Scientometrics (2021) 126:8471–8490
https://doi.org/10.1007/s11192-021-04112-9

Do research articles with more readable abstracts receive higher online attention? Evidence from *Science*

Tan Jin[1] · **Huiqiong Duan**[1] · **Xiaofei Lu**[2] · **Jing Ni**[3] · **Kai Guo**[4]

Received: 12 January 2021 / Accepted: 22 July 2021 / Published online: 5 August 2021

Abstract
The value of scientific research is manifested in its impact in the scientific community as well as among the general public. Given the importance of abstracts in determining whether research articles (RAs) may be retrieved and read, recent research is paying attention to the effect of abstract readability on the scientific impact of RAs. However, to date little research has looked into the effect of abstract readability on the impact of RAs among the general public. To address this gap, this study reports on an investigation into the relationship between abstract readability and online attention received by RAs. Our dataset consisted of the abstracts of 550 RAs from 11 disciplines published in *Science* in 2012 and 2018. Thirty-nine lexical and syntactic complexity indices were employed to measure the readability of the abstracts, and the *Altmetric* attention scores of the RAs were used to measure the online attention they received. Results showed that abstract readability is significantly related to the online attention RAs receive, and that this relationship is significantly affected by discipline and publication time. Our findings have useful implications for making RA abstracts accessible to the general public.

Keywords Abstract readability · Altmetrics · Online attention · Research articles

Introduction

With the rise of search engines and online databases, traditional approaches to accessing scientific research outputs such as browsing print journals are increasingly being replaced by keyword searches and online browsing (Houghton et al., 2004; Thelwall et al., 2013). In searching for and browsing information from research outputs, readers often use the

✉ Kai Guo
kaiguo@connect.hku.hk

[1] School of Foreign Languages, Sun Yat-Sen University, No. 135, Xingang Xi Road, Guangzhou 510275, China

[2] Department of Applied Linguistics, The Pennsylvania State University, 234 Sparks Building, University Park, PA 16802, USA

[3] Faculty of Nursing, Jiujiang University, No. 320, Xunyang East Road, Jiujiang 332000, China

[4] Faculty of Education, The University of Hong Kong, Pokfulam Road, Hong Kong 999077, China

titles and abstracts of specific research outputs to obtain a preliminary understanding of these outputs and locate materials of interest to them (Pinto & Lancaster, 1999). Widely recognized as "an abbreviated, accurate representation of a document" (Weil, 1970, p. 352), the abstract of a research article (RA) provides a synopsis of the main ideas, research design, and findings of the research that can help accelerate the readers' grasp of the essential information about the RA and facilitate their judgement of whether it is sufficiently relevant or interesting to warrant access to and reading of the full text (Gazni, 2011). As such, it is not surprising that abstracts tend to be read more frequently than other RA parts (Nicholas et al., 2003).

Meanwhile, thanks to the open access movement of scholarly publishing and the rapid rate of online knowledge dissemination, RAs are now not only referenced by scholars for academic purposes, but also read by an increasingly wide range of non-expert readers seeking information for various purposes (Sud & Thewal, 2014). It is thus more important than ever for RAs to be accessible to readers in both the scientific community and the general public in an era of information overload (Tankó, 2017). Among the factors that may affect readers' comprehension of RA abstracts, abstract readability has been found to play a fairly prominent role (Dronberger & Kowitz, 1975). At the same time, researchers have reported a lower level of readability of RA abstracts in comparison to full RA texts (King, 1976) and an upward trend of RA abstract difficulty (e.g., Lei & Yan, 2016). In light of the importance of RA abstracts in the scientific reach of RAs, recent research has started to pay attention to the relationship of abstract readability and the scientific impact of RAs (e.g., Didegah & Thelwall, 2013; Gazni, 2011; Sienkiewicz & Altmann, 2016).

Importantly, if RA abstracts prove difficult to read to the scientific community, they may be even more so to non-expert readers (Bornmann, 2014), and abstract readability may have an effect on the reach and impact of RAs among the general public (Lei & Yan, 2016). Exciting efforts to improve research accessibility are emerging. For example, the Open Accessible Summaries in Language Studies (OASIS) initiative "aims to make research into language learning and teaching openly available and easily accessible to anyone who might be interested for professional or other reasons" by asking authors of published RAs to provide a one-page summary in non-technical language (OASIS, 2020). The increasing popularity of the OASIS initiative among language-related academic journals speaks to the recognition of the effect of abstract readability on the dissemination of research outputs to the general public. However, systematic examinations of the relationship between abstract readability and the impact of RAs among general non-expert readers are scant. An understanding of this relationship will provide empirical evidence for the need for such initiatives as OASIS as well as insight into the specific aspects of language use that researchers can attend to in order to improve abstract readability and the social impact of their RAs.

Measuring abstract readability

The readability of a written text is generally understood as the ease with which the text can be read and understood (Klare, 1963). Previous research has shown that the readability of an abstract may be affected by its linguistic features (Crossley et al., 2019) and other factors such as structure and layout (Hartley, 2000; Hartley & Sydes, 1997). Much research on the linguistic aspect of abstract readability has employed traditional readability formulas (Graesser et al., 2011; Klare, 1974), such as the Flesch Reading Ease (FRE; Flesch, 1948) and the Simple Measure of Ginledygook statistics (SMOG; McLaughlin, 1969). Such formulas generally incorporate a measure of word difficulty or familiarity, most commonly number of letters or syllables per word or the proportion of familiar words, and a measure of sentence complexity, most commonly mean sentence length, and generate a single metric that can be used to compare the readability of different texts. Dronberger and Kowitz (1975), for example, observed significant differences between the FRE scores of abstracts and full texts of documents retrieved from the Educational Resources Information Center system and recommended the use of readability formulas to evaluate and modify the difficulty levels of abstracts. Similarly, Lei and Yan (2016) analyzed a corpus of RAs published in four academic journals (i.e., *Scientometrics*, *Journal of Informetrics*, *Research Policy*, and *Research Evaluation*) from 2003 to 2012 using FRE and SMOG and reported a significant gap between the readability of RA abstracts and full RAs.

Despite the widespread usage of readability formulas for measuring the readability and ease of comprehension of written texts, some scholars have questioned their construct validity because of the simplistic nature of the word and sentence difficulty measures they incorporate (Benjamin, 2012; Crossley et al., 2019; Graesser et al., 2011; Lu et al., 2014). In particular, research in cognitive science has modelled reading comprehension as a coordinated process involving multiple levels of linguistic and discursive complexity, including not only lexical and syntactic complexity but also cohesion and sentiment, among others (Graesser et al., 2011; Just & Carpenter, 1980; Kintsch & van Dijk, 1978). Accordingly, recent research has advocated the use of multiple indices that gauge different dimensions of linguistic complexity in the measurement of the formal linguistic aspect of text readability or difficulty (e.g., Stevens et al., 2015; Tankó, 2017; Vajjala & Meurers, 2012). For example, Vajjala and Meurers (2012) reported that the use of a combination of lexical and syntactic complexity indices outperformed traditional readability formulas by a large margin in predicting the grade levels of reading texts.

