'FORM' AND 'FUNCTION' IN THE SOCIAL SCIENCES

Byoung-Hyoun Hwang, Hugh Hoikwang Kim and Kai Wu*

This Draft: August 2018

Ideas proposed in scientific journal articles have the potential to touch many areas of our lives in significant ways. Here, we examine whether there are frictions in the scientific knowledge dissemination process. In particular, we test whether something as seemingly frivolous as the ease with which an article can be read affects the degree to which an article reaches its intended audience. Consistent with the presence of frictions, we find that when applying a copy-editing software that counts the pervasiveness of the most important "writing faults" to journal articles published from 2005 through 2014, articles with one more writing fault per one hundred words receive 7% fewer citations.

JEL Classification: A10, M30.

Keywords: Scientific Knowledge Dissemination, Readability.

* Hwang and Wu are affiliated primarily with the Cornell SC Johnson College of Business, Dyson School of Applied Economics and Management, Cornell University, Warren Hall, Ithaca, NY 14853; Hwang is also affiliated with Korea University Business School, Korea University, Seongbuk-gu, Seoul, Korea. Kim is affiliated with the Darla Moore School of Business, University of South Carolina, 1014 Green Street, Columbia, SC 29208. E-mail: bhwang@cornell.edu, hugh.kim@moore.sc.edu and kw445@cornell.edu. The authors would like to thank Richard Wu, who provided excellent research assistance and Alex Butler, Clifton Green, David Hirshleifer, Soo Kim and Siew Hong Teoh for helpful comments.

1. Introduction

Ideas proposed in scientific journal articles have the potential to touch many areas of our lives in significant ways. It is thus important that these ideas fully reach their intended audiences. In this paper, we examine whether there are frictions in the scientific dissemination process that prevent ideas from doing so. In particular, we test whether something as seemingly frivolous as the "form" of a scientific journal article, here understood to be the ease with which an article can be read, affects the degree to which a scientific article reaches its intended audience.

Our suspicion that a lack of clear writing may impede scientific knowledge dissemination is couched within a growing body of work that studies the effects of the readability of a text. The psychology literature provides evidence that low readability weakens readers' trust in the text; low readability also causes readers to subconsciously evaluate the text less favorably (McGlone and Tofighbakhsh 2000, Oppenheimer 2006, Alter and Oppenheimer 2008). Recent work in accounting and finance examines how the ease with which the text in corporate disclosure documents can be processed affects investors and the underlying firm. Miller (2010) finds evidence that difficult-to-read annual reports create disagreement among small investors. Lawrence (2013) and Elliott, Rennekamp and White (2015) provide evidence that investors are less likely to invest in firms that have difficult-to-read disclosure documents. Hwang and Kim (2017) provide evidence that companies with difficult-to-read annual reports trade at significant discounts.[1]

We conjecture that the above found effects of readability may carry over to academia. Scientists produce hundreds of journal articles every year. Staying current, even in one's "narrow" area of expertise, thus represents a non-trivial task. Given the limited amount of time that the intended audience can spend on each journal article, we speculate that the audience may stop reading a hard-to-read article after a few paragraphs, even if the quality of the underlying idea is high. Relatedly, if there are two journal articles of

[1] In line with the above work, regulators and government agencies increasingly emphasize the importance of clear writing. In 1998, the Securities and Exchange Commission (SEC) published the Plain English Handbook (1998), which provides both language-related and formatting suggestions intended to increase the readability of corporate disclosure documents. In 2010, Congress enacted the Plain Writing Act, which requires federal agencies to write "clear Government communication that the public can understand and use." (Public Law 111–274, 111th Congress, October 13, 2010).

the same "function," but one article is harder to read, readers may subconsciously develop negative sentiments towards the harder-to-read article and perceive that article to be less interesting than the article that is easier to read.

To test whether a lack of clear writing in a scientific journal article impedes an idea from reaching its intended audience, we turn to the field of financial economics (or, simply, finance).[2] We do not believe the observations we make in this paper are particular to finance, however, and we later also provide descriptive statistics for articles published in the fields of accounting, economics, management, marketing, and operations and information systems.

We download all papers that were published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014. As our measure of "reaching its intended audience," we use the number of Google Scholar citations that the relevant article garners as of September 2016. We apply copy-editing software to each paper and count the number of times various "writing faults" that make a document harder to read appear in the text. We use the (scaled) pervasiveness of these writing faults as our measure of readability. Writing faults include the use of passive verbs, hidden verbs, complex words, abstract words, overused words and clichés, legal words, wordy phrases, overwriting, foreign words and long sentences. Section 2 describes and presents examples of each writing fault.

