Journal of
Professional Communication jpc

Prose and cons of scholarly articles: How readability tests expose poor knowledge mobilization in academic publications

Peter J. McDonald★, Philip Savage

McMaster University (Canada)

Article Info

Article Type:
Research Article

Article History:
Received: *2020-01-14*
Revised: *2020-06-25*
Accepted: *2020-07-01*

Keywords:
Readability
Academic prose
Knowledge mobilization
Flesch Reading Ease
Gunning FOG Index
Humanities
Social science
Professional communication

Abstract

Current literature shows that poor and unclear writing is a significant barrier for non-academic audiences. Readability research is a growing interest among STEM and health science fields; however, the humanities and social science disciplines are neglected. To address this gap, articles from the humanities and social science disciplines were analyzed using the Flesch Reading Ease (FRE) and the Gunning FOG Index (GFI) readability tests. Results show that the FRE mean score for all analyzed articles is 29.04, and the total GFI mean score is 18.02, meaning they are extremely difficult to read and often require a post-secondary education for adequate comprehension. Empirically driven, quantitative articles had no significant difference in readability than sense-making, qualitative articles. Results also show that the humanities and social sciences have readability similar or equivalent to STEM and health science fields.

While there is an abundance of research being produced, there is, however, little attention given to how effective that research is disseminated among non-academic audiences (Cooper, Rodway, & Read, 2016; Shields & Evans, 2012). How research is communicated becomes as important as the research itself because "when a writer fails to communicate, the work is pointless, for scholarship is necessarily a social enterprise" (Selvin & Wilson, 1984, p. 206). Recently, there has been pressure among scholars to make their research more accessible, especially when that research is publicly funded (Cooper et al., 2018; Cain, Shore, Weston, & Sanders, 2018). This complements a shift in the role of research itself as more universities in Canada "pledge to bettering the lives of not only students but also of the greater public(s) and the communities they operate within" (Cain et al., 2018, p. 41). This shift is pressured at the behest of granting agencies, government, and

★*Corresponding author (Peter J. McDonald)*
Email: mcdonp1@mcmaster.ca

"presumably the general public" (Matheson & Edwards, 2016, p. 5). In particular, the Social Science and Humanities Research Council (SSHRC) specifically cites knowledge mobilization (KMb) as a policy which all research must abide to receive funding (Cain et al., 2018; Cooper et al., 2018). KMb is a transactional process between knowledge producers and knowledge users that connects information to communities through various strategies, specifically to promote research in an accessible format (Sá, Li, & Faubert, 2011; Cain et al., 2018; Shields & Evans, 2012). More simply, KMb "helps make research useful to society" (Matheson & Edwards, 2016, p. 4). Following KMb strategies increases the impact of research and encourages a collaborative approach to sense-making by involving various stakeholders—all in pursuit of the social enterprise (Nichols, Phipps, Provençal, & Hewitt, 2013).

Academic prose is notoriously uncharitable to readers, especially those lacking in formal education (Cheek & Rosenhaupt, 1968). Text readability refers to the ease of reading resulting from a low text difficulty (Benjamin, 2012). Although there is a healthy amount of readability research in STEM and health science fields, the humanities and social sciences have generally been ignored. The present endeavour attempts to remedy that and add to the growing literature on readability in the humanities and social sciences. To analyze readability, this study investigates scholarly, peer-reviewed publications in academic journals using the Flesch Reading Ease (FRE) and the Gunning FOG Index (GFI) readability tests.

Literature review

Research products are highly relevant to audiences outside the traditional academic spheres. For example, policymakers regularly consult available research to make evidence-based decisions and recommendations; practitioners, such as teachers, can use new research to guide pedagogical practices in the classroom; and advocates pursuing social change stimulated by innovative research can orient their policies and missions accordingly (Nichols et al., 2013; Cooper et al., 2016; Cain et al., 2018). These stakeholders, however, do not comprise the entire body of interested parties. Laypeople are increasingly encouraged to participate in research, be informed with contemporary information, and make decisions that are consistent with facts (Kirkpatrick et al., 2017). To maximize the impact research has on non-academic audiences, many policies, mandates, and initiatives have been conceived and implemented with varying levels of success (Kirkpatrick et al., 2017; Cain et al., 2018; Cooper et al., 2018). However, there is a discrepancy between the institutional

goals of disseminating research and the actual consumption of that research by non-academic audiences (Cain et al., 2018; Cooper et al., 2018). To address this concern, KMb has been increasingly discussed among research institutions, universities, and governmental organizations, such as funding agencies (Matheson & Edwards, 2016; Shields & Evans, 2012).

Barriers to knowledge mobilization

Since its inception in the 1990s, enacting KMb has been fraught with difficulties (Shields & Evans, 2012; Sá et al., 2011). These barriers exist at the institutional, professional, and conceptual level. At the institutional level, universities still favour the longstanding, traditional models of publishing, such as through peer-reviewed journals and acquiring grants outside KMb-conscious institutions (Cain et al., 2018). Additionally, KMb and academia are not strictly compatible because academic communities focus on generating research and not necessarily making that research accessible to non-academic audiences (Cooper et al., 2018). For many governmental grant agencies that explicitly require KMb considerations for funding, accountability to KMb practices is often overlooked or insufficient for research after it is completed (Cain et al., 2018). While many institutions value KMb activities, many still do not have the resources to support researchers (Cooper et al., 2018; Cain et al., 2018). Even when institutions have KMb support, researchers rarely use those resources when producing or disseminating information (Cooper et al., 2016). Additionally, the goals themselves have been the point of concern as many institutions have poorly defined objectives, vague strategies, or simply lack appropriate resources to address them (Matheson & Edwards, 2016; Cooper et al., 2018; Sá et al., 2011).

At the professional level, many institutions do not offer appropriate incentives to carry out the arduous activities needed for effective KMb (Sá et al., 2011). For example, many institutional structures still favour publications in "reputable outlets aimed at disciplinary audiences" (Sá et al., 2011, p. 504). This is because KMb-related activities are not always well-regarded among professional researchers (Cain et al., 2018), and pursuing KMb initiatives can be seen to distract a researcher seeking tenure or promotions in favour of professionally recognized activities (Sá et al., 2011). Because KMb activities are labour- and time-intensive, "expecting researchers to divert efforts from research production to engaging with non-academic target audiences is potentially an unrealistic goal" (Cooper et al., 2018, p. 10).