Recent studies of abstract readability are also starting to resort to indices of lexical and syntactic complexity. For example, Tankó (2017) employed a set of indices of lexical density, lexical sophistication, lexical variation, and clause length to analyze a corpus of RA abstracts and reported that they "have high syntactic complexity and lexical density and contain primarily low frequency words" (p. 42), which, along with their high information content, makes them difficult to process. Lu et al. (2019), while focusing on whole RAs rather than RA abstracts, showed that indices of lexical complexity (e.g., type-token ratio) and syntactic complexity (e.g., number of clauses per sentence) were useful in revealing the relationship between the readability and scientific impact of RAs.

Abstract readability and scientific impact

The scientific impact of research outputs, often measured quantitatively using citation-based analysis, is an important indicator of the value of the research and of the scholarly success of the researchers producing such outputs (Didegah & Thelwall, 2013; Lu et al., 2019; Sud & Thelwall, 2014). Indeed, RAs that are frequently cited are usually those that have laid an important foundation or made critical discoveries upon which new research can be built (D'Angelo & Di Russo, 2019). The citation counts of research outputs are now readily available from journal websites and/or research databases (Sud & Thelwall, 2014). In addition to the subject matter addressed and the quality of the research presented in RAs, many other factors may affect their scientific impact, such as the impact of the journals they are published in and the references they cite, the scope of collaboration involved in the research, and features pertaining to their writing and organization (Chen et al., 2020; Didegah & Thelwall, 2013).

Among the features examined in the last category, the readability or linguistic complexity of RA abstracts has received substantial attention. This is not necessarily surprising. In order for an RA to be cited by other researchers, it needs to first attract their attention and be read by them. The title and abstract are often the first parts of an RA that are read by researchers looking for relevant research sources and references. In this process, the readability or complexity of the abstract may affect their comprehension of the abstract and potentially their decision to pursue the full text or move on to other sources. Several recent studies have explored the relationship between abstract readability and the scientific impact of research outputs as indicated by their citation counts, although the results reported so far have been somewhat inconsistent. Sienkiewicz and Altmann (2016) reported significant positive correlations between the citation counts of RAs and the complexity of abstracts measured using the Gunning fog index, the *z*-index, and Herdan's C measure. Gazni (2011) reported a significant negative linear relationship between the citation counts of published RAs from five prestigious institutions and abstract readability assessed using the Flesch Reading Ease Score. Didegah and Thelwall (2013) considered seven readability formulas and settled on the Flesh Reading Ease Score as a proxy of abstract readability "since it seems to be the most popular and also has a high correlation with the other six scores" (p. 864). However, they reported no significant association between this measure of abstract readability and citation counts of RAs. Notably, most previous studies took discipline and time since publication into account in examining the relationship between abstract readability and scientific impact (D'Angelo & Di Russo, 2019; Htoo & Na, 2017; Maflahi & Thelwall, 2018).

While this line of research has mostly used readability formulas to assess abstract readability, in their analysis of the relationship of whole RA readability to scientific impact, Lu et al. (2019) moved beyond readability formulas but employed a set of lexical and syntactic complexity indices, as mentioned in "Measuring abstract readability" section. Specifically, they reported that high-impact articles tend to have a higher degree of lexical diversity (e.g., a higher type-token ratio), lexical sophistication (e.g., longer nouns), as well as syntactic complexity (e.g., longer sentences) than medium- and low-impact articles.

Online attention and social impact

In recent years, the rise of social media platforms such as Facebook, Google+, and Twitter and the increasing use of such platforms for sharing and discussing research outputs have dramatically improved the visibility of research outputs to the general public (Adie & Roe, 2013; Sugimoto et al., 2017). Conventional measures of scientific impact such as citation counts cannot readily capture the social impact of research outputs among the general public (Mohammadi et al., 2015). For this reason, bibliometrics researchers have started developing measures of social attention received by RAs online (Díaz-Faes et al., 2019). Altmetrics, short for "alternative metrics", is a term coined to describe web-based metrics of the impact of research outputs using data from online platforms such as social media (Bornmann, 2014). Such metrics are revolutionary in that, with the potential to take into account attention to and activities around research outputs from both research scholars and non-expert readers, they allow for an efficient and comprehensive assessment of the social impact of research outputs. It has been shown that altmetrics can serve as an effective indicator of attention to research outputs in a short time window, as it may take a long time for citations in other publications to appear and be indexed (Thelwall, 2018).

Citation counts and altmetrics usefully complement each other in assessing the scientific and social impact of research outputs, and some research has explored how these two mechanisms may be related to each other. Syamili and Rekha (2017) collected 62 RAs on the Ebola disease and compared their citation counts with several altmetrics variables, such as the number of views, saves, Mendeley shares, and Twitter shares. While they reported significant correlations between citation counts and views, saves, and Mendeley shares, they found that Twitter shares, which best captured attention from the general public, showed a very weak correlation with citation counts. Costas et al. (2015) reported only a weak correlation between the scientific impact and online attention of RAs, and further showed that 55% of the most highly-cited publications in their database received little attention from the general public. Similarly, Verma and Madhusudhan (2019) offered empirical evidence that highly-cited RAs are not necessarily widely disseminated on social media, while RAs receiving substantial online attention from non-expert readers may have no citation counts at all. All in all, these results support the notion that scientific impact measured using citation counts and social impact measured using altmetrics of online attention are indeed different aspects of RA impact that both deserve research attention.

While the title and abstract of an RA may affect whether researchers retrieve, read, and subsequently cite the RA, they may have a more immediate effect on whether non-expert readers pay attention to and share it online, and the readability of the title and abstract may play an even more important role in this process as non-expert readers generally have less expertise in the subject matter of the RA than expert researchers. While many studies have investigated the effect of abstract readability on the scientific impact of RAs, few studies have examined its effect on the online attention received by RAs. In a study of the virality of RAs, Guerini et al. (2012) examined the relationship between abstract readability, measured using the Gunning fog and Flesch indices, and three metrics of responses to RAs, namely, citations, bookmarks, and downloads. They reported that highly and lowly cited RAs did not differ in abstract readability, and that most bookmarked RAs tended to have less readable abstracts, while the most downloaded RAs tended to have more readable abstracts. The altmetrics of online attention (i.e., bookmarks and downloads) used in their study likely still reflected attention more from the scientific community rather than non-expert readers, many of whom may simply browse through RAs or share them on social media rather than actually bookmark or download them. In a recent study of the relationship between abstract readability and online attention, Ngai and Singh (2020) paid special attention to non-expert readers' interest in RAs. Specifically, they employed altmetrics data provided by Altmetric.com to measure public attention to RAs on new media and social media and reported a positive correlation between the use of metadiscourse and altmetrics attention score. However, the study focused on a single aspect of abstract readability only, i.e., the facilitative power exerted by metadiscoursal devices to persuade and engage readers, rather than measuring it with multiple indices of formal linguistic features. A better understanding of the relationship between abstract readability and online attention would call for the use of both multidimensional readability models that can comprehensively predict linguistic complexity and integrative altmetrics measurement that can better reflect the wide range and types of ways through which non-expert readers pay attention to and interact with RAs.