To assess the validity of our measure, we randomly assign finance PhD students the introduction sections of articles that, as per our measure, earn "low readability" scores. We repeat this procedure with introduction sections of articles that earn "high readability" scores. We find that students largely agree with the output generated by our readability measure, as they perceive introductions with low readability scores to be significantly more difficult to read than those with high readability scores. We also find hints in the

[2] We focus on one field (only) as the construction of many of our variables is labor-intensive. We choose finance because we, the authors of this study, all work in finance departments and, as such, are familiar with the publication process in the finance area.

data that introductions with low readability scores are perceived to be less interesting than introductions with high readability scores.

When relating the readability of scientific journal articles to their number of citations within a regression framework, we find that papers of lower readability have substantially fewer citations than papers of higher readability. In particular, our estimates imply that one more writing fault per one hundred words comes with 7% fewer citations. For reference, the average paper in our sample receives 204 citations; the median number of citations is 115; the average article in our sample has 6.1 writing faults per one hundred words; the standard deviation is 1.1. Our results easily survive the inclusion of various controls, such as years-since-publication, number of authors, author affiliation, number of conference and seminar presentations, whether the paper won a best-paper award, whether the paper is a theory paper, title length, the subfield of the paper (e.g., "Portfolio choice and investment decisions," "Bankruptcy and litigation") and journal fixed effects.

The interpretation of our observed negative association between citations and writing faults is complicated by the fact that we cannot perfectly control for the true quality of an article's underlying idea. The lack of a perfect control raises the possibility that part of the observed difference in citations stems from differences in function rather than from differences in form.

While we cannot perfectly rule out this possibility, we find the results from the following additional analysis revealing: Each year, we select articles that are of "Highest Function yet Bad Form," and we contrast their impact to that of articles that are of "Regular Function yet Good Form". In particular, the *Journal of Finance*, the *Journal of Financial Economics* and *The Review of Financial Studies* each award best-paper prizes to the best articles that were published in a given year. These awards represent some of the most prestigious prizes in the field of finance. Award-winning articles are not free of writing faults, however, and the correlation between winning an award and number of writing faults is essentially zero.

When we compare (1) the impact of award-winning papers that are in the bottom decile with respect to their readability scores with (2) that of non-award-winning papers that are in the top decile with respect to their readability scores, we find that award-winning papers receive just as many citations as non-award-

winning papers, even though the average award-winning paper is arguably of substantially higher function than the average non-award-winning paper.

In the end, our observations from both the experimental and the regression-based settings are at the very least consistent with the notion that packaging an idea in confusing writing impedes an idea from reaching its intended audience. Articles with high function, but bad form, thus, do not receive the attention they deserve, thereby distorting the knowledge dissemination process.

2. Data and Key Variables

2.1. Measure of Impact

We download all papers that were published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014. In total, we cover 2,618 papers, 716 of which are from the *Journal of Finance*, 1,048 of which are from the *Journal of Financial Economics*, and 854 of which are from *The Review of Financial Studies*. For each of these papers, we manually search the number of citations in Google Scholar (https://scholar.google.com/intl/en/scholar/about.html) via title, authors, and year of publication. We collect the number of citations as of September 16–20 2016, which should ensure that differences in citations do not reflect differences in points of data collection. As shown in Table 1, the average paper in our sample generates 204 citations. The median number of citations is 115; the 10^{th} and 90^{th} percentiles are 25 and 465, respectively.

2.2. Other Article and Author Characteristics

For each article in our sample, we compute the number of years since publication (as of 2016), the number of authors, an *Affiliation Ranking* variable, the number of conferences and seminars the paper has been invited for presentation prior to publication, whether the article won a "best-paper award," whether the article is a theory paper, the length of the title (in words), and the number of JEL codes.

To elaborate on some of the above variables, the *Affiliation Ranking* variable captures the "prestige" of the institutions the authors of the relevant articles are affiliated with. We turn to the *Finance Research*

Ranking compiled by the Arizona State University (http://apps.wpcarey.asu.edu/fin-rankings/rankings). The *Finance Research Ranking* counts the number of publications in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* by institution and runs from 1990 through the present. To avoid look-ahead bias, we save the Top 50 ranking as of 2004, the year before our sample begins.[3] To facilitate interpretation, the number one institution receives a score of 50, the number two institution receives a score of 49, and so on. Any institution that is not represented in the Top 50 receives a score of zero. A high score therefore indicates that the institution in question has a strong reputation. We compute the average score across the institutions with which the author or authors of the paper in question are primarily affiliated. Taking the score of the highest ranking institution instead of the average score does not materially alter any of our results (results available upon request).

Our *Award Paper* variable is constructed as follows: Each year, the *Journal of Finance*, the *Journal of Financial Economics* and *The Review of Financial Studies* award best-paper prizes to the best published articles in that year. Our *Award Paper* variable equals one if the relevant article won such a prize, and zero otherwise. The awards are the Amundi Smith Breeden Prize and the Brattle Group Prize for the *Journal of Finance*; the Jensen Prize and the Fama-DFA Prize for the *Journal of Financial Economics*; and the Michael J. Brennan Best Paper Award for *The Review of Financial Studies*.

