes readability using two common measures, the Flesch Reading Ease (FRE) and the New Dale-Chall formula (NDC). The study finds a concerning decrease in readability over time. Supplementary analyses show that an increasing number of co-authors partially explains this trend, but that the trend persists even after controlling for this variable. Implications for the field of information systems and future research opportunities are discussed.

Keywords: Readability, information systems research, Flesch Reading Ease, New Dale-Chall

Introduction

It is generally important that scientific documents like articles or their abstracts be comprehensible for other researchers (Loveland et al. 1973). In fact, it might even be desirable that such works of science be accessible to the informed layperson as well (Scharrer et al. 2013). Prior researchers have consequently invested time and effort into studying how readable scientific articles—and especially their abstracts—are. Scholars recently found disconcerting downward trends in readability in various disciplines in the natural, behavioral, and social sciences as specifically related to the life sciences (Plavén-Sigray et al. 2017). Others have followed up with field-specific studies in fields like neuroscience and found similar results (Yeung et al. 2018). Partially because of these findings, scholars have begun to study how researchers can write more readable abstracts (Freeling et al. 2019) and how to generally improve how they report their findings (Hanel and Mehler 2019).

The accessibility of research on information systems is particularly important against the backdrop of the ubiquitous processes of digital transformation (Morakanyane et al. 2017; Nadkarni and Prügl 2020; Vial 2019). This is the case for a variety of reasons. First, the phenomenon is evolving quickly, so it is important that researchers can efficiently stay abreast of the latest findings. Second, practitioners (in business, but also in other sectors such as education) need to be able to access and process scientific material to transfer novel ideas from research into practice. Third, regulators must be able and willing

to consume scientific material about digital transformation to adequately set policy based on sound theory and evidence. And finally, a variety of other members of society should be able to access information systems research to be able to make sense of ongoing societal and technological developments.

Consequently, and in the spirit of calls for replication research in information systems (Dennis and Valacich, 2015), it is important to understand if the negative trends regarding readability observed in other fields are also present in information systems research. However, it is currently unknown if such a trend also plagues our field. For one, prior research has not specifically focused on information systems research. For another, it has demonstrated substantial heterogeneity between different fields (Plavén-Sigray et al. 2017). Therefore, we cannot simply assume that findings from other fields will generalize to information systems research.

To find out whether a negative trend regarding readability exists in the information systems literature, I study 5,467 abstracts of eight particularly important information systems journals. I analyze their readability over time using two different established measures of readability and find that the trend in the information systems literature mirrors the trends in other fields in that readability decreases over time. An increasing number of co-authors per article partially but not fully explains the trend.

The remainder of the article is organized as follows. I first describe the method of analysis, then proceed to present the empirical results, and conclude with a discussion comprising implications for the field as well as an acknowledgement of this study's limitations and opportunities for future research.

Method

Sample and Pre-Processing

I considered eight journals that were selected by the Association for Information Systems' "College of Senior Scholars" into the so-called "Senior Scholars' Basket" (Association for Information Systems 2011). These journals are particularly reputable and are frequently used in tenure decisions in information systems departments of universities around the globe. The basket currently includes the *European Journal of Information Systems*, the *Information Systems Journal*, *Information Systems Research*, the *Journal of AIS*, the *Journal of Information Technology*, the *Journal of MIS*, the *Journal of Strategic Information Systems*, and the *MIS Quarterly*.

I collected information on all articles published in these journals from 1990 through the end of 2020 in January 2021. Specifically, I obtained author names, publication dates, and the full text of the abstract. My initial sample included 6,750 records. I removed all records that did not represent regular journal articles (e.g., editorial material, corrections, or retraction notices) or which did not have an abstract. This reduced my sample to 5,485.