Research questions

The goal of this study is to examine the relationship between abstract readability and the online attention received by RAs. Specifically, the study aims to answer the following two research questions:

(1) What is the relationship between lexical and syntactic features of abstract readability and the level of online attention received by RAs?
(2) How is the relationship affected by publication time and discipline?

Method

The corpus

Our data consisted of a corpus of abstracts of articles published in the Reports column of the *Science* journal. The journal *Science* was selected for two reasons. First, it is ranked among the world's most high-impact outlets of cutting-edge research, has broad disciplinary coverage (with 43 distinct disciplines represented), and enjoys a wide international authorship and readership. Second, as indicated in its mission and scope statement, *Science* seeks to publish influential articles that merit recognition not only by the scientific community but also by the general public (*Science*, 2020a), making it especially appropriate for the analytical focus of the current study on online attention to RAs from both the scientific community and the general public. Articles in the Reports column "present important new research results of broad significance" (*Science*, 2020b), are about 2500 words in length, and are required to contain a standalone abstract of around 125 words that "explain[s] to the general reader why the research was done, what was found and why the results are important" (*Science*, 2020c).

Data collection proceeded in three steps. We first downloaded all articles published in the Reports column of the *Science* journal in 2012 and 2018 from its official website (https://science.sciencemag.org/content/collections.) in May, 2019. These two years were selected to examine the potential effect of time on the amount of online attention received by the articles. Next, we filtered disciplines with fewer than 25 articles in either year. A total of 11 disciplines remained, including Biochemistry, Cell Biology, Chemistry, Ecology, Evolution, Geophysics, Immunology, Materials Science, Molecular Biology, Neuroscience, and Physics. Finally, we randomly sampled 25 articles for each of the 11 disciplines for each of the two years. As summarized in Table 1, the final corpus contained 550 abstracts, with a total of 68,859 running words (Mean = 125.20, SD = 13.94).

Readability measurement

Following recent research that advocated for and empirically validated the use of indices reflecting different dimensions of linguistic complexity for readability measurement (e.g., Crossley et al., 2019; Graesser et al., 2011; Lu et al., 2019; Tankó, 2017; Vajjala & Meurers, 2012), we assessed the readability of each abstract in the corpus using a comprehensive set of lexical and syntactic complexity indices.

Table 1 Descriptive statistics of the corpus

Discipline	Year of publication	Number of texts	Words	Mean	SD
Biochemistry	2012	25	3275	131.00	11.04
	2018	25	3119	124.76	8.13
Cell Biology	2012	25	3137	125.48	11.21
	2018	25	3267	130.68	7.55
Chemistry	2012	25	3084	123.36	17.90
	2018	25	3053	122.12	14.35
Ecology	2012	25	3142	125.68	20.20
	2018	25	3265	130.60	17.73
Evolution	2012	25	3132	125.28	13.25
	2018	25	3105	124.20	10.11
Geophysics	2012	25	3089	123.56	11.97
	2018	25	3247	129.88	20.78
Immunology	2012	25	3076	123.04	12.62
	2018	25	3414	136.56	12.52
Materials Science	2012	25	2851	114.04	14.24
	2018	25	3080	123.20	8.43
Molecular Biology	2012	25	3094	123.76	9.30
	2018	25	3247	129.88	12.94
Neuroscience	2012	25	3010	120.40	13.88
	2018	25	3144	125.76	10.78
Physics	2012	25	2951	118.04	14.76
	2018	25	3077	123.08	9.33
All		550	68,859	125.20	13.94

The Lexical Complexity Analyzer (LCA; Lu, 2012) was employed to measure the lexical complexity of each abstract. LCA contains 25 measures that gauge three dimensions of lexical complexity, namely, lexical density, lexical sophistication, and lexical variation, as summarized in Table 2. Lexical density is calculated as the ratio of the number of lexical words to the total number of words in a text. Lexical sophistication refers to the degree to which the words used in a text are advanced, and, following Laufer and Nation (1995), LCA operationalizes sophisticated words as low frequency words (i.e., words beyond the top 2000 most frequent words in the British National Corpus). This operationalization is both practical and useful, as more frequent words have been found to be processed or decoded faster in reading (Haberlandt & Graesser, 1985; Perfetti, 2007). LCA offers five lexical sophistication measures that each tap into the proportion of sophisticated lexical word tokens (LS1), sophisticated word types (LS2), and sophisticated verb types (VS1, CVS1, and VS2). Lexical variation assesses the degree of diversity of the words used in a text. Texts with greater lexical variation may create more gaps in cohesion and lead to greater difficulty in reading comprehension than texts with greater lexical repetition (Graesser et al., 2011). LCA offers 19 lexical variation measures of three broad categories. The first four are based on the idea of number of different words (NDW). In addition to NDW in the text, three additional measures are included to mitigate the effect of text length on NDW, i.e. NDW in the

Table 2 Lexical complexity measures from Lexical Complexity Analyzer (Lu, 2012)

Number	Measure	Code	Formula
Lexical density			
1	Lexical density	LD	N_{lex}/N
Lexical sophistication			
2	Lexical sophistication-I	LS1	N_{slex}/N_{lex}
3	Lexical sophistication-II	LS2	T_s/T
4	Verb sophistication-I	VS1	T_{sverb}/N_{verb}
5	Corrected VS1	CVS1	$T_{sverb}/\sqrt{2N_{verb}}$
6	Verb sophistication-II	VS2	T^2_{sverb}/N_{verb}
Lexical variation			
7	Number of different words	NDW	T
8	NDW (first 50 words)	NDW-50	T in the first 50 words of sample
9	NDW (expected random 50)	NDW-ER50	Mean T of 10 random 50-word samples
10	NDW (expected sequence 50)	NDW-ES50	Mean T of 10 random 50-word sequences
11	Type-token ratio	TTR	T/N
12	Mean Segmental TTR (50)	MSTTR-50	Mean TTR of all 50-word segments
13	Corrected TTR	CTTR	$T/\sqrt{2N}$
14	Root TTR	RTTR	$T/\sqrt{N}$
15	Bilogarithmic TTR	LogTTR	$LogT/LogN$
16	Uber Index	Uber	$Log^2N/Log(N/T)$
17	Lexical word variation	LV	T_{lex}/N_{lex}
18	Verb variation-I	VV1	T_{verb}/N_{verb}
19	Squared VV1	SVV1	T^2_{verb}/N_{verb}
20	Corrected VV1	CVV1	$T_{verb}/\sqrt{2N_{verb}}$
21	Verb variation-II	VV2	T_{verb}/N_{lex}
22	Noun variation	NV	T_{noun}/N_{lex}
23	Adjective variation	AdjV	T_{adj}/N_{lex}
24	Adverb variation	AdvV	T_{adv}/N_{lex}
25	Modifier variation	ModV	$T_{adj} + T_{adv}/N_{lex}$