Theory Paper equals one if the corresponding text contains the term "proof" along with either one of the following terms: "proposition," "theorem," "lemma," "corollary".

The *Number of JEL Codes* variable counts the number of Journal of Economic Literature (JEL) codes for the relevant article. For papers published in the *Journal of Finance*, which do not have JEL codes, the *Number of JEL Codes* variable is set at zero. JEL codes represent subfields in economics.[4] For instance, a paper with JEL code "G11" is a paper on "Portfolio Choice and Investment Decisions"; a paper with JEL code "G33" is a paper on "Bankruptcy and Litigation." Most articles have multiple JEL codes. In our

[3] The top-ten schools in our *Affiliation Ranking* are (1) NYU, (2) the University of Pennsylvania, (3) the University of Chicago, (4) Harvard University, (5) UCLA, (6) the University of Michigan, (7) Northwestern University, (8) Duke University, (9) Columbia University and (10) MIT. The full top-fifty is available upon request.
[4] The complete list of JEL codes can be found here: https://www.aeaweb.org/econlit/jelCodes.php.

regression analysis, we include journal-fixed effects along with fixed effects based on the article's first JEL code.

2.3. Measure of Readability

Our approach to measuring readability is very similar to that adopted in Hwang and Kim (2017), who study how the readability of corporate disclosure documents affects firm value. We save each paper as a separate Microsoft Word document. We then use a program called StyleWriter, a manuscript editor that, once installed on a computer, searches Word documents for "writing faults."

The writing faults (along with examples of the writing faults and how to avoid them) are:

(1) Passive verbs:

Example: "*We must re-think how our resources will be best used to provide world-class customer service*."

Possible correction: "*We must re-think how to best use our resources to provide world-class customer service*."

(2) Hidden verbs:

A hidden verb is a verb used as a noun. It is often hooked to an extra (weak) verb.

Example: "*. . . to make an application for employment*."

Possible correction: "*. . . to apply for employment*."

(3) Complex words:

Example: "*While third parties sometimes endeavor to ameliorate relationships . . .*"

Possible correction: "*While third parties sometimes try to improve relationships . . .*"

(4) Abstract words:

By avoiding abstract words, writers can clarify the message they are trying to convey.

Example: "*We need to install more output devices*."

Possible correction: "*We need to install more printers*."

(5) Overused words and Clichés:

The former are popular terms used in a variety of settings; they can essentially mean anything you want them to mean. The latter are phrases that have become devalued through overuse.

Examples of overused words: "*The patient was then informed about the parameters of treatment available . . . we must more carefully study the parameters of our health care system.*"

Possible corrections: "*The patient was then informed about the types of treatment available . . . we must more carefully study the limitations of our health care system.*"

Examples of clichés: "*open a can of worms*," "*we beg to differ*," "*wakeup call*"

(6) Legal words:

Example: "*forthwith*"

Possible correction: "*immediately*"

(7) Wordy phrases:

Examples: "*an appreciable number of*," "*has a requirement for*"

Possible corrections: "*many*," "*requires*"

(8) Overwriting:

Example: "*It is completely unnecessary.*"

Possible correction: "*It is unnecessary.*"

(9) Foreign words:

Examples: "*The results show a high urban crime rate, inter alia . . . our paper helps fill a lacuna in the literature . . .*"

Possible corrections: "*The results show a high urban crime rate, among others . . . our paper helps fill a gap in the literature . . .*"

(10) Long sentences:

There is no objective criterion as to what constitutes a long sentence. Cutts (2013) in the *Oxford Guide to Plain English*, for instance, recommends an average sentence length of 15–20 words. In our study, we follow our software's definition of a long sentence, which is a sentence with more than 35 words.

Our readability measure, *Readability*, is the number of occurrences of the above writing faults, scaled by the number of words and multiplied by (100) and (-1).

$$Readability = \frac{\sum_{i=1}^{10} WritingFaults_i}{\#Words} \times (100) \times (-1). \quad (1)$$

Multiplying by one hundred later helps us interpret the coefficient estimates. We multiply by negative one so that higher readability scores imply more easily readable documents. For instance, a score of negative ten implies that there are ten writing faults per one hundred words. A more positive score of negative one implies that there is only one writing fault per one hundred words.

We acknowledge alternate measures of readability, perhaps the two most popular of which are the *Fog Index* and the *Flesch-Kincaid Index*. Both measures are designed to gauge the number of years of formal education needed to comprehend a text on a first reading. The *Fog Index* is 0.4 × (Average Number of Words per Sentence + Fraction of Complex Words × 100). The *Flesch-Kincaid Index* is 0.39 × (Total Number of Words / Total Number of Sentences) + 11.8 × (Total Number of Syllables / Total Number of Words) - 15.59 (Kincaid, Fishburne, Rogers and Chissom, 1975). Other measures of readability include document length and the file size of an electronic document (e.g., Li, 2008; Lawrence, 2013; Loughran and McDonald, 2014).