To ensure clean data, I randomly selected 100 abstracts from my sample. I reviewed them to confirm there were no data quality problems or—if there were problems—to apply appropriate pre-processing to the entire sample. During this manual inspection it became apparent that some publishers included copyright statements that required removal. I implemented corresponding removal routines in an automated pre-processing script. In addition, I removed all periods from abbreviations (e.g., "U.S.") because they might interfere with the identification of sentence boundaries. Also, I dropped all enumerators like "(1), (2)." I further expanded common abbreviations such as "vs.". Because they are not an input for the readability measures used in this paper (which I will discuss below), I also replaced all numbers and percentages in the abstracts with blank spaces. In a last step, I removed all text in parentheses and square brackets because such formatting frequently indicated references to other articles rather than substantive content.

Because the used readability measures require the number of sentences as an input, I decided to remove all abstracts with less than three sentences to reduce outliers. The removal of 18 such cases reduced the final sample to 5,467. Note that all analyses reported below are robust to the inclusion of these articles.

Figure 1 provides an overview over the number of journals in the sample, as well as the number of articles published in each year across the sampling frame. It is apparent that both the number of journals as well as the annual volume of published information systems research in them increases over time.

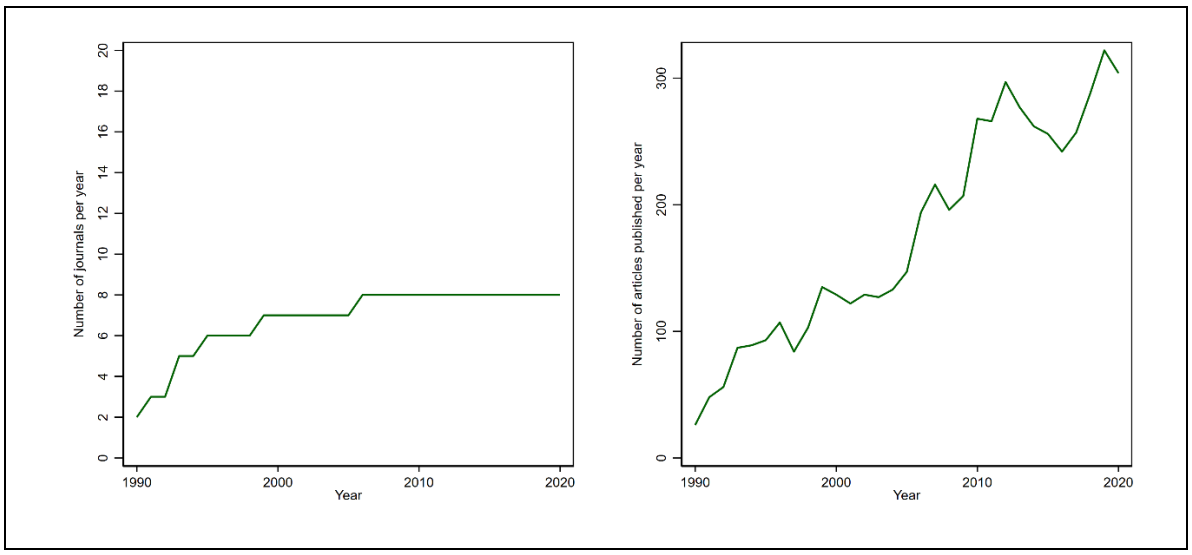

Figure 1. Number of Journals and Articles per Year over Time

Measurement

To measure readability, I employed two of the simplest and most frequently used measures (Plavén-Sigray et al. 2017). These are the Flesch Reading Ease (FRE) measure, which measures how easy it is to read a text, and the New Dale-Chall readability formula (NDC), which measures how hard it is to read a text.