Challenges at the conceptual level tend to discuss how research ought to operate. For example, Shields and Evans (2012) suggest that there is a delicate balance between "curiosity-driven research and excellence and relevance" (p. 255). This, however, may not fully represent the conceptual challenges of KMb initiatives. As Sá et al. (2011) write, aggressive KMb activities privilege a utilitarian and instrumentalist paradigm rather than free, academic inquiry. Chubb and Watermeyer (2016) argue that new criteria are now "modifying the ways in which academics approach research and behave as researchers" by requiring "credible statements of how they will ensure economic and/or societal returns from their research" (p. 2362). Consequently, researchers have expressed concern that these mandates are interfering with academic freedom and forcing research to produce outcomes that follow prescribed results (Nichols et al., 2013).

Strategies for knowledge mobilization

Through several key steps, organizations can better ensure that the research being disseminated can maximize its impact on society (Cain et al., 2018). Clear and feasible goals are needed at the institutional level to orient policies that ensure KMb is carried out (Sá et al., 2011; Moore et al., 2016). Follow-ups are essential at all levels and among peers as successful KMb requires monitoring to ensure research products match outlined goals (Cooper et al, 2016; Moore et al., 2016; Cain et al., 2018). KMb is a contextual process and requires support structures that can accommodate cross-discipline research products (Cooper et al., 2016; Sá et al., 2011). Partnerships have also been proven to be effective. Outreach initiatives aiming to build meaningful networks among academics, non-academics, and communities have had increasing success (Nichols et al., 2013). Universities especially are effective sites for community-based research with students because KMb is an interactive process enhanced when end-users facilitate knowledge exchange (Nichols et al., 2013). Nurturing strong partnerships between researchers and policymakers also has benefits as practitioners typically have strong connections to communities already (Cooper et al., 2016). Self-awareness of the average academic publication can also be useful because if impactful research is overlooked, it can have little to no effect on society. Therefore, Shields and Evans (2012) recommend that articles get a "social life" to introduce or extend their relevance in popular discourse (p. 255).

Strategies to overcome the culture of academia also require attention. For example, academic values of tenure, promotions, and esteem can be re-evaluated by integrating KMb as a respected and legitimate research process (Cooper et al., 2016). Appropriate funding streams to finance initiatives, such as partnerships, events, training, outreach, and social media similarly require attention (Cooper et al., 2016; Sá et al., 2011). These support structures can also incorporate training for researchers on best practices for KMb and how to network with non-academics (Sá et al., 2011). Importantly, training and support to make publications more readable have been a commonly cited strategy for KMb—a strategy that many publications do not employ (Kirkpatrick et al., 2017; Cooper et al., 2016; Cain et al., 2018; Shields & Evans, 2012; Sá et al., 2011). The specialized terminology and jargon make traditional publications inaccessible (Matheson & Edwards, 2016), and poor articulation can obscure the implications of research (Cooper et al., 2016; Cain et al., 2018). While many KMb initiatives can be in place, end-users must have the ability to access and understand that knowledge. Enhancing readability can clearly communicate research to non-academics and improve KMb (Kirkpatrick et al., 2017; Cooper et al., 2016; Cain et al., 2018). In keeping with the spirit of KMb, the readability of research products becomes increasingly important.

Issues of access versus knowledge mobilization

The internet has turned issues of access to information into inundation of information in mere decades. Because of the internet's affordances, non-academics increasingly express their desire to obtain information (Zethsen, 2018). The digital age has made people skeptical, and acquiring information, particularly research, is now a social need (Zethsen, 2018). However, access to information "is not the same as understanding" information (Zethsen, 2018, p. 86). While many research products may be accessible, poor writing can prevent many people from understanding the research. Issues of accessibility can also be extended to providing research in other languages, dialects, and registers (Zethsen, 2018). Timeliness can also be included in issues of accessibility as timely information can swiftly address issues a society is facing, such as misinformation. Access to information is not the same as providing knowledge, "and this distinction is crucial if the goal is an empowered society" (Zethsen, 2018, p. 92).

Readability

Researchers have noted that comprehension and readability share a close relationship (Sawyer, Laran, & Xu, 2008). Although comprehension is notably difficult to measure, "researchers often use readability as a proxy" (Lee & French, 2011, p. 694). Readability tests have a long and accepted tradition in academic research (Sawyer et al., 2008; Lee & French, 2011). For assessing text difficulty, readability tests typically rely on two fundamental assumptions: smaller words and shorter sentences allow for easier comprehension. Generally, readability tests target these variables to determine a score. Thus, complex sentences have larger words and longer sentences, where more syllables can indicate longer and more complex words (Choudhry et al., 2016; Modiri et al., 2018). Complex sentences "require the reader to maintain more concentration to understand the meaning of a sentence" (Choudhry et al., 2016, p. 632). Improving readability is correlated to faster reading speeds, increased knowledge retention, and increased reader comprehension (Lee & French, 2011; Sawyer et al., 2008).

Out of academia, writers and editors in various industries regularly use readability tests to ensure writing is appropriate for target audiences (Kirkpatrick et al., 2016), such as insurance and financial services (Van Boom, Desmet, & Van Dam, 2016), health (Walfish & Watkins, 2005), education (Crossley, Greenfield, & McNamara, 2008), and government administrations (Ficzere, Hagan, Ness, Greene, & Hill, 2016). Readability is not related to issues of legibility or typography. Additionally, it does not analyze the content of the writing, such as quality, accuracy, level of interest, or writing style (Sawyer et al., 2008).

Methods

Using the Flesch Reading Ease (FRE) and Gunning FOG Index (GFI) readability tests, 100 articles from humanities and social sciences disciplines in English-language journals were analyzed (see Appendix A for detailed list). Both readability tests assessed the same 100 articles. These readability tests employ quantitative measures to assess text complexity and thus a reader's ability to comprehend the material (Choudhry et al., 2016). The FRE has been cited as the most reliable and most frequently used readability test (Hartley, 2016). To determine a text's readability, divide words by number of sentences

and syllables by number of words and apply the following formula:

$$206.835 - 1.015 \left(\frac{Words}{Sentences} \right) - 84.6 \left(\frac{Syllables}{Words} \right)$$

Higher scores indicate a text's ease of reading and lower scores indicate text difficulty. However, as readability is highly contextual, there is potentially no lower limit a reading score can have. For example, it is possible for a text to receive a negative score indicating its extreme reading difficulty.