All of the above measures have their place in the literature. At the same time, we conjecture that using a measure based on actual writing faults that writing classes and textbooks teach us to avoid increases the power of the analysis. Hwang and Kim (2017) provide experimental and regression-based evidence to this regard.[5]

Table 1 shows that there is great variation in our readability measure in our sample. The 10th and 90th percentiles for *Readability* are -7.5 and -4.8; the mean and standard deviation are -6.13 and 1.12,

[5] Still, our measure of readability is admittedly crude and ignores important attributes of readability such as logical structure or "getting to the point." We believe that whatever effect we find in this study may thus be considered a mere prelude to the overall effect of readability on the intended audience.

respectively. The mean of -6.13 implies that, on average, for every one hundred words, there are 6.1 writing faults. For reference, the *Readability* of this paper is -5.4.

Figure 1 shows the time-series average *Readability* for each of the three finance journals separately. *Readability* is fairly stable through time. On average, the *Journal of Financial Economics* has the highest readability score, followed by the *Journal of Finance* and *The Review of Financial Studies*. Overall, the differences in *Readability* across the finance journals are not economically meaningful, however.

Table 2 extends the across-journal comparison to all fields in business and (general) economics. As (general) economics "A-level" journals, we include the *American Economic Review*, the *Journal of Political Economy* and the *Quarterly Journal of Economics*. Our list of A-level business journals is that of the *Business School Research Rankings*^{TM} compiled by the University of Texas (Dallas) (http://jindal.utdallas.edu/the-utd-top-100-business-school-research-rankings). In addition to the *Journal of Finance*, the *Journal of Financial Economics* and *The Review of Financial Studies*, the list includes: *Journal of Accounting and Economics*, *Journal of Accounting Research*, *The Accounting Review*, *Academy of Management Journal*, *Academy of Management Review*, *Administrative Science Quarterly*, *Journal of International Business Studies*, *Organization Science*, *Strategic Management Journal*, *Journal of Consumer Research*, *Journal of Marketing*, *Journal of Marketing Research*, *Marketing Science*, *Information Systems Research*, *MIS Quarterly*, *Journal on Computing*, *Journal of Operations Management*, *Manufacturing and Service Operations Management*, *Operations Research*, *Production and Operations Management* and *Management Science*. Since the computation of *Readability* is labor-intensive, we only consider the 1,752 articles in the above journals that were published in 2014, the last year of our sample period.

Table 2 reveals strong differences across journals and fields. Economics- and accounting journals fare relatively well with *Readability* ranging from -6.02 for the *Journal of Political Economy* to -6.64 for *The Accounting Review*. Marketing journals also fare relatively well with the quantitative and managerial journals having slightly fewer writing faults than the more psychology-based journal (*Journal of Marketing*

= -6.55, *Journal of Marketing Research* = -6.56, *Management Science* = -6.65, *Marketing Science* = -6.95, versus *Journal of Consumer Research* = -7.19).

Management journals tend to have lower readability scores. The journals with the lowest readability scores, however, are in the field of Operations and Information Systems. The average article in the *Journal of Operations Management* has 8.7 writing faults per one hundred words; the average article in *Information Systems Research* has 8.3 writing faults per one hundred words. Some of the lower *Readability* may be the result of subject matters. At the same time, there are journals in the field of Operations and Information Systems that have substantially fewer writing faults, on average, such as the *Journal on Computing* (-6.68) and *Production and Operations Management* (-6.82).

Returning to our main sample of articles published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014, Table 3 presents a correlation matrix of our primary variables. Table 3 shows that both *Affiliation Ranking* and *Number of Presentations* are significantly positively correlated with *Citations*. *Theory Paper* is significantly negatively correlated with *Citations*. However, none of these measures reliably associate with *Readability*. That is, there are no systematic differences in the occurrence of writing faults between "high-affiliation-rank" papers and "less-high-affiliation-rank" papers, between papers that have been presented many times prior to publication and papers that have been presented fewer times, and between theory papers and non-theory papers.

3. Experimental Evidence of the Validity and Effectiveness of our Readability Measure

Before estimating the impact of clear writing, we pause to assess the validity of our readability measure in an experimental setting. We also attempt to shed light on how the readability of scientific journal articles influences readers' moods and perceptions of how interesting an underlying idea is.

We randomly select twenty papers that are in the top *Readability* quartile ("high readability papers") and twenty papers that are in the bottom quartile ("low readability papers"). We then assign these papers to finance PhD students and ask them to rate the readability of the introduction section. We focus on the

introduction section, as reading the entire paper would require too much time for the PhD students. The average *Readability* of the introductions of high readability papers is -5.68; the average *Readability* of the introductions of low readability papers is -8.71; the difference is -3.04 (*t*-statistic = -9.01).