I calculated FRE according to the following formula (Flesch 1948):

$$FRE = 206.835 - 1.015\left(\frac{words}{sentences}\right) - 84.6\left(\frac{syllables}{words}\right)$$

I calculated NDC according to the following formula (Chall and Dale 1995; Kincaid et al. 1975):

$$NDC = \begin{cases} 0.1579\left(\frac{difficult}{words} * 100\right) + 0.0496\left(\frac{words}{sentences}\right) + 3.6365 & \text{if } \left(\frac{difficult}{words}\right) > 5\% \\ 0.1579\left(\frac{difficult}{words} * 100\right) + 0.0496\left(\frac{words}{sentences}\right) & \text{if } \left(\frac{difficult}{words}\right) \leq 5\% \end{cases}$$

Because the two measures have slightly different constructions, words were counted in the following way. For FRE, contractions like "we've" or "isn't" were counted as two words each since they ultimately represent two separate words. A list of common words that was developed for use with the NDC measure, however, includes contractions like "here's" and thus implies that such cases should be counted as one word. Consequently, the word count logic differs mildly between the calculations. This is not, however, a substantial issue because contractions are extremely rare in the abstracts included in the sample. As a brief test, I calculated the Pearson correlation between the two word counts. It was almost total with $r = .99$ ($p < .001$).

I performed sentence counts after all pre-processing steps were completed. Periods, exclamation points, question marks, and semicolons were treated as sentence boundaries. For syllable counts, I relied on the Perl module *Lingua::EN::Syllable*.

I followed prior literature in the identification of difficult words for the NDC measure (Plavén-Sigray et al. 2017). Specifically, words were classified as difficult if they were not explicitly listed on the abovementioned NDC list of common words (which includes words like "quick," "touch," or "today").

Analytical approach

To test for a time trend in readability, I used ordinary least squares (OLS) regression models to analyze the effect of publication year on readability. Since articles are nested in journals, I included dummy variables for each journal in each estimation to account for journal fixed effects. This allows me to capture, for instance, journal-specific conventions, focuses, and styles that might influence the readability of their abstracts. To further account for the nested data structure, I clustered standard errors at the journal. I performed all analyses using the *areg* estimator in Stata 16.1.

In particular, I estimated models for the effect of publication year on readability as measured by FRE as well as for the effect on readability as measured by NDC. For each readability measure, I estimated null models without publication year as an independent variable and models with it. Furthermore, I estimated two models that additionally include the number of co-authors as a covariate.

Results

Table 1 represents the correlation matrix of all variables used in this study. The FRE and NDC scores are highly significantly correlated at $r = -.66$ ($p < .001$). This is in line with correlations between these measures observed in prior research (Plavén-Sigray et al. 2017) and suggests that both measures actually capture a common construct.

Table 1. Correlation Matrix (*p*-Values in Parentheses)

	FRE	NDC	Number of authors	Publication year
FRE	1.000			
NDC	-0.662 (0.000)	1.000		
Number of authors	-0.068 (0.000)	0.148 (0.000)	1.000	
Publication year	-0.107 (0.000)	0.208 (0.000)	0.2963 (0.000)	1.000

Figure 2 depicts the mean FRE and NDC scores per year across all articles. As is readily apparent, there is a clear downward in readability, as indicated by decreasing FRE and increasing NDC scores. This already suggests a negative trend over time, as do the negative correlations between publication year and both readability measures in Table 1.