Table 1[1]

FRE scores and their associated interpretations

FRE Score	*Approx. reading age*	*Difficulty*	*Example*
90-100	10-11	Very easy	Comics
80-90	11-12	Easy	Pulp fiction
70-80	12-13	Fairly easy	Popular novels
60-70	14-15	Average	Tabloid newspapers
50-60	16-17	Fairly difficult	Introductory textbook
30-50	18-20	Difficult	Undergraduates' essay
0-30	Graduate	Very difficult	Academic prose

The GFI is another standard assessment tool used by writers for "analyzing textbooks and technical writing" (Whitt & Creech, 1983, p. 42). This readability test derives its readability score by calculating total number of words, sentences, and complex words, where complex words are defined as containing three or more syllables (Modiri et al., 2018). The formula:

$$4\left[\left(\frac{Words}{Sentences} \right) + 100 \left(\frac{Complex\ Words}{Words} \right)\right]$$

The reading score indicates the number of years of formal education required to comprehend 75% of the text in a single reading (Whitt & Creech, 1983). Grades 13 through 15 correspond to years of post-secondary education. Like with the FRE, the GFI has no limit on a text's readability score and can endlessly rise higher according to the text's complexity.

1 Hartley, 2016, p. 1524

Table 2[2]

GFI scores and their associated interpretations

Score	*Grade*
4.9 or lower	Easily understood by an average 4th grade student or lower
5.0-5.9	Easily understood by an average 5th or 6th grade student
6.0-6.9	Easily understood by an average 7th or 8th grade student
7.0-7.9	Easily understood by an average 9th or 10th grade student
8.0-8.9	Easily understood by an average 11th or 12th grade student
9.0-9.9	Easily understood by an average 13th – 15th grade student
10.0 or higher	Easily understood by an average university graduate

Academic fields

A list of humanities and social sciences disciplines was alphabetized, then a random number generator selected 12 as the starting point and determined an interval of 6 (Appendix B). This process selected Political Science, Communication Studies, Literature, Religious Studies, and Education. The 100 articles were taken from these academic fields.

Sampling

Articles were taken from peer-reviewed, English-language academic journals. Units of analysis were the articles but excluded abstracts and bibliographies. Samples were found using Google Scholar Metrics[3] and choosing the top 20 most cited articles by academic field. This method produced articles from 2013–2017, though this is only incidental. Instead, an article's citation count was used because it indicates impact on the respective field and suggests a higher readership and dissemination into the broader public. Ulrichsweb[4] was used to verify that journals from which articles were taken were peer-reviewed.

Articles were converted to a plain text format, then copied and pasted

2 Lynch et al., 2017, p. 129
3 https://scholar.google.com/citations?view_op=top_venues&hl=en
4 http://ulrichsweb.serialssolutions.com/

into an online readability test for analysis[5]. Slight modifications of the texts were necessary to ensure accurate analysis. Two such modifications occurred: all forms of in-text citations were removed and hyphenated words due to line breaks were reconstituted. All modifications facilitate an accurate readability analysis without compromising the text. Sample sizes were approximately 3,000 words for each article. The Pearson coefficient was used to determine the relationship between normalized citation counts (NCC; total number of citations divided by the number of years since publication), academic field, and readability.

Results

Table 3
Mean readability scores and standard deviation by field and formula

Academic Field	*Flesch Reading Ease Score*	*Gunning FOG Index Score*
Political Science	28.0 (SD 5.4)	18.4 (1.5)
Communication Studies	28.8 (8.5)	17.8 (1.6)
Literature	32.3 (9.6)	18.4 (2.2)
Religious Studies	25.9 (9.3)	18.2 (2.4)
Education	30.2 (8)	17.6 (1.8)

Of the five academic fields tested, the total mean for the FRE and GFI is 29.04 and 18.02, respectively. These scores fall within the very difficult to read category and require a post-secondary education for adequate comprehension. The academic field with the poorest FRE mean score is Religious Studies (25.9), while Political Science and Literature receive the poorest mean GFI score (18.4). Despite this, Literature still receives the best mean readability score on the FRE (Appendix C has detailed listing).

For the total mean, the range between academic fields is minimal with a 6.36-point spread in FRE scores and a .82 spread in GFI scores. This indicates that each academic field does not deviate far from the total mean.

Within individual fields, however, wider ranges were noted. Religious Studies had the widest range of FRE scores (1.9 to 44.2), while Political Science had the widest range for GFI scores (15.6 to 26). These outliers can explain why they received the poorest scores in their respective analyses—their high standard deviation. For the narrowest range, Education was the most

5 http://www.readabilityformulas.com/free-readability-formula-tests.php

consistent in FRE scores (13.7 to 43.2) and GFI scores (14.9 to 20.9). This narrow range, however, did not make it the most readable academic field.

Individual articles had a wide range of readability scores. The worst scoring article on the FRE received a 1.4 (Political Science) while the best was 51 (Communication Studies). The worst GFI scoring article received a 26 (Political Science) and the best scoring article was 13.4 (Religious Studies).

No strong correlation was found between the NCC and its readability score. Literature was found to have a relationship between NCC and readability, though this association is extremely weak.

Table 4

Pearson coeff. analysis of normalized citation count and readability score by academic field.

Academic Field	*FRE r2 value*	*GFI r2 value*	*FRE p value*	*GFI p value*
Political Science	0.1	0.1	0.2	0.3
Communication Studies	0.0	0.0	0.9	1.0
Literature	0.2	0.1	0.0*	0.2
Religious Studies	0.1	0.1	0.3	0.2
Education	0.0	0.0	0.6	0.6

* positive correlation

Discussion

Based on literature, readability will be operationalized as attaining a 50 or above on the FRE and a grade 8 level of text difficulty on the GFI (Kirkpatrick et al., 2017; Wasike, 2018; Dubé & Lapane, 2014). An FRE score of 50 or above is considered best for effective communication among "non-subject specialists" (Kirkpatrick et al., 2017, p. 3) and a GFI score of 8 is needed for "near-universal understanding" (Lynch et al., 2017, p. 127). Articles failing to meet this combined requirement will be deemed poor in readability. No articles met this requirement. Only 1% of articles achieved a 50 or above on the FRE, and no articles gained an 8 or lower on the GFI.

Results compared to other fields

How do these scores fare with other fields? Results from this study show that scores do not stray far from typical academic prose common to many other fields. Health science fields, such as internal medicine, surgery, and neuroimaging, have been found to score in the 20s to low 30s on the FRE (Yeung, Goto, & Leung, 2018). This study is also congruent with the little research that exists on the readability of the humanities and social sciences. Political science, for instance, has been found to have a mean score of 33 on the FRE (Cann, Goelzhauser, & Johnson, 2014). While older research also reports that humanities and social science articles received a 25.4 and 17.4 on the FRE, respectively (Hartley, Sotto, & Fox, 2004). Complicated prose is not exclusive to traditional academic texts. For example, articles in marketing journals averaged a mid-30 on the FRE (Sawyer, Laran, & Xu, 2008). Similarly, articles in investment and finance journals had a mean FRE score of 30.4 (Lee & French, 2011).