To ensure that our sample of papers represents all areas of finance, we adopt the following procedure: Of the twenty randomly chosen high-readability papers (low readability papers),

- five are from the pool of papers that are in the area of financial markets (JEL codes: G10-G19) and are purely empirical
- five are from the pool of papers that are in the area of financial markets and contain a theoretical model
- five are from the pool of papers that are in the areas of Financial Institutions & Services and Corporate Finance & Governance (JEL codes: G20-G39) and are purely empirical
- five are from the pool of papers that are in the areas of Financial Institutions & Services and Corporate Finance & Governance and contain a theoretical model

We are unable to match papers with survey participants based on area of expertise. Our survey participants report that 36.67% of the papers that they were assigned to read are in their area of expertise. There is no reliable difference in this fraction between the twenty high readability papers and the twenty low readability papers.

Our subject pool consists of twenty-one finance PhD students: eighteen PhD students were asked to read six introductions (three high readability introductions and three low readability introductions); three PhD students were asked to read four introductions (two high readability introductions and two low readability introductions). Each of the forty introductions is read by three finance PhD students. The students are from the following schools: Cornell University, Emory University, Indiana University, University of Southern California, University of Washington, and Yale University.

We ask the students to answer the following set of questions after reading each introduction:

(1) "*How easy to read was the introduction? The scales are 7 ("Very easy") to 1 ("Not at all easy").*"

(2) "*How do you feel right now (1/2)? The scales are 7 ("Calm") to 1 ("Bothered").*"

(3) "*How do you feel right now (2/2)? The scales are 7 ("Relaxed") to 1 ("Tense").*"

(4) "*How interesting do you find the paper? The scales are 7 ("Very interesting") to 1 ("Not at all interesting").*"

Table 4 reports the average readability score produced by the students for the group of twenty high-readability papers and the group of twenty low-readability papers. Because each paper is read by three students, each of the two cells contains sixty observations.

Papers that are in the top quartile based on *Readability* receive an average score of 5.38. In comparison, papers that are in the bottom quartile receive an average score of 4.70. The difference is +0.68 (*t*-statistic = 2.70). Since students generally avoided the extremes and mostly assigned scores of four, five or six, the difference of +0.68 is economically meaningful. The relatively strong agreement of survey participants with the outputs generated by our readability measure helps build confidence in the validity of our measure.

Table 4 also reveals that introductions of lower readability are associated with slightly worse moods, as subjects felt slightly more "bothered" and "tense" when reading introductions of low readability than when reading introductions of high readability. For the first sentiment-related question ("calm" versus "bothered"), the average score given for high- and low-readability articles are 5.13 and 4.80, respectively (Δ = 0.33, *t*-statistic = 1.23). For the second sentiment-related question ("relaxed" versus "tense"), the average score given for high- and low-readability articles are 5.00 and 4.85, respectively (Δ = 0.15, *t*-statistic = 0.59).

Since our papers were randomly drawn, the intrinsic level of interest should be similar across high- and low readability papers. Yet, we observe that papers with highly readable introductions are perceived to be slightly more interesting. Papers that are in the top quartile based on *Readability* receive an average score of 4.57; papers that are in the bottom quartile receive an average score of 4.32. The difference is +0.25 (*t*-statistic = 0.91).

The above results are much stronger when we pool all 120 observations together and correlate the response to the first question ("*How easy to read was the introduction?*") with the responses to the second question (*How do you feel right now? 1/2*), the responses to the third question (*How do you feel right now? 2/2*), and the responses to the fourth question (*How interesting do you find the paper?*). The Pearson Correlation Coefficients are +0.68 (p-value < 0.01), +0.59 (p-value < 0.01) and +0.56 (p-value < 0.01), respectively.

Ultimately, while the results reported in Section 3 should be interpreted with some caution given the relatively small sample sizes, we find reasonably strong evidence of the utility of our measure of readability. We also find hints in the data that the readability of a text alters how subjects view the text and the idea couched within the text.

4. Regression-Based Evidence on the Effect of Readability

To quantify the effect of clear writing, we estimate the following regression equation:

$$Y_{i,t} = \alpha + \beta\, Readability_{i,t} + X\delta + \varepsilon_{i,t}. \qquad (2)$$

The dependent variable is the natural logarithm of the number of citations. We take the natural logarithm since *Citations* is highly right-skewed. *X* includes *Years since Publication*, *Number of Authors*, *Affiliation Ranking*, *Number of Presentations*, *Award Paper*, *Theory Paper*, *Length of Title*, and *Number of JEL Codes*, all of which are described above. We also include journal fixed effects and fixed effects based on the paper's first JEL code. *T*-statistics are based on standard errors adjusted for heteroscedasticity and clustered by the first JEL code.

We present our regression results in Table 5. Depending on the set of controls, the coefficient estimate for *Readability* ranges from 0.063 (*t*-statistic = 3.83) to 0.084 (*t*-statistic = 5.35). Our regression analysis thus indicates that, holding all else equal, one more writing fault per one hundred words comes with around 7% fewer citations. This result is consistent with the notion that a lack of clear writing hinders an idea from reaching its intended audience.