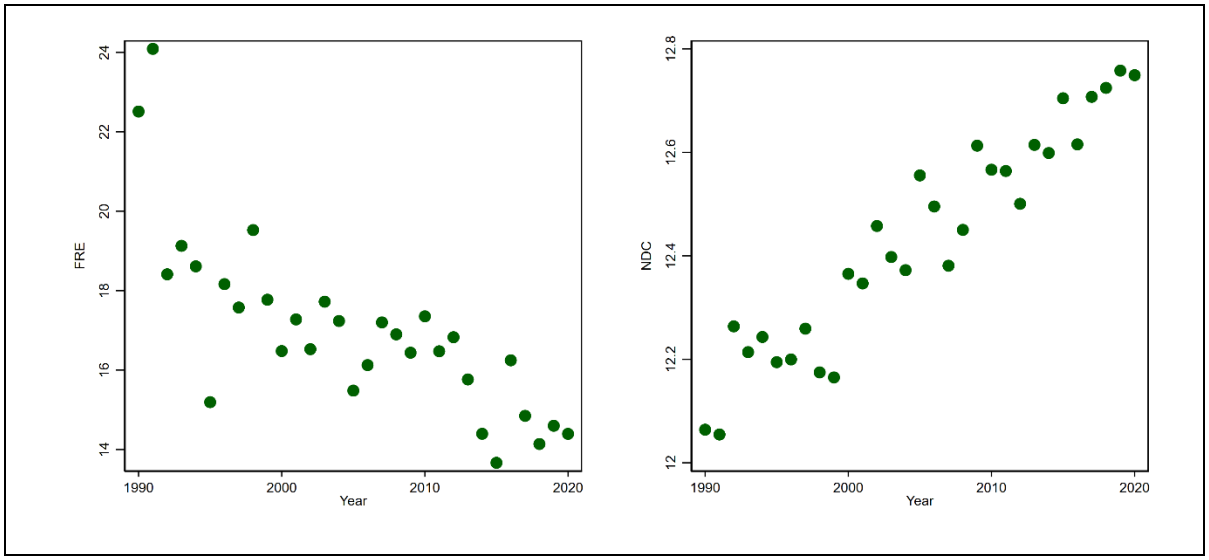

Figure 2. Mean Flesch Reading Ease (FRE) and New Dale-Chall (NDC) Readability over Time

FRE and NDC rely on three characteristics of the analyzed text: The number of syllables per word (only FRE), the number of words per sentence (both FRE and NDC), and the share of difficult words (only NDC). One might therefore wonder if the apparent decrease in readability might be due to one of these characteristics in particular. The three panels in Figure 3 show that this appears to not be the case. All three characteristics seem to change in parallel, with all three increasing over time. Thus, all three characteristics appear to contribute to the negative trend in readability.

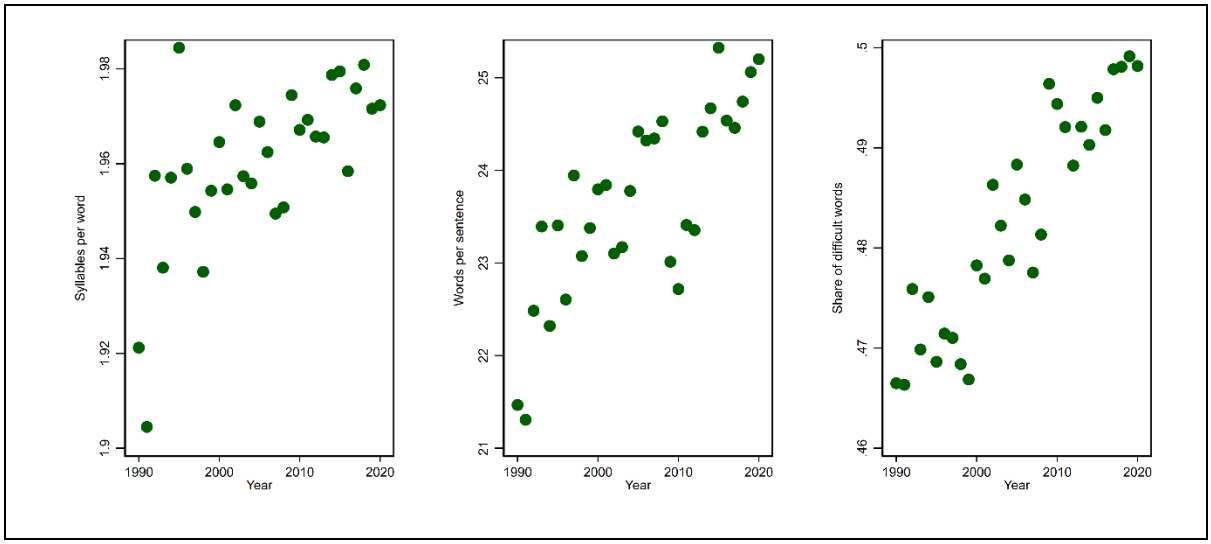

Figure 3. Number of Syllables per Word, Number of Words per Sentence, and Share of Difficult Words over Time

The results of my regression analyses can be seen in Table 2. Models 1 and 4 are the null models for FRE and NDC, respectively. They only include journal dummies as predictors. Models 2 and 5 add publication year as an additional predictor to analyze our main relationship of interest. In both models, readability is significantly ($p < 0.01$) related to publication year, indicating that there is indeed a trend

of readability over time. The negative sign of the coefficient for FRE and the positive sign of the coefficient for NDC both confirm that this trend is towards decreased readability.