Scores and academic paradigm

Readability research is somewhat inconsistent when scoring two broad academic paradigms: empirically driven research and sense-making scholarship. Empirically driven research tends to follow along traditional lines, such as formal methods and approaches, quantifiable data, and a structured format (Sawyer et al., 2008; Cann et al., 2014). Sense-making, however, often involves an exploratory approach, more ambiguous language, and analytical frameworks to assess complex issues (Sawyer et al., 2008; Hartley et al., 2004). Differences in readability scores between these paradigms, if there are any, tend to be nominal. Some researchers found a 5-point difference in empirically driven versus sense-making articles, with the former being more readable (Cann et al., 2014). One possible explanation is that numbers can give a false-positive because they are coded as monosyllabic, and thus the subsequent readability scores are a mere "artifactual relationship" and not inherent to the academic paradigm (Sawyer et al., 2008, p. 109). Other researchers found empirically driven writing was marginally less readable (Hartley et al., 2004), while others found that it made no difference (Sawyer et al., 2008). The same inconsistency was found in this study. Political Science, often associated with empirically driven research, tied with Literature to receive the worst GFI score. In a likewise manner, two sense-making fields, Literature and Religious

Studies, received the highest and lowest FRE scores, respectively. However, the selected fields in this study can blur the lines that delineate the two broad academic paradigms. For example, Communication Studies can use empirical methods to analyze the frequency of social media posts or employ semiotic analyses to explore the content of them.

Readability and NCC

Similarly, no strong relationship between normalized citation count (NCC) and readability was found in this study, only a weak positive correlation with Literature. While this study corroborated other findings (Swayer et al. 2008), other research shows that there can be a positive correlation (McCannon, 2019) and a negative correlation (Yeung et al. 2018). McCannon (2019), however, analyzed all articles from the same journal over a nine-year span and found that, when articles are compared against each other irrespective of NCC, data indicate that research tends to be cited more if it is more readable. This is significant for one major reason: Citation numbers was a common sampling method employed in many studies—including this one, the argument being that citation count is representative of an article's impact. However, collecting the most cited articles may bias data because there is sometimes a negative correlation between readability and higher numbers of citations.

Implications for practitioners

Poor readability is a barrier to KMb at the professional level as well. In the present study, Education's FRE (30.2) and GFI (17.6) scores are rated more readable than the total mean readability scores. Education therefore represents an above average level of readability, yet even this field is a barrier to its own practitioners. Researchers and educators alike have criticized this field because "papers are not written for practical application" (See, Gorard, & Siddiqui, 2016, p. 69). When teachers were asked to read and apply academic research to their own pedagogy, complaints emerged surrounding the research's poor writing, with one participant bemoaning, "I need a translator to understand what this article is saying" (See et al., 2016, p. 65). This poor readability prevented educators from implementing pedagogical research and practices discovered and discussed in academia (See et al., 2016). In other

cases, practitioners misunderstood findings and thought they already were incorporating them into the classroom (See et al., 2016). These issues do not result from the discussion of inherently complex scientific theories, "but of poor or overly complex writing" (See et al., 2016, p. 65). If readers possess the requisite education, then why is there difficulty? Although a reader may have sufficient education according to a readability test, prolonged reading of complex texts can cause mental fatigue (Choudhry et al., 2016) and hinder comprehension further as more time is spent reading. Further, Education's above-average status is only marginal compared to the total mean scores.

Journalists have received criticism for how they handle their inherited role of knowledge brokers. One prominent example comes from a study that concluded the use of parachutes while jumping from a plane had no effect on the health compared to others who only used empty backpacks (Harris, 2018). The study in question was satirical, though some journalists took the bait and failed to read that all participants of the study jumped from planes still on the ground (Harris, 2018). The poor readability of scholarly writing has some responsibility to bear. Clear and readable prose can help mitigate these issues and promote KMb.

Readability and academic status

If readability is so poor, why does it remain so? Poor readability is associated with scholarly prestige, credibility, and status (Lee & French, 2011). In a seminal study, Armstrong (1980) gave faculty participants different excerpts with varying degrees of readability. Results from Armstrong's (1980) study report that participants favoured the less readable excerpts, despite all excerpts having the same content. In fact, pursuing readability may actually harm one's academic standing as findings conclude "more easily read articles were rated less prestigious" (Sawyer et al., 2008, p. 110). This coupling of academic prose and prestige is extended to scholarly journals. For example, some journals tend to accept more manuscripts with poor readability while rejecting manuscripts with better readability (Lee & French, 2011). Other findings indicate that journals with higher impact factors tended to have worse readability (Yeung et al., 2018). This further perpetuates poor readability in scholarship as communication is sidelined in favour of appearing credible and impressing the reader (Lee & French, 2011). Some scholars postulate that poor readability may simply be a characteristic of academic prose (Whitt & Creech, 1983). To uncouple the notion of poor writing with prestigious scholarship,

C.W. Mills (2000) recommends: "To overcome the academic *prose* you have first to overcome the academic *pose*" (p. 219). One strategy Mills (2000) proposes to overcome the academic pose is assume your audience has the right to know, then write accordingly.

Readability for understandability

Aggressive actions toward improving readability based on scores alone can result in less readable prose. In particular, "gaming" the system and blindly chasing favourable readability scores can sacrifice a text's understandability (Kirkpatrick et al., 2017). For example, readability tests favour shorter sentences and smaller words, but carelessly over-editing can impede understandability (Crossley et al., 2008). Readable prose, instead, comes from the judgment of a proficient writer, not a simplistic formula.

Limitations and weaknesses

Readability tests only analyze the number of words, complex words, and sentence length, and cannot evaluate extra-linguistic content, such as the structure. For example, subheadings help organize the flow of content by breaking up sections into manageable and logical chunks (Kirkpatrick et al., 2017). Extra-linguistic content such as this was not considered in the readability analyses. Similarly, readability tests cannot assess writing styles, such as grammaticality, syntactic coherence, or word choice.

Likewise, readability formulae cannot analyze symbols, such as equations, specialized notations, and glyphs (Benjamin, 2012). Figures, graphs, charts, and images were also not considered in the analysis, despite these elements especially being able to clearly illustrate research (Hartley, 2016).

Using the most cited articles as samples proved the most useful. However, citations do not necessarily reflect the articles' impact on laypeople—a key audience who would be especially disadvantaged with poor readability. Additionally, citation numbers were found to not always be consistent in Google Scholar Metrics. For example, Google bibliometric data could differ if an article was found through Metrics versus a manual search query.