T = Number of word types; N = Number of word tokens. The subscript under T or N denotes the number of types of tokens of particular category of words: lex = lexical words; slex = sophisticated lexical words; s = sophisticated words; sverb = sophisticated verbs; adj = adjectives; adv = adverbs

first 50 words of the text (NDW-50) and the average NDW of 10 random 50-word samples from the text (NDW-ER50) or 10 random 50-word sequences from the text (NDW-ES50). The next six are based on the idea of type-token ratio (TTR), including both the original TTR measure and five transformations of the measure to mitigate the effect of text length (MSTTR-50, CTTR, RTTR, LogTTR, and Uber). The final nine measures gauge the degree of variation of specific types of words, including lexical word variation (LV), verb variation (VV1, SVV1, CVV1, and VV2), noun variation (NV), adjective variation (AdjV), adverb variation (AdvV), and modifier variation (ModV).

The L2 Syntactic Complexity Analyzer (L2SCA; Lu, 2010) was used to assess the syntactic complexity of each abstract. Based on Wolfe-Quintero et al. (1998)'s recommendation and categorization, L2SCA offers 14 measures that reflect five dimensions of syntactic complexity, including length of production unit, amount of subordination, amount of

Table 3 Syntactic complexity measures from L2 Syntactic Complexity Analyzer (Lu, 2010)

Number	Label	Description
Length of production unit		
1	MLC	Mean length of clause
2	MLS	Mean length of sentence
3	MLT	Mean length of T-unit
Amount of subordination		
4	C/T	Number of clauses per T-unit
5	CT/T	Complex T-unit ratio
6	DC/C	Number of dependent clauses per clause
7	DC/T	Number of dependent clauses per T-unit
Amount of coordination		
8	CP/C	Number of coordinate phrases per clause
9	CP/T	Number of coordinate phrases per T-unit
10	T/S	Number of T-units per sentence
Degree of phrasal sophistication		
11	CN/C	Number of complex nominals per clause
12	CN/T	Number of complex nominals per T-unit
13	VP/T	Number of verb phrases per T-unit
Overall sentence complexity		
14	C/S	Number of clauses per sentence

coordination, degree of phrasal complexity, and overall sentence complexity, as summarized in Table 3.

The LCA and L2SCA are both freely accessible through command-line and web-based interfaces, as detailed in Lu (2014); additionally, the L2SCA is also accessible through a graphic-user interface and as a R package.[1] The reliability of LCA depends on the reliability of the tagger and lemmatizer used for POS tagging and lemmatization. For example, the web-based LCA uses the Stanford tagger (Toutanova et al., 2003) and the MORPHA lemmatizer (Minnen et al., 2001), which have accuracy levels of over 97% and 99%, respectively. Lu (2010) evaluated the reliability of the L2SCA on a set of English essays written by Chinese college students and reported correlations ranging from 0.834 to 1.000 between the scores calculated by human annotators and those generated by the L2SCA. In terms of their usefulness for text readability analysis, Vajjala and Meurers (2012) reported that using 15 of the 25 measures from the LCA measures (including the LD measure and all lexical variation measures other than the four NDW-based ones) along with an additional measure of lexical variation achieved an accuracy of 68.1% for predicting the grade levels of reading texts, while using the 14 measures from the L2SCA achieved an accuracy of 71.2%.

[1] http://www.personal.psu.edu/xxl13/downloads.

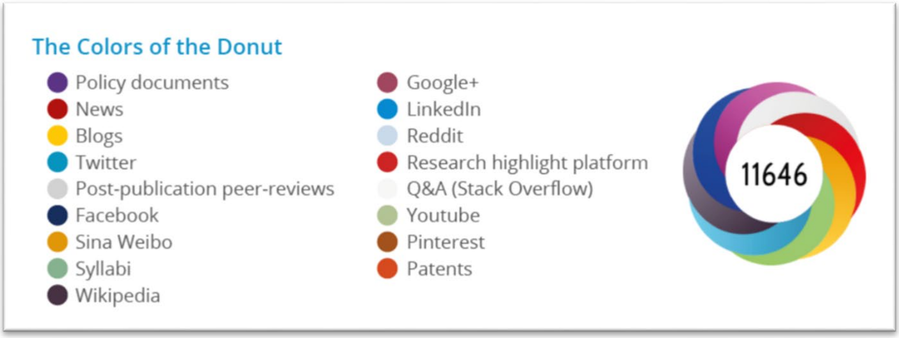

Fig. 1 Diagrammatic representation of the Donut (Altmetric, 2020)

Attention measurement

The online attention received by each RA in our dataset was quantitatively assessed using the *Altmetric* attention score (AAS, provided by Altmetric.com) for each RA retrieved from the official website of *Science* on May 14, 2019. The AAS of an RA has been considered a reliable and convenient reflection of the quantity and quality of the online attention it receives, as it weights the mentions of the article by different authors from different types of sources (Adie & Roe, 2013). In particular, three variables play an important role in computing the AAS of an RA, namely, the volume, authors, and sources of the online mentions of the RA (Trueger et al., 2015). The AAS of an RA goes up with the increase of the volume of its online mentions, i.e. the number of times it is shared or discussed. Meanwhile, the contribution of an author's mention of the article to its AAS is weighted by the author's profile, including reach (number of followers), promiscuity (frequency of mentioning research outputs), and bias (frequency of mentioning research outputs from the same journal) (Trueger et al., 2015). Finally, the AAS of an RA is presented in a colored "donut" with different colors representing different types of sources of its mentions, such as news outlets, policy documents, Wikipedia, and social media (Brigham, 2014). This is illustrated by Fig. 1, in which the number at the center of the donut is the AAS received by a certain article, and the various colors represent the proportions of its mentions from different sources. Table 4 shows the descriptive statistics of the AASs for all RAs in our corpus by discipline and publication time.