Table 2. Regression Results
(Standard Errors in Parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

	Model 1 (FRE)	Model 2 (FRE)	Model 3 (FRE)	Model 4 (NDC)	Model 5 (NDC)	Model 6 (NDC)
Publication year		-0.161**	-0.145*		0.021***	0.018***
		(0.05)	(0.05)		(0.00)	(0.00)
Number of authors			-0.417			0.061**
			(0.29)			(0.01)
Journal dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	16.215	340.593**	309.556*	12.522***	-28.788**	-24.245**
	(.)	(92.54)	(94.06)	(0.00)	(5.96)	(6.05)
N	5,467	5,467	5,467	5,467	5,467	5,467
R^2	0.007	0.017	0.018	0.018	0.055	0.061
Resid. deg. of freedom	7	7	7	7	7	7
BIC	42906.0	42858.0	42859.8	13366.3	13168.3	13143.3

Prior literature mentions an increasing number of co-authors as a possible explanation for decreasing readability, suggesting that too many authors might have a detrimental effect on writing quality (Drenth 1998; Epstein 1993; Plavén-Sigray et al. 2017). Graf-Vlachy (2001), for example, suggests that larger author teams might have greater difficulties in accommodating competing editing suggestions while maintaining readable text or that a diffusion of responsibility for readability reduces it. Indeed, other researchers' empirical results support the notion that a greater number of authors is negatively related to readability, but that the decrease in readability remained even after accounting for the number of authors (Plavén-Sigray et al. 2017). As can be seen in Figure 4, the mean number of co-authors per article also increases substantially over time in my sample.

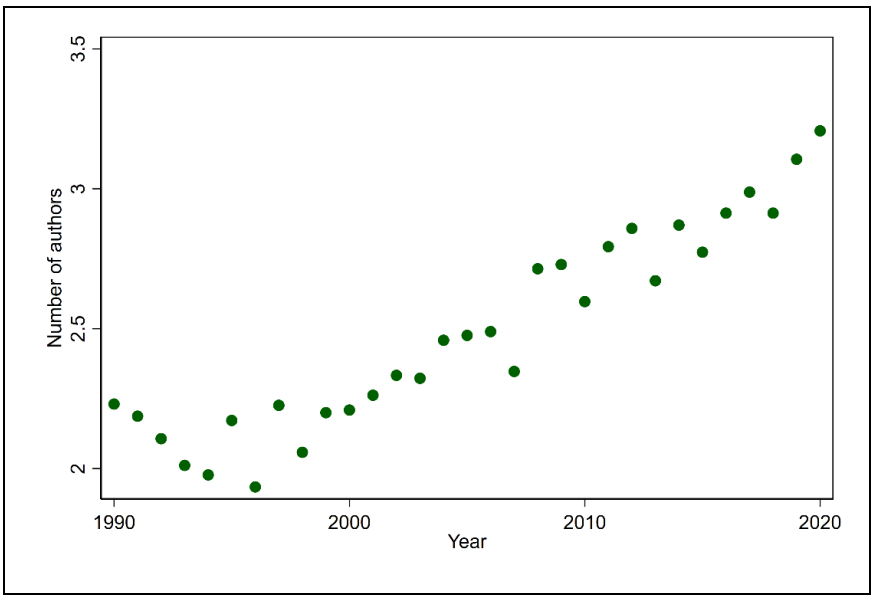

Figure 4. Mean Number of Authors per Article over Time

Models 3 and 6 in Table 2 therefore include the number of co-authors of each paper as an additional variable. For Model 3, which uses FRE as the readability measure, the significance of publication year is reduced, but the coefficient remains significant ($p < 0.05$). The number of co-authors does not have a significant influence on readability although the effect is in the expected direction. In Model 6, which uses NDC as the readability measure, publication year retains its high significance ($p < 0.001$), and the number of authors also exhibits a significant influence on readability ($p < 0.01$). In sum, my results regarding a time trend remain essentially unaffected, and the number of authors had a significant effect on readability only when it was measured using NDC. The latter is interesting, suggesting that the number of authors might particularly influence the number of difficult words, the only aspect of readability that is present in NDC but not in FRE.