To the authors' best knowledge, no bibliometric database could account for citation numbers, date of publication, and social sciences and humanities disciplines simultaneously to a sufficient degree. Generally, bibliometric data-

bases cater to STEM and health sciences fields and offer much richer, multi-dimensional services. Therefore, there could be no diachronic analysis to assess how readability scores change over time. Future research could analyze readability over an extended period of time to note of any changes if there are any.

Another finding should serve as a cautionary warning to further readability research; many different applications produce different scores that are not consistent with each other (See Appendix D). This finding echoes Hartley (2016) and calls into question much of the research that has been done already on readability because there is no consensus among researchers on using a particular website or application.

Conclusion and future research

Academic prose falls far below readable levels for the average person. Regularly pitched beyond undergraduate reading levels, no article achieved the operationalized readability standard for this study. In fact, 99% of articles did not meet the FRE threshold, and no articles met the GFI threshold. These scores are, however, in line with scores seen from STEM and health science fields.

The articles in Political Science, Communication Studies, Literature, Religious Studies, and Education shared similar readability scores. The greatest variation was not across fields, but within them. Scores within fields had margins of over 40 points on the FRE in some cases and over 10 on the GFI. No research has found a margin this dramatic across fields.

The relationship between NCC, readability, and academic field is uncertain. Research shows that many factors may influence readability, such as journal impact factor, academic prestige, time, and writing skill. Many KMb strategies rely on knowledge producers treating readability as important as the research itself. However, researchers focus on outputs, not making it accessible.

Readability tests are an imperfect solution to a complex issue. Rather than relying on readability tests as editors, they serve a more honest role as a reflective tool for communicators. The scores derived from the formulae can hint that a writer may need a course correct. However, readability tests are quite limited as they cannot assess the quality of writing, grammar, or factual elements. Readability tests cannot replace an experienced editor trained specifically for KMb.

Readability of scholarly articles also has implications for pedagogy. If undergraduate students are tasked with reading and understanding scholarly articles, yet their reading levels fall short, then many may struggle with the expectations of post-secondary curricula. In this study, most articles analyzed were beyond the undergraduate reading level. If first-year undergraduates are taught using scholarly articles far above the expected reading level, much of the curricula may require transformation to adequately teach and instruct students. However, it can be argued that sound pedagogical practice involves challenging students, building on that, and guiding them to self-sustainability, a concept known as scaffolding in education. This concept, however, requires resources for those who struggle.

How or if readability analyses should be integrated into edited, academic prose merits its own research. Guides, best practices, and tangible solutions to the chronic issue of poor readability are likely discipline specific and require several other considerations, such as target audience, medium, and the socio-cultural space the message will inhabit.

Readability research does not belong to any single field of research. For example, readability research can employ a variety of cross-discipline approaches, such as eye-tracking tests to determine if readers require reparsing an excerpt; electroencephalography (EEG) to capture event-related potentials while parsing passages; maze or cloze procedures to assess reader comprehension; or focus groups to discuss with participants what their attitudes are toward one excerpt compared with another.

References

Armstrong, J. S. (1980). Unintelligible management research and academic prestige. *Interfaces*, 10(2), 80-6.

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

Cain, K., Shore, K., Weston, C., & Sanders, C. B. (2018). Knowledge Mobilization as a Tool of Institutional Governance: Exploring Academics' Perceptions of "Going Public". *Canadian Journal of Higher Education*, 48(2), 39-54.

Cann, D. M., Goelzhauser, G., & Johnson, K. (2014). Analyzing Text Complexity in Political Science Research. *PS: Political Science & Politics*, 47(3), 663-666.

Cheek, F. E., & Rosenhaupt, M. (1968). Are Sociologists Incomprehensible? An Objective Study. *American Journal of Sociology*, 73(5), 617-627.

Choudhry, A. J., Baghdadi, Y. M., Wagie, A. E., Habermann, E. B., Heller, S. F., Jenkins, D. H., ... & Zielinski, M. D. (2016). Readability of Discharge Summaries: With what Level of Information are we Dismissing our Patients?. *The American Journal of Surgery*, 211(3), 631-636.

Chubb, J. & Watermeyer, R. (2017). Artifice or Integrity in the Marketization of Research impact? Investigating the Moral Economy of (Pathways to) Impact Statements Within Research Funding Proposals in the UK and Australia. *Studies in Higher Education*, 42(12), 2360-2372, DOI: 10.1080/03075079.2016.1144182

Cooper, A., Rodway, J., & Read, R. (2018). Knowledge mobilization practices of educational researchers across Canada. *Canadian Journal of Higher Education*, 48(1), 1-21.

Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing text readability using cognitively based indices. *Tesol Quarterly*, 42(3), 475-493.

Dubé, C. E., & Lapane, K. L. (2014). Lay abstracts and summaries: Writing advice for scientists. *Journal of Cancer Education*, 29(3), 577-579.

Ficzere, C. H., Hagan, A., Ness, G. L., Greene, E. M., & Hill, K. (2016). Comparison of reading levels of pharmacy students and reading level of primary literature. *INNOVATIONS in pharmacy*, 7(3).

Hartley, J. (2016). Is time up for the flesch measure of reading ease?. *Scientometrics*, 107(3), 1523-1526.

Hartley, J., Sotto, E., & Fox, C. (2004). Clarity across the disciplines: An analysis of texts in the sciences, social Sciences, and arts and humanities. *Science Communication*, 26(2), 188-210.

Harris, R. (2018). Researchers show parachutes don't work, but there's a catch. *NPR*. Retrieved from https://www.npr.org/sections/health shots/2018/12/22/679083038/researchers-show-parachutes-dont-work-but-there-s-a-catch

Kirkpatrick, E., Gaisford, W., Williams, E., Brindley, E., Tembo, D., & Wright, D. (2017). Understanding plain english summaries. A comparison of two approaches to improve the quality of plain english summaries in research reports. *Research Involvement and Engagement*, 3(17), 1-14.

Lee, S., & French, N. (2011). The readability of academic papers in the journal of property investment & finance. *Journal of Property Investment & Finance*, 29(6), 693–704. https://doi-org.libaccess.lib.mcmaster.ca/10.1108/14635781111171814

Lynch, N. P., Lang, B., Angelov, S., McGarrigle, S. A., Boyle, T. J., Al-Azawi, D., & Connolly, E. M. (2017). Breast reconstruction post mastectomy—Let's Google it. Accessibility, readability and quality of online Information. *The Breast*, 32, 126-129.

Matheson, K., & Malcolm Edwards, C. (2016). Perspectives on knowledge mobilization: An introduction to the special issue. *Technology Innovation Management Review*, 6(9) 4-8.

McCannon, B. C. (2019). Readability and Research Impact. Economics Letters, 180, 76-79.

Mills, C. (2000). *The sociological imagination*. Oxford [England]: Oxford University Press.