Statistical analysis

We first conducted a two-way ANOVA with publication time and discipline as independent variables and the AASs of the *Science* articles as dependent variable to determine the effects of the two control variables and their interaction on the online attention received by the articles. To account for the considerable variation across years of publication and disciplines, the raw AASs were normalized using the field-normalization approach (e.g., Huang et al., 2018; Lei & Yan, 2016; Lu et al., 2019). Specifically, we calculated normalized AAS (NAAS) using the following formula:

NAAS = article's AAS/mean AAS of the year and discipline in which the article is published

We then calculated the correlation coefficients between each of the 39 linguistic complexity indices and the NAASs and also examined the correlations between different

Table 4 Descriptive statistics of AASs by discipline and publication time

Discipline	Year of publication	Number of texts	Minimum	Maximum	Mean	SD
Biochemistry	2012	25	6.30	74.79	30.08	19.79
	2018	25	9.10	920.34	163.92	218.58
Cell Biology	2012	25	4.75	76.02	20.70	16.58
	2018	25	12.75	722.79	132.47	141.16
Chemistry	2012	25	0.25	111.47	38.99	35.24
	2018	25	1.00	378.37	91.25	82.83
Ecology	2012	25	5.45	560.45	112.86	141.84
	2018	25	54.75	1708.36	711.33	463.22
Evolution	2012	25	2.50	449.55	82.23	90.99
	2018	25	119.00	1368.53	387.73	318.81
Geophysics	2012	25	0.25	222.31	41.86	58.30
	2018	25	24.70	1978.78	325.94	407.86
Immunology	2012	25	1.75	377.32	54.49	87.70
	2018	25	19.20	1046.77	195.98	253.93
Materials Science	2012	25	2.50	333.02	50.85	67.63
	2018	25	1.00	380.32	130.48	94.53
Molecular Biology	2012	25	3.35	176.39	24.81	34.62
	2018	25	27.05	1026.42	261.86	298.67
Neuroscience	2012	25	5.05	357.53	62.80	70.76
	2018	25	12.15	534.46	208.90	131.58
Physics	2012	25	1.50	222.56	30.39	42.88
	2018	25	2.25	182.03	78.02	50.03

publication times and among different disciplines. Finally, we constructed a series of multiple regression models with the NAASs as the dependent variable and the linguistic complexity indices as independent variables, using the stepwise method to exclude insignificant predictors (Larson-Hall, 2015).

Results

The two-way ANOVA showed significant effects of publication time (F (1, 528) = 143.98, $p < 0.001$, $\eta_p^2 = 0.214$), discipline (F (10, 528) = 14.68, $p < 0.001$, $\eta_p^2 = 0.218$), as well as the interaction between the two (F (10, 528) = 8.87, $p < 0.001$, $\eta_p^2 = 0.144$) on the AASs of the articles. Tukey's HSD post hoc tests showed that the AASs of the articles published in 2018 (Mean = 244.35) were significantly higher than those published in 2012 (Mean = 50) ($p < 0.001$), and the AASs of the articles in Ecology (Mean = 412.09) were significantly higher than those in all other disciplines ($p < 0.001$). Therefore, to exclude the effect of these two variables, we used NAASs in our subsequent analysis.

Table 5 presents the correlation coefficients between the NAASs of the articles and the 39 linguistic complexity indices computed using the entire dataset. The results showed that overall, three lexical complexity measures (i.e., LS2, NDW-50, and AdvV) significantly correlated with NAASs.

Table 5 Correlations between *NAASs* and linguistic complexity indices ($N = 550$)

Linguistic complexity indices	*Normalized Altmetric* attention scores	Significance
LD	−0.048	0.264
LS1	−0.076	0.076
LS2	−0.118**	0.005
VS1	−0.032	0.451
VS2	−0.024	0.581
CVSL	−0.032	0.454
NDW	0.028	0.513
NDW-50	0.088*	0.04
NDW-ER50	0.042	0.322
NDW-ES50	0.06	0.157
TTR	0.062	0.146
MSTTR	0.055	0.201
CTTR	0.047	0.267
RTTR	0.048	0.26
LOGTTR	0.052	0.222
UBER	0.071	0.098
LV	0.046	0.28
VVL	−0.035	0.418
SVVL	−0.02	0.632
CVVL	−0.026	0.543
VV2	0.011	0.8
NV	0.057	0.185
AdjV	0.024	0.567
AdvV	0.124**	0.003
ModV	0.068	0.111
MLS	0.045	0.294
MLT	0.047	0.276
MLC	−0.006	0.884
C/S	0.047	0.267
VP/T	−0.014	0.75
C/T	0.059	0.165
DC/C	0.028	0.505
DCT	0.053	0.217
T/S	−0.009	0.832
CT/T	0.02	0.633
CP/T	0.039	0.361
CP/C	0.024	0.58
CN/T	0.079	0.063
CN/C	0.029	0.496

$^{**}p < 0.01$, $^{*}p < 0.05$

We further examined the correlation coefficients between the linguistic complexity indices and NAASs for articles in each of the 11 disciplines published in each of the two publication times ($N = 25$ for each of the 22 groups)—the size of the correlation matrix (25

Table 6 The proportions of variance explained by the regression models

Model	R	R^2	Adjusted R^2	Standard error of the estimate	R^2 change	F change	df1	df2	Sig. F change
1	0.124	0.015	0.014	1.07716	0.015	8.616	1	548	0.003
2	0.16	0.026	0.022	1.07265	0.01	5.626	1	547	0.018
3	0.188	0.035	0.03	1.06822	0.01	5.539	1	546	0.019
4	0.211	0.044	0.037	1.06417	0.009	5.17	1	545	0.023
5	0.229	0.053	0.044	1.06055	0.008	4.722	1	544	0.03

Model 1 predictors: (constant), AdvV; Model 2 predictors: (constant), AdvV, LS2; Model 3 predictors: (constant), AdvV, LS2, CN/T; Model 4 predictors: (constant), AdvV, LS2, CN/T, NDW-50; Model 5 predictors: (constant), AdvV, LS2, CN/T, NDW-50, VP/T

Table 7 Coefficients of the regression models

Model		Unstandardized coefficients		Standardized coefficients	t	Sig	Collinearity statistics	
		B	Standard error	Beta			Tolerance	VIF
1	(Constant)	0.791	0.085		9.348	0		
	AdvV	5.535	1.886	0.124	2.935	0.003	1	1
2	(Constant)	1.465	0.296		4.945	0		
	AdvV	4.832	1.901	0.109	2.542	0.011	0.976	1.025
	LS2	−1.628	0.687	−0.101	−2.372	0.018	0.976	1.025
3	(Constant)	1.089	0.336		3.245	0.001		
	AdvV	5.133	1.898	0.115	2.705	0.007	0.971	1.03
	LS2	−1.77	0.686	−0.11	−2.578	0.01	0.968	1.033
	CN/T	0.113	0.048	0.1	2.353	0.019	0.986	1.014
4	(Constant)	−0.309	0.7		−0.441	0.659		
	AdvV	4.724	1.899	0.106	2.488	0.013	0.963	1.039
	LS2	−2	0.691	−0.125	−2.894	0.004	0.947	1.056
	CN/T	0.114	0.048	0.1	2.373	0.018	0.986	1.015
	NDW-50	0.037	0.016	0.097	2.274	0.023	0.973	1.027
5	(Constant)	0.007	0.712		0.01	0.992		
	AdvV	5.268	1.909	0.118	2.76	0.006	0.946	1.057
	LS2	−2.264	0.699	−0.141	−3.237	0.001	0.919	1.088
	CN/T	0.158	0.052	0.139	3.039	0.002	0.837	1.195
	NDW-50	0.038	0.016	0.1	2.371	0.018	0.972	1.029
	VP/T	−0.224	0.103	−0.1	−2.173	0.03	0.823	1.215

indices, 2 publication times, and 11 disciplines) makes it impractical to present it here. Many more significant correlation coefficients emerged for different groups. Notably, many indices exhibited substantial differences in their correlation coefficients with NAASs across these groups. For example, the correlation coefficient between MLS and NAASs was −0.448 ($p < 0.05$) for Physics articles published in 2012, 0.174 ($p = 0.405$) for Physics articles published in 2018, and 0.045 ($p = 0.294$) for all articles in the dataset.