A few notes on the coefficient estimates for the control variables: The number of authors on a paper strongly positively associates with the number of citations. One interpretation of this finding is that having co-authors improves the quality of a paper. Another possibility is that co-authors help raise awareness of a paper through their own personal networks.

The coefficient estimate for *Affiliation Ranking* is 0.009 (*t*-statistic = 9.05). Rankings may represent another friction in the knowledge dissemination process and work produced by authors not affiliated with "high reputation" institutions may not receive the attention it deserves. Alternatively, "high reputation" institutions may produce higher quality work.

The estimate for *Number of Presentations* is 0.025 (*t*-statistic = 7.46), suggesting that presenting the paper one more time at a conference or a university prior to publication increases subsequent citation counts by 2.5%. Conferences and seminars at universities are primary vehicles through which researchers can draw attention to their work prior to publication. In addition, papers that have benefitted from comments received during talks are likely of higher quality. Vice versa, high-quality papers by high-quality authors are also more likely to get accepted/invited for presentation at conferences and universities.

Not surprisingly, receiving an award strongly and positively contributes to citation counts. Our estimate implies that, holding all else equal, award-winning papers, on average, receive 46.2% more citations than non-award papers. Theory papers, on the other hand, receive 45.4% fewer citations.

Interestingly, the length of a paper's title negatively associates with the number of citations. Papers with short titles may be broader ("*The Cross-Section of Expected Stock Returns*" versus "*The Effect of Introducing a Non-Redundant Derivative on the Volatility of Stock-Market Returns When Agents Differ in Risk Aversion*"). Shorter titles may also be more attention-grabbing. For instance, when conducting a randomized field experiment on *Seeking Alpha*, a leading investments-related social media platform, Umar (2016) finds that articles posted on *Seeking Alpha* with twice the title length receive 30% fewer page views. In an online survey, Umar also finds that participants perceive the news that "Apple *Inc.* Sees iPhone Sales Slump" to be more interesting than the news that "Apple *Incorporated* Sees iPhone Sales Slump".

5. Portfolio-Based Evidence on the Effect of Readability

Our analysis demonstrates that articles with more writing faults receive fewer citations. Our interpretation of this empirical phenomenon is that lack of clear writing represents a friction in the knowledge dissemination process that prevents potentially important knowledge from reaching its intended audience. An alternative perspective is that articles with better ideas also happen to have fewer writing faults.

As alluded to before, in the data, we do not observe that articles of seemingly higher function come with more writing faults. The correlations between *Readability* and *Award Paper*, the correlation between *Readability* and *Number of Presentations,* and the correlation between *Readability* and *Affiliation Ranking* are all close to zero.

To further gauge the plausibility of the alternative interpretation, we conduct the following experiment: Each year, we sort articles independently by *Readability* and various proxies for article quality. Our proxies for article quality are whether the paper won a best-paper award, whether the paper has been invited for presentation disproportionately many times, the affiliation of the authors and the number of authors. We then compare the impact of two groups of articles: (1) articles that score high based on the above proxies, yet have low readability ("Highest-Function articles") and (2) articles that score low based on the above proxies, yet have high readability ("Regular-Function articles"). Even if *Readability* positively correlates with quality, our "direct" proxies for article quality, such as whether the article won a best-paper award, should correlate much more strongly with the true quality of the article. The average Highest-Function article should therefore have higher intrinsic quality than the average Regular-Function article, irrespective of the number of writing faults. That is, the average award-winning paper should be of higher quality than the average non-award-winning paper. Any lower citation count that the group of award-winning (yet bad form) papers receives can then be attributed to bad form rather than lower function.

In our first sort, we form portfolios of articles that won a best-paper award, yet are in the bottom decile with respect to *Readability*. We also form portfolios of articles that did not win a best-paper, yet are in the top decile with respect to *Readability*. We do this each year from 2005 through 2014. We then

compare the average annual citations for articles in the Highest-Function Portfolio with that for articles in the Regular-Function Portfolio.

Table 6 reports the time-series mean of the cross-sectional means. We find that the average Regular-Function article receives 37 citations/year. In comparison, the average Highest-Function article receives 38 citations/year and, as such, has similar impact even though the average best-paper article is arguably of higher quality than the average non-award-winning paper.

We make similar observations for the other proxies of article quality. In our second sort, Highest-Function articles are defined as top-decile articles with respect to *Number of Presentations*, yet bottom-decile articles with respect to *Readability*. Regular-Function articles are defined as bottom-decile articles with respect to *Number of Presentations*, yet top-decile articles with respect to *Readability*. Here, we find that Regular-Function articles have substantially higher citation counts than their Highest-Function counterparts (40 versus 28).