Modiri, O., Guha, D., Alotaibi, N. M., Ibrahim, G. M., Lipsman, N., & Fallah, A. (2018). Readability and quality of wikipedia pages on neurosurgical topics. *Clinical Neurology and Neurosurgery*, 166, 66-70.

Moore, G., Fitzpatrick, T., Lim-Carter, I., Haynes, A., Flego, A., & Snelgrove, B. (2016). Implementing knowledge translation strategies in funded research in Canada and Australia: A case study. *Technology Innovation Management Review*, 6(9).

Nichols, N., Phipps, D., Provencal, J. & Hewitt, A. (2013). Knowledge mobilization, collaboration, and social innovation: Leveraging investments in higher education. *Canadian Journal of Nonprofit and Social Economy Research*, 4(1), 25-42.

Sá, C. M., Li, S. X., & Faubert, B. (2011). Faculties of education and institutional strategies for knowledge mobilization: An exploratory study. *Higher Education*, 61(5), 501-512.

Sawyer, A. G., Laran, J., & Xu, J. (2008). The readability of marketing journals: Are award-winning articles better rritten?. *Journal of Marketing*, 72(1), 108-117.

See, B. H., Gorard, S., & Siddiqui, N. (2016). Teachers' Use of Research Evidence in Practice: A Pilot Study of Feedback to Enhance Learning. Educational Research, 58(1), 56-72.

Selvin, H. C., & Wilson, E. K. (1984). On sharpening sociologists' prose. *The Sociological Quarterly*, 25(2), 205-222.

Shields, J., & Evans, B. (2012). Building a policy-oriented research partnership for knowledge mobilization and knowledge transfer: The case of the Canadian metropolis project. *Administrative Sciences*, 2(4), 250-272.

Van Boom, W. H., Desmet, P., & Van Dam, M. (2016). "If it's easy to read, it's easy to claim" — The effect of the readability of insurance contracts on consumer expectations and conflict behaviour. *Journal of Consumer Policy*, 39(2), 187-197.

Walfish, S., & Watkins, K. M. (2005). Readability level of health insurance portability and accountability act notices of privacy practices utilized by academic medical centers. *Evaluation & The Health Professions*, 28(4), 479-486.

Wasike, B. (2018). Preaching to the choir? An analysis of newspaper readability vis-a-vis public literacy. *Journalism*, 19(11), 1570-1587.

Whitt, H. P., & Creech, J. C. (1983). Gunning down the fog: A test of the unintelligibility and illiteracy hypotheses. *Mid-American Review of Sociology*, 39-49.

Yeung, A. W. K., Goto, T. K., & Leung, W. K. (2018). Readability of the 100 most-cited neuroimaging papers assessed by common readability formulae. *Frontiers In Human Neuroscience*, 12, 308.

Young, N., Nguyen, V. M., Corriveau, M., Cooke, S. J., & Hinch, S. G. (2016). Knowledge users' perspectives and advice on how to improve knowledge exchange and mobilization in the case of a co-managed fishery. *Environmental Science & Policy*, 66, 170-178. doi:10.1016/j.envsci.2016.09.002

Zethsen, K. K. (2018). Access is not the same as understanding. Why intralingual translation is crucial in a world of information overload. *Across Languages and Cultures*, 19(1), 79-98.

Appendix A

Table A1

Peer-reviewed articles ordered by field and highest to lowest number of citations

Political Science Articles:	***Number of Citations (as of July, 2019)***
King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. American *Political Science Review*, 107(2), 326-343.	1,000
Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267-297.	972
Fukuyama, F. (2013). What is governance?. *Governance*, 26(3), 347-368.	688
Schmidt, V. A. (2013). Democracy and legitimacy in the European Union revisited: Input, output and 'throughput'. *Political Studies*, 61(1), 2-22.	528
Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2), 495-510.	526
Bakker, R., De Vries, C., Edwards, E., Hooghe, L., Jolly, S., Marks, G., ... & Vachudova, M. A. (2015). Measuring party positions in Europe: The Chapel Hill expert survey trend file, 1999–2010. *Party Politics*, 21(1), 143-152.	514
Boix, C., & Svolik, M. W. (2013). The foundations of limited authoritarian government: Institutions, commitment, and power-sharing in dictatorships. *The Journal of Politics*, 75(2), 300-316.	496
Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review*, 107(1), 57-79.	430
Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59(3), 690-707.	418

Boix, C., Miller, M., & Rosato, S. (2013). A complete data set of political regimes, 1800–2007. *Comparative Political Studies*, 46(12), 1523-1554.	403
Stegmueller, D. (2013). How many countries for multilevel modeling? A comparison of frequentist and Bayesian approaches. *American Journal of Political Science*, 57(3), 748-761.	387
Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, 16, 101-127.	387
Persson, A., Rothstein, B., & Teorell, J. (2013). Why anticorruption reforms fail—systemic corruption as a collective action problem. *Governance*, 26(3), 449-471.	387
Hainmueller, J., & Hopkins, D. J. (2014). Public attitudes toward immigration. *Annual Review of Political Science*, 17, 225-249.	362
Hanmer, M. J., & Ozan Kalkan, K. (2013). Behind the curve: Clarifying the best approach to calculating predicted probabilities and marginal effects from limited dependent variable models. *American Journal of Political Science*, 57(1), 263-277.	332
Mudde, C. (2013). Three decades of populist radical right parties in Western Europe: So what?. *European Journal of Political Research*, 52(1), 1-19.	320
Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*, 23(1), 76-91.	311
Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, 58(2), 367-386.	299
Bormann, N. C., & Golder, M. (2013). Democratic electoral systems around the world, 1946–2011. *Electoral Studies*, 32(2), 360-369.	292
Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., ... & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064-1082.	291

Table A2

Peer-reviewed articles ordered by field and highest to lowest number of citations

Communication Studies Articles	***Number of Citations (as of July, 2019)***
Leonardi, P. M., Huysman, M., & Steinfield, C. (2013). Enterprise social media: Definition, history, and prospects for the study of social technologies in organizations. *Journal of Computer-Mediated Communication*, 19(1), 1-19.	498
Gerlitz, C., & Helmond, A. (2013). The like economy: Social buttons and the data-intensive web. *New Media & Society*, 15(8), 1348-1365.	441
Van Dijck, J. (2013). 'You have one identity': Performing the self on Facebook and LinkedIn. *Media, Culture & Society*, 35(2), 199-215.	431
Van Deursen, A. J., & Van Dijk, J. A. (2014). The digital divide shifts to differences in usage. *New Media & Society*, 16(3), 507-526.	426
Ellison, N. B., Vitak, J., Gray, R., & Lampe, C. (2014). Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes. *Journal of Computer-Mediated Communication*, 19(4), 855-870.	412
Couldry, N., & Hepp, A. (2013). Conceptualizing mediatization: Contexts, traditions, arguments. *Communication Theory*, 23(3), 191-202.	368
Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The "nasty effect:" Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19(3), 373-387.	358
Sundar, S. S., & Limperos, A. M. (2013). Uses and grats 2.0: New gratifications for new media. *Journal of Broadcasting & Electronic Media*, 57(4), 504-525.	326