Given these results, we entered all 39 linguistic complexity indices as predictors of NAASs in constructing multiple regression models, using the stepwise method to exclude insignificant predictors. Tables 6 and 7 summarize the proportion of variance of NAASs explained by each regression model and the coefficients of each regression model, respectively. As these results indicated, the best model (Model 5) included five significant predictors (AdvV, LS2, CN/T, NDW-50, and VP/T) and accounted for 5.3% of the total variance of NAASs. The Tolerance (> 0.10) and VIF (< 10) values of these predictors indicated that multicollinearity was not a concern.

Discussion

Our analysis of the relationship between abstract readability, measured using 39 lexical and syntactic complexity indices, and the online attention received by 550 RAs from 11 disciplines published in *Science* in 2012 and 2018, measured using their AASs, revealed a number of substantive findings. First, our analysis confirmed significant effects the two control variables, publication time and discipline, on the online attention received by the RAs. Second, while only three complexity indices correlated significantly with the NAASs of the RAs in the whole dataset, the correlation coefficients between these indices and the NAASs were found to vary substantially by publication time and discipline. Finally, a multiple regression analysis using the 39 complexity indices as predictors generated a model that accounted for 5.3% of the total variance of the NAASs, with five significant predictors. We discuss these results in relation to the two research questions and their implications below.

The relationship between abstract readability and the online attention received by RAs

The results pertaining to our first research question showed that abstract readability significantly affects the online attention received by RAs. The five indices that made their way into the best regression model included one lexical sophistication index (LS2), two lexical variation indices (AdvV and NDW-50), and two phrasal complexity indices (CN/T and VP/T). A few previous studies have examined the relationship between abstract readability and the scientific impact of RAs (e.g., Didegah & Thelwall, 2013; Lei & Yan, 2016). As noted earlier, online attention differs from scientific impact, which is usually measured using citation counts, in that the major contributors of online attention are non-expert readers rather than researchers and the two groups of readers may be sensitive to the readability and complexity of RA abstracts in different ways.

Lexical sophistication negatively affected the online attention received by RAs, as indicated by the significant negative coefficient of LS2. This result suggests that non-expert readers' comprehension of and attention to RAs may be negatively impacted by the proportion of sophisticated word types, many of which are likely academic and technical words, in RA abstracts. Lu et al. (2019) reported that high-impact RAs used more sophisticated words, suggesting that the use of more sophisticated academic or technical words does not necessarily present a greater comprehension challenge to researchers, but may in fact contribute to a more positive perception of the research and/or writing quality of the RA, making researchers more likely to read the RAs carefully and cite them (e.g., Dolnicar & Chapple, 2015). In other words, the different effects of lexical sophistication on online attention and scientific impact are consistent with the differential level of academic and technical expertise between non-expert readers and expert researchers. The contrast between our results and those reported by Lu et al. (2019) highlights the difference between researchers' and non-expert readers' perceptions of RA readability (Newbold & Gillam, 2010), considering that scholarly references are the only source of citations while media outlets (e.g., online news and blogs) are weighted heavily in the calculation of AASs (Altmetric, 2021).

Lexical variation had the greatest effect on online attention. Two lexical variation measures (i.e., AdvV and NDW-50) were positive predictors of NAASs, accounting for nearly half of the total variance explained by all indices in the final regression model. These results suggest that a greater number of different words in the abstract positively impacted online attention. Particularly, a higher degree of adverb variation was useful in attracting more online attention. Although greater lexical variation may create a larger gap in text cohesion and greater processing burden for readers (Graesser et al., 2011), our results show that the richer content encoded by more diverse words may outweigh the processing challenge they pose in sustaining readers' attention and interest. Lu et al. (2019) reported that high-impact RAs had a higher degree of lexical diversity measured using TTR, suggesting similar attractiveness of lexically diverse RAs to researchers. Taken together, these results indicate that greater lexical variation may contribute to both researchers and non-expert readers' positive perception of the writing quality of RAs, and subsequently increase their likelihood to read and share the RAs.

Two syntactic complexity indices, i.e. CN/T and VP/T, were included in the final regression model. CN/T was a positive predictor of NAASs, while VP/T was a negative predictor. Complex nominals are commonly used in RAs to achieve informativeness and concision at the same time through the use of modifiers (Lu et al., 2020; Snow, 2010). A higher density of complex nominals in RA abstracts not only makes the abstracts more academic and formal but also allows the authors to pack more information in the limited space allowed. Our results show that these features may have positively affected non-expert readers' attention to RAs and their positive effect was not outweighed by the additional structural complexity brought about by the greater use of complex nominals. On the contrary, greater use of verbal phrases negatively affected online attention, as indicated by the significant negative coefficient of VP/T. This may be due to non-expert readers' reliance on linguistic style as a criterion of RA quality. Complex nominals have been shown to be a highly important feature of academic writing (e.g., Lu et al., 2020), and many complex nominals in RAs may have arisen from nominalizations, which are recognized to be "crucial to the conciseness expected in academic language" (Snow, 2010, p.452). It is possible that non-expert readers may judge RAs with the concise and authoritative style realized through complex nominals in general and nominalizations in particular more positively than those employing more clausal structures (both finite and non-finite).

The impact of publication time and discipline

The AASs of the RAs varied significantly by publication time, and the correlations between the complexity indices and AASs varied substantially by publication time as well. We stretched the time span from the first year AASs became available (2012) to the year before our data collection (2018) to maximize the potential effect of time on cumulative online attention. Time generally positively affects citation counts, and we expected a similar effect of time on cumulative online attention. The result that RAs published in 2018 had a higher mean AAS than those published in 2012 was thus initially surprising. Several reasons may help explain this difference. First, while it takes time for RAs to be cited and the scientific impact of RAs tend to grow over time, online shares and mentions of RAs happen instantaneously and tend to concentrate on recent RAs. Second, the past decade witnessed a rapid expansion of the use of the Internet for knowledge discovery and dissemination as well as the use of social media in general, likely leading to increased online reading and sharing by a wider non-expert reader base. The increase in online reading by a wider non-expert reader base likely contributed to greater online attention of RAs published in 2018 than those published in 2012. Meanwhile, as more readers with more diverse levels of expertise participate in online reading and sharing, it is not surprising that the relationship between abstract readability and online attention shifts accordingly.