In our third sort, Highest-Function articles are defined as articles for which at least one of the authors is from a Top 25 institutions as per the *Finance Research Ranking*, yet reside in the bottom decile with respect to *Readability*; Regular-Function articles are defined as articles for which none of the authors is from a Top 25 institutions, yet reside in the top decile with respect to *Readability*. Again, we find that Regular-Function articles have similar impact as Highest-Function articles.

6. Conclusion

Our study examines whether the impact of a scientific journal article is solely determined by its function, or whether something as seemingly frivolous as form also matters. On the "consumption of knowledge side," we propose that when a paper becomes difficult to read, readers only partially absorb the ideas proposed in the paper and/or subconsciously develop negative sentiments and perceive the paper to be less interesting. Holding function constant, this causes ideas packaged in bad form to lose impact.

On the "production of knowledge side," we observe considerable variation in readability, with some papers reading rather well while others suffer from passive verbs, complex words and wordy phrases,

among other problems. Some scientists may be overconfident and erroneously believe their writing to be superb. Others may remain unconvinced of the incremental benefit of writing easier-to-read papers. Still others may take pride in their ability to construct complex phrases and use terms such as “inter alia” and “lacuna.” Whatever causes the observed differences in readability, our evidence suggests that lack of clear writing can significantly lower the impact of an article and hinder the dissemination of potentially important knowledge.

The “form” of an article is only one out of many possible frictions in the scientific dissemination knowledge and, ex ante, perhaps not even the most important friction. Our findings may thus be considered a mere prelude to the overall presence of frictions in the scientific knowledge dissemination process.

References

Alter, A.L., Oppenheimer, D.M., 2008. Easy on the mind, easy on the wallet: The effects of familiarity and fluency on currency valuation. Psychonomic Bulletin and Review 15, 985-990.

Cutts, M., 2013. Oxford guide to Plain English. Oxford University Press (4th edition).

Garner, B.A., 2013. Legal writing in Plain English. The University of Chicago Press (2nd edition).

Hwang, B.H., Kim, H., 2017. It pays to write well. Journal of Financial Economics 124(2), 373–394..

Kincaid, J.P., Fishburne, R.P., Rogers, R.L., Chissom, B.S., 1975. Derivation of new readability formulas (automated readability index, fog count, and flesch reading ease formula) for Navy enlisted personnel. Research Branch Report *8-75*, Chief of Naval Technical Training.

Lawrence, A., 2013. Individual investors and financial disclosure. Journal of Accounting and Economics 56(1), 130-147.

Lehavy, R., Li, F., Merkley, K., 2011. The effect of annual report readability on analyst following and the properties of their earnings forecasts. Accounting Review 86(3), 1087-1115.

Li, F., 2008. Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics 45(2-3), 221-247.

Loughran, T., McDonald, B., 2014. Measuring readability in financial disclosures. Journal of Finance 69(4), 1643–1671.

McGlone, M.S., Tofighbakhsh, J., 2000. Birds of a feather flock conjointly: Rhyme as reason in aphorisms. Psychological Science 11(5), 424–428.

Miller, B.P., 2010, "The effects of reporting complexity on small and large investor trading," Accounting Review, 85(6), 2107-2143.

Oppenheimer, D.M., 2006. Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. Applied Cognitive Psychology 20(2), 139-156.

Rennekamp, K. 2012. Processing fluency and investors' reactions to disclosure readability. Journal of Accounting Research 50(5), 1319–1354.

Securities and Exchange Commission (SEC), 1998. A Plain English handbook: How to create clear SEC disclosure documents. Washington DC, USA.

Umar, T. 2016. Attention grabbers when Seeking Alpha. University of Chicago Working Paper.

Figure 1
Readability of Scientific Journal Articles by Finance Journal over Time

This figure plots our measure of readability, *Readability*, across 2,618 scientific journal articles published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014. *Readability* is the number of writing faults per one hundred words, multiplied by negative one. The lines represent the average *Readability* by journal and year.