Majchrzak, A., Faraj, S., Kane, G. C., & Azad, B. (2013). The contradictory influence of social media affordances on online communal knowledge sharing. *Journal of Computer-Mediated Communication*, 19(1), 38-55.	308
Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communication*, 18(2), 154-174.	291
John, N. A. (2013). Sharing and Web 2.0: The emergence of a keyword. *New Media & Society*, 15(2), 167-182.	286
Marwick, A. E., & Boyd, D. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, 16(7), 1051-1067.	285
Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics*, 46(1), 8-23.	273
Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication*, 64(2), 317-332.	273
Nah, S., & Saxton, G. D. (2013). Modeling the adoption and use of social media by nonprofit organizations. *New Media & society*, 15(2), 294-313.	271
Utz, S., Schultz, F., & Glocka, S. (2013). Crisis communication online: How medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster. *Public Relations Review*, 39(1), 40-46.	261
Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, 63(2), 221-243.	252

Ceron, A., Curini, L., Iacus, S. M., & Porro, G. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society*, 16(2), 340-358.	252
Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social media as information source: Recency of updates and credibility of information. *Journal of Computer-Mediated Communication*, 19(2), 171-183.	250
Hermida, A., Lewis, S. C., & Zamith, R. (2014). Sourcing the Arab Spring: A case study of Andy Carvin's sources on Twitter during the Tunisian and Egyptian revolutions. *Journal of Computer-Mediated Communication*, 19(3), 479-499.	248

Table A3

Peer-reviewed articles ordered by field and highest to lowest number of citations

Literature Articles	***Number of Citations (as of July, 2019)***
Latour, B. (2014). Agency at the Time of the Anthropocene. *New Literary History*, 45(1), 1-18.	341
Ryan, M. L. (2013). Transmedial storytelling and transfictionality. *Poetics Today*, 34(3), 361-388.	124
Hodder, I. (2014). The entanglements of humans and things: A long-term view. *New Literary History*, 45(1), 19-36.	100
Nielsen, H. S., Phelan, J., & Walsh, R. (2015). Ten theses about fictionality. *Narrative*, 23(1), 61-73.	60
Birke, D., & Christ, B. (2013). Paratext and digitized narrative: Mapping the field. *Narrative*, 21(1), 65-87.	56
Gubar, M. (2013). Risky Business: Talking about Children in Children's Literature Criticism. *Children's Literature Association Quarterly*, 38(4), 450-457.	55

Hayles, N. K. (2014). Cognition everywhere: The rise of the cognitive nonconscious and the costs of consciousness. *New Literary History*, 45(2), 199-220.	53
Gill, R. B. (2013). The uses of genre and the classification of speculative fiction. *Mosaic: A Journal for The Interdisciplinary Study of Literature*, 71-85.	52
Goldstone, A., & Underwood, T. (2014). The quiet transformations of literary studies: What thirteen thousand scholars could tell us. *New Literary History*, 45(3), 359-384.	51
Saldívar, R. (2013). The Second Elevation of the Novel: Race, Form, and the Postrace Aesthetic in Contemporary Narrative. *Narrative*, 21(1), 1-18.	50
McCracken, E. (2013). Expanding Genette's epitext/peritext model for transitional electronic literature: centrifugal and centripetal vectors on kindles and iPads. *Narrative*, 21(1), 105-124.	48
Pederson, J. (2014). Speak, trauma: Toward a revised understanding of literary trauma theory. *Narrative*, 22(3), 333-353.	46
Meretoja, H. (2014). Narrative and human existence: Ontology, epistemology, and ethics. *New Literary History*, 45(1), 89-109.	39
Cheah, P. (2014). World against globe: Toward a normative conception of world literature. *New Literary History*, 45(3), 303-329.	39
Jagoda, P. (2013). Gamification and other forms of play. *Boundary 2*, 40(2), 113-144.	37
Savarese, R. J., & Zunshine, L. (2014). The critic as neurocosmopolite; Or, what cognitive approaches to literature can learn from disability studies: Lisa Zunshine in conversation with Ralph James Savarese. *Narrative*, 22(1), 17-44.	35
Kuzmičová, A. (2014). Literary narrative and mental imagery: A view from embodied cognition. *Style*, 48(3), 275-293.	32

Bernaerts, L., Caracciolo, M., Herman, L., & Vervaeck, B. (2014). The storied lives of non-human narrators. *Narrative*, 22(1), 68-93.	32
Nail, T. (2017). What is an Assemblage?. *SubStance*, 46(1), 21-37.	30
Vanwynsberghe, H., & Verdegem, P. (2013). Integrating social media in education. *ClCWeb-Comparative Literature and Culture*, 15(3).	29

Table A4
Peer-reviewed articles ordered by field and highest to lowest number of citations

Religious Studies Articles	***Number of Citations (as of July, 2019)***
Bonelli, R. M., & Koenig, H. G. (2013). Mental disorders, religion and spirituality 1990 to 2010: a systematic evidence-based review. J*ournal of Religion and Health*, 52(2), 657-673.	238
Ammerman, N. T. (2013). Spiritual but not religious? Beyond binary choices in the study of religion. *Journal for the Scientific Study of Religion*, 52(2), 258-278.	198
Exline, J. J., Pargament, K. I., Grubbs, J. B., & Yali, A. M. (2014). The Religious and Spiritual Struggles Scale: Development and initial validation. *Psychology of Religion and Spirituality*, 6(3), 208.	139
Ivtzan, I., Chan, C. P., Gardner, H. E., & Prashar, K. (2013). Linking religion and spirituality with psychological well-being: Examining self-actualisation, meaning in life, and personal growth initiative. *Journal of Religion and Health*, 52(3), 915-929.	137
Benefiel, M., Fry, L. W., & Geigle, D. (2014). Spirituality and religion in the workplace: History, theory, and research. *Psychology of Religion and Spirituality*, 6(3), 175.	106
Smith, J. M. (2013). Creating a godless community: The collective identity work of contemporary American atheists. *Journal for the Scientific Study of Religion*, 52(1), 80-99.	104