In addition, the AASs of the RAs and the relationship between abstract readability and online attention varied by discipline as well. The mean AAS of Ecology RAs reached 412.09, significantly higher than the means of RAs in all other disciplines, and noticeably higher than even the second highest mean AAS, which was 234.98 for Evolution RAs. Physics RAs had the lowest mean AAS, which was 54.20. Among the top 100 RAs with the highest AASs in our corpus, more than a quarter (26) were from Ecology, while only one was from Physics. Previous studies also reported substantial between-discipline variation in AASs and Mendeley readership (Htoo & Na, 2017; Zahedi et al., 2014). The disciplinary variation in AASs may to some extent reflect the interest of the general public in research in different disciplines. Non-expert readers' differential levels of interest and expertise in research in different disciplines likely also affected the role of abstract readability in the online attention received by RAs from different disciplines.

Implications

In an era of mass online reading and sharing facilitated by digitalization and social media platforms, the impact of research outputs is increasingly extending from within the research community to outside the academia. The differences in disciplinary knowledge expertise and reading motives between researchers and non-expert readers will result in differences in their comprehension of and interest in research outputs (Coiro, 2021). As the part of research outputs read most frequently by online readers (Nicholas et al., 2003), abstracts play an especially important role in expanding the reach of RAs among online readers, given the vast information available online that competes for their attention and their usually short attention span for any particular information source. Our results highlight the importance to make RA abstracts more readable and accessible to non-expert readers in order to increase the impact of research outputs among the general public and facilitate knowledge dissemination in the digital era. The results from our regression analysis also offer several specific lexical and syntactic complexity features that authors can pay attention to in an effort to make RA abstracts less challenging to and more appealing to non-expert readers.

Our results also have useful methodological implications for further research in this area. First, while traditional readability formulas merge lexical and syntactic difficulty into a single metric, our results indicate that it is useful to differentiate lexical and syntactic complexity measures in assessing abstract readability both to better understand their respective effect on online attention and to pinpoint the specific features that researchers can focus on in their efforts to improve abstract readability and increase online attention. Second, while traditional readability formulas usually use mean length of sentence as the measure of syntactic difficulty, our analysis using 14 indices reflecting multiple different dimensions of syntactic complexity indicated the importance of considering phrasal complexity in this line of research. Third, our results also showed that publication time and discipline should both be taken into account in future research on online attention. Discipline-specific analysis of the relationship between abstract readability and online attention of RAs will provide more relevant insight for researchers in particular disciplines.

Conclusion

This study examined the relationship between abstract readability, measured using 39 indices of lexical and syntactic complexity, and the online attention received by RAs, measured using *Altmetric* attention scores, with a corpus of 550 RAs from 11 disciplines published in 2012 and 2018 in the Reports column of *Science*. Our analysis revealed a number of lexical sophistication, lexical variation, and phrasal complexity measures that significantly predicted the amount of online attention received by the RAs. The analysis further showed that both publication time and discipline may affect the online attention received by the RAs as well as the relationship of abstract readability to online attention. Our findings provide empirical evidence for the importance of making RA abstracts accessible to non-expert readers in order to increase the impact of research outputs among the general public. Our findings also offer several specific linguistic complexity features that researchers can tap into to increase the readability of RA abstracts and maximize reader comprehension and interest.

Several limitations of the current study and avenues for future research exist. First, the scale and scope of our dataset can be expanded to include a greater number of RAs from more diverse journals, disciplines, and publication periods. Second, emerging initiatives to make research output accessible, such as the OASIS initiative, are highly plausible, and it will be revealing to assess non-expert readers' perception of the readability differences between original RA abstracts and the accessible summaries as well as the effect of those differences on their understanding of and interest in the research outputs and their likelihood to share the research outputs. Third, in addition to lexical and syntactic features, future research may include additional readability indices, such as those of cohesion and sentiment analysis (e.g., Crossley et al., 2019), to examine abstract readability in a more comprehensive way and better understand its effect on online attention. Finally, abstract readability is only one of many factors that can influence non-expert readers' interest in and attention to research outputs, and it will be useful to examine the interaction between abstract readability and such factors as research topic, title readability, number and complexity of keywords, researcher or institution prestige, and journal impact, among others.

Authors' contributions TJ: Conceptualization, Methodology, Writing—review & editing. HD: Data curation, Investigation, Writing—review & editing. XL: Conceptualization, Methodology, Writing—review & editing. JN: Methodology, Investigation, Writing—review & editing. KG: Methodology, Investigation, Writing—review & editing.

Funding This research was supported by a grant from the National Social Science Fund of China (18BYY110) to the first author.

References

Adie, E., & Roe, W. (2013). Altmetric: Enriching scholarly content with article-level discussion and metrics. *Learned Publishing, 26*(1), 11–17.

Altmetric Website. (2020). Colors of the donut. Retrieved on 30 August 2020 from https://www.altmetric.com/about-our-data/the-donut-and-score/
Altmetric Website. (2021). How is the Altmetric Attention Score calculated?. Retrieved on 25 April 2021 from https://help.altmetric.com/support/solutions/articles/6000233311-how-is-the-altmetric-attention-score-calculated-
Benjamin, R. G. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review, 24*(1), 63–88.
Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics, 8*(4), 895–903.
Brigham, T. J. (2014). An introduction to Altmetrics. *Medical Reference Services Quarterly, 33*(4), 438–447.
Chen, B., Deng, D., Zhong, Z., & Zhang, C. (2020). Exploring linguistic characteristics of highly browsed and downloaded academic articles. *Scientometrics, 122*(3), 1769–1790.
Coiro, J. (2021). Toward a multifaceted heuristic of digital reading to inform assessment, research, practice, and policy. *Reading Research Quarterly*, *56*(1), 9–31.
Costas, R., Zahedi, Z., & Wouters, P. (2015). Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology, 66*(10), 2003–2019.
Crossley, S. A., Skalicky, S., & Dascalu, M. (2019). Moving beyond classic readability formulas: New methods and new models. *Journal of Research in Reading, 42*(3–4), 541–561.
D'Angelo, C. A., & Di Russo, S. (2019). Testing for universality of Mendeley readership distributions. *Journal of Informetrics, 13*(2), 726–737.
Díaz-Faes, A. A., Bowman, T. D., & Costas, R. (2019). Towards a second generation of 'social media metrics': Characterizing Twitter communities of attention around science. *PLoS ONE, 14*(5), 1–18.
Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics, 7*(4), 861–873.
Dolnicar, S., & Chapple, A. (2015). The readability of articles in tourism journals. *Annals of Tourism Research, 52*, 161–166.
Dronberger, G. B., & Kowitz, G. T. (1975). Abstract readability as a factor in information systems. *Journal of the American Society for Information Science, 26*(2), 108–111.
Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221–233.
Gazni, A. (2011). Are the abstracts of high impact articles more readable? Investigating the evidence from top research institutions in the world. *Journal of Information Science, 37*(3), 273–281.
Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher, 40*(5), 223–234.
Guerini, M., Pepe, A., & Lepri, B. (2012, June). Do linguistic style and readability of scientific abstracts affect their virality? In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media* (pp. 475–478). Dublin, Ireland.
Haberlandt, K. F., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General, 114*(3), 357–374.
Hartley, J. (2000). Clarifying the abstracts of systematic literature reviews. *Bulletin of the Medical Library Association, 88*(4), 332–337.
Hartley, J., & Sydes, M. (1997). Are structured abstracts easier to read than traditional ones? *Journal of Research in Reading, 20*(2), 122–136.
Houghton, J. W., Henty, M., & Steele, C. (2004). Research practices and scholarly communication in the digital environment. *Learned Publishing, 17*(3), 231–249.
Htoo, T. H. H., & Na, J. C. (2017). Disciplinary differences in altmetrics for social sciences. *Online Information Review, 41*(2), 235–251.
Huang, W., Wang, P., & Wu, Q. (2018). A correlation comparison between Altmetric Attention Scores and citations for six PLOS journals. *PLoS One*, *13*(4), e0194962.
Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.
King, R. (1976). A comparison of the readability of abstracts with their source documents. *Journal of the American Society for Information Science, 27*(2), 118–121.
Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363–394.
Klare, G. R. (1963). *The measurement of readability*. Iowa State University Press.
Klare, G. R. (1974). Assessing readability. *Reading Research Quarterly, 10*(1), 62–102.
Larson-Hall, J. (2015). *A guide to doing statistics in second language research using SPSS and R* (2nd ed.). Routledge.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307–322.