Since the left panel of Figure 2 might hint at the years 1990 and 1991 being outliers with regard to FRE, I re-ran Models 1 through 3 as a robustness check without these two years. The obtained results were fully consistent with the main analysis.

Discussion

Already fifty years ago, researchers in organization and management studies have noted a decrease in the readability of academic journals (Loveland et al. 1973). Others more recently showed that this trend is present in other disciplines as well (Plavén-Sigray et al. 2017). My analyses demonstrate that the trend is also prevalent in information systems research.

Of course, such a trend would not be particularly concerning if the overall readability of abstracts was still high. Unfortunately, this is not the case. The FRE readability measure is calibrated such that scores below zero are assigned to texts that college graduates might not understand anymore (Kincaid et al. 1975; Flesch 1948). The share of abstracts in information systems research that one might consider incomprehensible for such educated laypeople increased steadily. In the first decade of the sample, 7.1% of abstracts received FRE scores below zero. In the second decade, this share rose mildly to 8.1%. In the last decade in the sampling frame, 11.1% of abstracts received such a score.

Implications for Information Systems Researchers

The results of my analyses are concerning because they suggest that the field's research findings might have become less accessible to potential readers over time. First, other academics might understand abstracts less when they are hard to read and they might even cite them less (Hartley 1994, Freeling et al. 2019). Second, practitioners might be demotivated to read scientific articles or even find themselves unable to comprehend them at all. Third, journalists might have a similar reaction. This would be unfortunate because they can be a powerful conduit to publicize research output to a wider audience (Bubela et al. 2009).

I thus propose that there are implications for researchers in the field in three roles. First, I urge researchers who write scientific texts to be mindful of the readability of their works. Naturally, scientific research entails some degree of complexity, and this will always at least partially translate into texts that are not necessarily simple (Knight 2003). However, this does not mean that there is nothing authors can do. When writing abstracts, they might at least make a deliberate attempt to write in a particularly simple and accessible way (Hartley 1994; Freeling et al. 2019). There are many free online tools to assess readability as measured by FRE and NDC which can help authors check the readability of their writing.

Second, in their role as reviewers, researchers might push for ever greater simplicity in writing. Reviewers have a substantial impact on what ultimately gets published—and in what form. I thus encourage reviewers to ask manuscript authors to simplify their language where feasible and appropriate. I particularly call on reviewers to give specific feedback, for instance on excessively long sentences or on unnecessary use of scientific jargon.

In the end, of course, journal editors and conference (track) chairs are the ultimate arbiters. It is thus also incumbent on them to encourage readable writing. They can do so in various ways. For example, they could explicitly make readability a criterion for decisions, and encourage reviewers to attend to it. They could also help authors produce abstracts that are easier to read. For instance, they could allow slightly longer abstracts, which might reduce the need of authors to cram as much information as possible into a very limited amount of text. They could also institute structured abstracts. Such abstracts provide a "template" to fill in and have been shown to aid readability (Hartley 2003; Hartley and Benjamin 1998). Alternatively, they might offer authors the opportunity to provide so-called "lay summaries" in addition to conventional scientific abstracts (Kuehne and Olden 2015).

Limitations and Future Research

This article has several limitations, which in turn open avenues for future research. First, the IS Senior Scholar's basket includes only journals and is "focused on behavioral, business-oriented IS research, which might reflect a majority, but is not a universal model" (Association for Information Systems 2011). Future researchers might thus consider replicating my analysis using other IS publication outlets, including conference proceedings, journals focused more on a design-science approach to information systems, as well as practitioner-oriented journals.