Shonin, E., Van Gordon, W., & Griffiths, M. D. (2014). The emerging role of Buddhism in clinical psychology: Toward effective integration. *Psychology of Religion and Spirituality*, 6(2), 123.	100
Xygalatas, D. (2013). Effects of religious setting on cooperative behavior: A case study from Mauritius. *Religion, Brain & Behavior*, 3(2), 91-102.	99
Vieten, C., Scammell, S., Pilato, R., Ammondson, I., Pargament, K. I., & Lukoff, D. (2013). Spiritual and religious competencies for psychologists. *Psychology of Religion and Spirituality*, 5(3), 129.	99
Davis, D. E., Worthington Jr, E. L., Hook, J. N., & Hill, P. C. (2013). Research on religion/spirituality and forgiveness: A meta-analytic review. *Psychology of Religion and Spirituality*, 5(4), 233.	96
Aldwin, C. M., Park, C. L., Jeong, Y. J., & Nath, R. (2014). Differing pathways between religiousness, spirituality, and health: A self-regulation perspective. *Psychology of Religion and Spirituality*, 6(1), 9.	94
Currier, J. M., Mallot, J., Martinez, T. E., Sandy, C., & Neimeyer, R. A. (2013). Bereavement, religion, and posttraumatic growth: A matched control group investigation. *Psychology of Religion and Spirituality*, 5(2), 69.	93
Reddish, P., Bulbulia, J., & Fischer, R. (2014). Does synchrony promote generalized prosociality?. *Religion, Brain & Behavior*, 4(1), 3-19.	80
Schjoedt, U., Sørensen, J., Nielbo, K. L., Xygalatas, D., Mitkidis, P., & Bulbulia, J. (2013). Cognitive resource depletion in religious interactions. Religion, *Brain & Behavior*, 3(1), 39-55.	77
Davis, E. B., Moriarty, G. L., & Mauch, J. C. (2013). God images and god concepts: Definitions, development, and dynamics. *Psychology of Religion and Spirituality*, 5(1), 51.	77

Unterrainer, H. F., Lewis, A. J., & Fink, A. (2014). Religious/spiritual well-being, personality and mental health: A review of results and conceptual issues. *Journal of Religion and Health*, 53(2), 382-392.	76
Shonin, E., Van Gordon, W., & Griffiths, M. D. (2014). Meditation awareness training (MAT) for improved psychological well-being: A qualitative examination of participant experiences. *Journal of Religion and Health*, 53(3), 849-863.	73
LeDrew, S. (2013). Discovering atheism: Heterogeneity in trajectories to atheist identity and activism. *Sociology of Religion*, 74(4), 431-453.	69
Kul, S., Savaş, E., Öztürk, Z. A., & Karadağ, G. (2014). Does Ramadan fasting alter body weight and blood lipids and fasting blood glucose in a healthy population? A meta-analysis. *Journal of Religion and Health*, 53(3), 929-942.	68
Padela, A. I., & Curlin, F. A. (2013). Religion and disparities: Considering the influences of Islam on the health of American Muslims. *Journal of Religion and Health*, 52(4), 1333-1345.	63

Table A5

Peer-reviewed articles ordered by field and highest to lowest number of citations

Education Articles	***Number of Citations (as of July, 2019)***
Liyanagunawardena, T., Adams, A., & Williams, S. (2013). MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distributed Learning*, 14(3), 202-227.	828
Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, 50(1), 4-36.	807
Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, 83(3), 357-385.	748

Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. *Educational Researcher*, 42(1), 38-43.	582
Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, 15(1).	526
Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, 38(5), 758-773.	489
Kim, C., Kim, M. K., Lee, C., Spector, J. M., & DeMeester, K. (2013). Teacher beliefs and technology integration. *Teaching and Teacher Education*, 29, 76-85.	464
Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70-120.	421
Hanushek, E. A. (2013). Economic growth in developing countries: The role of human capital. *Economics of Education Review*, 37, 204-212.	365
Cheung, A. C., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review*, 9, 88-113.	365
Cook, B. G., & Odom, S. L. (2013). Evidence-based practices and implementation science in special education. *Exceptional children*, 79(2), 135-144.	364
Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, 115(3), 1-47.	363
Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review*, 12, 45-58.	340
Hanushek, E. A. (2013). Economic growth in developing countries: The role of human capital. *Economics of Education Review*, 37, 204-212.	330

Wang, X. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, 50(5), 1081-1121.	328
Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144.	316
Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, 28, 12-23.	310
McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education*, 64(5), 378-386.	306
Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44(4), 237-251.	303
Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79-122.	290

Appendix B

Table B1

Academic fields listed alphabetically

1. Anthropology	9. Linguistics
2. Communication Studies*	10. Media Studies
3. Economics	11. Performing Arts
4. Education*	12. Political Science*
5. Geography	13. Psychology
6. History	14. Religious Studies*
7. Law	15. Sociology
8. Literature*	16. Visual Arts

* Asterisks indicate disciplines selected by the random number generator.

Appendix C

Table C1
Number of FRE articles' scores by range and field

Academic Field	*FRE Score Range 0-9*	*FRE Score Range 10-19*	*FRE Score Range 20-29*	*FRE Score Range 30-39*	*FRE Score Range 40-49*	*FRE Score 50+*
Political Science	1	1	8	10	0	0
Communication studies	0	3	7	9	0	1
Literature	0	2	6	9	3	0
Religious Studies	1	4	9	5	1	0
Education	0	3	6	9	2	0

Table C2
Number of GFI articles' scores by range and field

Academic Field	*GFI Score Range 12-13*	*GFI Score Range 14-15*	*GFI Score Range 16-17*	*GFI Score Range 18-19*	*GFI Score Range 20-21*	*GFI Score 22+*
Political Science	0	3	6	9	0	2
Communication studies	0	1	13	4	2	0
Literature	0	2	8	4	5	1
Religious Studies	1	2	5	9	1	2
Education	0	5	7	5	3	0

Appendix D

To test the consistency between commonly used readability test applications, the abstract of this study was the dependent variable while the readability test application was the independent variable. Results show that there can be considerable variability depending on the application used. This calls into question the methods used by readability research to date,

Table D1
Different Readability Test Applications Compared for Replicability

Readability Test Application	*FRE Score*	*GFI Score*
Microsoft Word (version 2012)	21.8	n/a
http://www.readabilityformulas.com/free-readability-formula-tests.php	33.2	17.3
https://www.online-utility.org/english/readability_test_and_improve.jsp	29.5	15.3
http://gunning-fog-index.com/	n/a	16.7
https://www.webfx.com/tools/read-able/	33.2	17.3
https://readable.com/text/	26	17.8
https://www.prepostseo.com/readability-checker	14.5	n/a
https://www.perrymarshall.com/grade/	n/a	17.3