Lei, L., & Yan, S. (2016). Readability and citations in information science: Evidence from abstracts and articles of four journals (2003–2012). *Scientometrics, 108*(3), 1155–1169.

Lu, C., Bu, Y., Dong, X., Wang, J., Ding, Y., Larivière, V., Sugimoto, C. R., Paul, L., & Zhang, C. (2019). Analyzing linguistic complexity and scientific impact. *Journal of Informetrics, 13*(3), 817–829.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474–496.

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *Modern Language Journal, 96*(2), 190–208.

Lu, X. (2014). *Computational methods for corpus annotation and analysis*. Springer.

Lu, X., Casal, J. E., & Liu, Y. (2020). The rhetorical functions of syntactically complex sentences in social science research article introductions. *Journal of English for Academic Purposes*, *44*, 100832.

Lu, X., Gamson, D. A., & Eckert, S. A. (2014). Lexical difficulty and diversity in American elementary school reading textbooks: Changes over the past century. *International Journal of Corpus Linguistics, 19*(1), 94–117.

Maflahi, N., & Thelwall, M. (2018). How quickly do publications get read? The evolution of Mendeley reader counts for new articles. *Journal of the Association for Information Science and Technology, 69*(1), 158–167.

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading, 12*(8), 639–646.

Minnen, G., Carroll, J., & Pearce, D. (2001). Applied morphological processing of English. *Natural Language Engineering, 7*(3), 207–223.

Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (2015). Who reads research articles? An altmetrics analysis of Mendeley user categories. *Journal of the Association for Information Science and Technology, 66*(9), 1832–1846.

Newbold, N., & Gillam, L. (2010). The linguistics of readability: The next step for word processing. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics and Writing: Writing Processes and Authoring Aids* (pp. 65–72). Los Angeles, California: Association for Computational Linguistics.

Ngai, C. S. B., & Singh, R. G. (2020). Relationship between persuasive metadiscoursal devices in research article abstracts and their attention on social media. *PLoS One*, *15*(4), e0231305.

Nicholas, D., Huntington, P., & Watkinson, A. (2003). Digital journals, Big Deals and online searching behavior: A pilot study. *Aslib Proceedings: New Information Perspectives, 55*(1/2), 84–109.

OASIS Website. (2020). Why accessible summaries. Retrieved on 30 August 2020 from https://oasis-database.org/about

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383.

Pinto, M., & Lancaster, F. W. (1999). Abstracts and abstracting in knowledge discovery. *Library Trends*, *48*(1), 234–248.

Science Website. (2020a). Mission and scope. Retrieved on 30 August 2020 from https://www.sciencemag.org/about/mission-and-scope

Science Website. (2020b). Information for authors. Retrieved on 30 August 2020 from https://www.sciencemag.org/authors/science-information-authors

Science Website. (2020c). Instructions for preparing an initial manuscript. Retrieved on 30 August 2020 from https://www.sciencemag.org/authors/instructions-preparing-initial-manuscript

Sienkiewicz, J., & Altmann, E. G. (2016). Impact of lexical and sentiment factors on the popularity of scientific papers. *Royal Society Open Science*, *3*(6), 160140.

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science, 328*(5977), 450–453.

Stevens, R. J., Lu, X., Baker, D. P., Ray, M. N., Eckert, S. A., & Gamson, D. A. (2015). Assessing the cognitive demands of elementary school reading curricula: An analysis of reading text and comprehension tasks from 1910 to 2000. *American Educational Research Journal, 52*(3), 582–617.

Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics, 98*(2), 1131–1143.

Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology, 68*(9), 2037–2062.

Syamili, C., & Rekha, R. V. (2017). Do altmetric correlate with citation?: A study basedon PLOS ONE journal. *COLLNET Journal of Scientometrics and Information Management, 11*(1), 103–117.

Tankó, G. (2017). Literary research article abstracts: An analysis of rhetorical moves and their linguistic realizations. *Journal of English for Academic Purposes, 27*, 42–55.

Thelwall, M. (2018). Early Mendeley readers correlate with later citation counts. *Scientometrics, 115*(3), 1231–1240.

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PLoS ONE, 8*(5), 1–7.

Toutanova, K., Klein, D., Manning, C., & Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 252–259). Edmonton, Alberta, Canada: The Association for Computational Linguistics.

Trueger, N. S., Thoma, B., Hsu, C. H., Sullivan, D., Peters, L., & Lin, M. (2015). The altmetric score: A new measure for article-level dissemination and impact. *Annals of Emergency Medicine, 66*(5), 549–553.

Vajjala, S., & Meurers, D. (2012, June). On improving the accuracy of readability classification using insights from second language acquisition. In *Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 163–173). Montréal, Canada: Association for Computational Linguistics.

Verma, S., & Madhusudhan, M. (2019). An altmetric comparison of highly cited digital library publications of India and China. *Annals of Library and Information Studies, 66*(2), 71–75.

Weil, B. H. (1970). Standards for writing abstracts. *Journal of the American Society for Information Science, 21*(5), 351–357.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. University of Hawaii Press.

Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics, 101*(2), 1491–1513.