Second, I only analyze the abstracts of articles and not their full texts. While prior research has found substantial correlations between the readability of abstracts and that of the entire articles (Plavén-Sigray et al. 2017), future researchers might wish to replicate my analyses using the full texts.

Third, there might be limitations pertaining to the used readability measures. For one, the measures and subjective assessments of readability may not always match (Griesinger and Klene 1984). It is, for example, possible that as a scientific field develops over time, the capability of actors in the field might increase. It is thus conceivable that scientific audiences might not actually perceive readability to have decreased over time. Future research with subjective readability assessments might address these issues. For another, the two measures of readability might not capture every aspect of texts that is relevant to their readability (Benjamin 2012). For example, they do not measure the length of causal chains which has also been demonstrated to affect readability (Otero et al. 2004). Similarly, one can imagine that the clarity of transitions between sentences or the choice of adequate vocabulary influences readability as well. Other work has also demonstrated that more complex automated measures of readability may hold potential to make even more precise assessments of readability (Crossley et al., 2017; Si and Callan 2001). Finally, other contextual factors might further influence the readability of abstracts.

Fourth, a potential alternative explanation for the observed time trend in readability is that journals might have changed in their orientation over time. It is, for instance, conceivable that some journals started out with a stronger practitioner focus and gradually transformed into journals catering almost exclusively to an academic audience. Paired with a presumed tendency of academic authors to use harder-to-read language than practitioners might use (Knight 2003), this might be another cause for the decline in readability. Devising a measure for the scholarliness of academic journals—without recourse to the readability of the texts therein, of course—would allow testing this notion.

In conclusion, my research indicates that there might be improvement potential regarding the readability of abstracts (and potentially entire articles) in information systems research. We—as a discipline and as individual researchers—should have a vested interest in making our research as easy to understand as possible (without compromising accuracy and detail, of course). In fact, there is evidence that better-written scientific works are cited more frequently (Freeling et al. 2019). Editors and gatekeepers in the community should also not be afraid to enforce—or at least nudge authors towards—more readable writing, as journal prestige is not tainted by greater readability (Hartley and Benjamin 1998). I therefore call on all information systems scholars to keep the reader in mind when writing and reviewing. Making scientific work easier to read likely benefits us all.

References

Association for Information Systems 2011. "Senior Scholars' Basket of Journals," (https://aisnet.org/page/SeniorScholarBasket, accessed January 28, 2021).

Benjamin, R. G. 2012. Reconstructing Readability: "Recent Developments and Recommendations in the Analysis of Text Difficulty," *Educational Psychology Review* (24:1), pp. 63–88.

Bubela, T., Nisbet, M. C., Borchelt, R., Brunger, F., Critchley, C., Einsiedel, E., Geller, G., Gupta, A., Hampel, J., Hyde-Lay, R., Jandciu, E. W., Jones, S. A., Kolopack, P., Lane, S., Lougheed, T., Nerlich, B., Ogbogu, U., O'Riordan, K., Oullette, C., Spear, M., Strauss, S., Thavaratnam, T., Willemse, L., and Caulfield, T. 2009. "Science communication reconsidered," *Nature Biotechnology* (27:6), pp. 514–518.

Chall, J. S., and Dale, E. 1995. *Readability revisited: The New Dale-Chall readability formula*. Brookline Books, Massachusetts: Cambridge.

Crossley, S. A., Skalicky, S., Dascalu, M., McNamara, D. S., and Kyle, K. 2017. "Predicting Text Comprehension, Processing, and Familiarity in Adult Readers: New Approaches to Readability Formulas," *Discourse Processes* (54:5-6), pp. 340–359.

Dennis, A. R., and Valacich, J. S. 2015 "A Replication Manifesto," AIS Transactions on Replication Research, (1), 1.

Drenth, J. P. 1998. "Multiple authorship: the contribution of senior authors," *Journal of the American Medical Association* (280:3), pp. 219–221.

Epstein, R. J. 1993. "Six authors in search of a citation: villains or victims of the Vancouver convention?," *British Medical Journal* (306:6880), pp. 765–767.

Freeling, B., Doubleday, Z. A., and Connell, S. D. 2019. "Opinion: How can we boost the impact of publications? Try better writing," in *Proceedings of the National Academy of Sciences of the United States of America*, (116:2), pp. 341–343.

Flesch, R. 1948. "A new readability yardstick," *Journal of Applied Psychology* (32:3), pp. 221–233.

Graf-Vlachy, L. 2021. "Is the readability of abstracts decreasing in management research?," *Review of Managerial Science*, forthcoming.

Griesinger, W. S., and Klene, R. R. 1984. "Readability of Introductory Psychology Textbooks: Flesch versus Student Ratings," *Teaching of Psychology* (11:2), pp. 90–91.

Hartley, J. 2003. "Improving the Clarity of Journal Abstracts in Psychology," *Science Communication*, (24:3) pp. 366–379.

Hartley, J. 1994. "Three ways to improve the clarity of journal abstracts," *British Journal of Educational Psychology* (64:2), pp. 331–343.

Hartley, J., and Benjamin, M. 1998. "An evaluation of structured abstracts in journals published by the British Psychological Society," *British Journal of Educational Psychology* (68:3), pp. 443–456.

Hanel, P. H., and Mehler, D. M. 2019. "Beyond reporting statistical significance: Identifying informative effect sizes to improve scientific communication," *Public Understanding of Science* (28:4), pp. 468–485.

Kincaid, J. P., Fishburne Jr., R. P., Rogers, R. L., and Chissom, B. S. 1975. "Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel," *Naval Technical Training Command Millington TN Research Branch*, RBR-8-75.

Knight, J. 2003. "Scientific literacy: Clear as mud," *Nature* (423:6938), pp. 376–378.

Kuehne, L. M., and Olden, J. D. 2015. "Opinion: Lay summaries needed to enhance science communication," in *Proceedings of the National Academy of Sciences of the United States of America* (112:12), pp. 3585–3586.

Loveland, J., Whatley, A., Ray, B., and Reidy, R. 1973. "An Analysis of the Readability of Selected Management Journals," *Academy of Management Journal* (16:3), pp. 522–524.

Morakanyane, R., Grace, A., and O'Reilly, P. 2017. "Conceptualizing Digital Transformation in Business Organizations: A Systematic Review of Literature," in *Proceedings of the 30th Bled eConference*, pp. 427–444.

Nadkarni, S., and Prügl, R. 2020. "Digital transformation: a review, synthesis and opportunities for future research," *Management Review Quarterly*, in press.

Otero, J., Caldeira, H., and Gomes, C. J. 2004. "The influence of the length of causal chains on question asking and on the comprehensibility of scientific texts," *Contemporary Educational Psychology* (29:1), pp. 50–62.

Plavén-Sigray, P., Matheson, G. J., Schiffler, B. C., Thompson, W. H., and King, S. 2017. "The readability of scientific texts is decreasing over time," *eLife*, (6), e27725.

Si, L., and Callan, J. 2001. "A statistical model for scientific readability," in *Proceedings of the Tenth International Conference on Information and Knowledge Management,* pp. 574–576.

Scharrer, L., Britt, M. A., Stadtler, M., and Bromme, R. 2013. "Easy to Understand but Difficult to Decide: Information Comprehensibility and Controversiality Affect Laypeople's Science-Based Decisions," *Discourse Processes* (50:6), pp. 361–387.

Vial, G. 2019. "Understanding digital transformation: A review and research agenda," *Journal of Strategic Information Systems* (28:2), pp. 118–144.

Yeung, A. W. K., Goto, T. K., and Leung, W. K. 2018. "Readability of the 100 Most-Cited Neuroimaging Papers Assessed by Common Readability Formulae," *Frontiers in Human Neuroscience* (12), pp. 308.