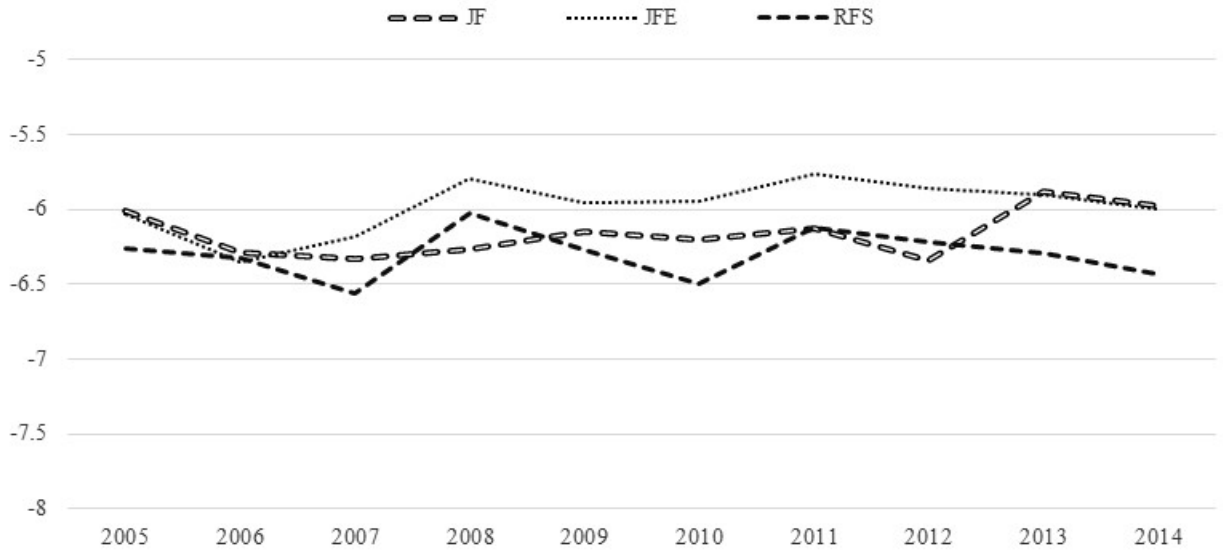

Table 1
Descriptive Statistics

This table presents summary statistics for our main variables. The sample includes 2,618 scientific journal articles published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014. The *Journal of Finance* has 716 papers, the *Journal of Financial Economics* has 1,048 papers, and *The Review of Financial Studies* has 854 papers. *Citations* is the Google citations the paper receives as of September 16–20, 2016. *Readability* is the number of writing faults per one hundred words multiplied by negative one. *Years since Publication* is the number of years since publication (as of year 2016). *Number of Authors* is the number of authors listed in the paper. *Affiliation Ranking* is the average "ranking score" across the schools the authors are primarily affiliated with as detailed in Section 2.2. *Number of Presentations* is the number of conferences and seminars the paper was presented at prior to publication. Each year, the *Journal of Finance*, the *Journal of Financial Economics* and *The Review of Financial Studies* award best-paper prizes to the best published papers in that year. *Award Paper* equals one if the paper won such a prize. *Theory Paper* equals one if a paper contains the term "proof" along with either one of the following terms: "proposition," "theorem," "lemma," "corollary". *Length of Title* is the number of words in the title. *Number of JEL Codes* is the number of JEL codes listed in the paper.

	N	Mean	StDev	10^{th} Percentile	Median	90^{th} Percentile
Citations	2,618	203.56	293.14	25.00	115.00	465.00
Readability	2,618	-6.13	1.12	-7.50	-6.10	-4.80
Years since Publication	2,618	6.19	2.83	2.00	6.00	10.00
Number of Authors	2,618	2.39	0.87	1.00	2.00	3.00
Affiliation Ranking	2,618	18.00	16.40	0.00	15.67	44.50
Number of Presentations	2,618	8.68	6.50	1.00	8.00	17.00
Award Paper	2,618	0.05	0.21	0.00	0.00	0.00
Theory Paper	2,618	0.14	0.35	0.00	0.00	1.00
Length of Title	2,618	8.80	3.48	5.00	8.00	13.00
Number of JEL Codes	2,618	2.28	1.84	0.00	2.00	5.00

Table 2
Readability of Scientific Journal Articles by Business and (General) Economics Journals

This table reports the average readability score, *Readability*, across all scientific journal articles published in the relevant journal in 2014.

Readability Ranking	(1) Journal	(2) Field(s)	(3) *Readability*
1	*Journal of Finance*	Finance	-5.98
2	*Journal of Financial Economics*	Finance	-6.00
3	*Journal of Political Economy*	Economics (General)	-6.02
4	*American Economic Review*	Economics (General)	-6.02
5	*Quarterly Journal of Economics*	Economics (General)	-6.12
6	*Journal of Accounting Research*	Accounting	-6.21
7	*Journal of Accounting and Economics*	Accounting	-6.40
8	*The Review of Financial Studies*	Finance	-6.43
9	*Journal of Marketing*	Marketing	-6.55
10	*Journal of Marketing Research*	Marketing	-6.56
11	*The Accounting Review*	Accounting	-6.64
12	*Management Science*	Accounting, Finance, Management, Marketing, Operations and Information	-6.65
13	*Journal on Computing*	Operations and Information	-6.68
14	*Production and Operations Management*	Operations and Information	-6.82
15	*Operations Research*	Operations and Information	-6.92
16	*Marketing Science*	Marketing	-6.95
17	*Strategic Management Journal*	Management	-6.95
18	*Journal of Consumer Research*	Marketing	-7.19
19	*MIS Quarterly*	Operations and Information	-7.20
20	*Administrative Science Quarterly*	Management	-7.22
21	*Manufacturing and Service Operations*	Operations and Information	-7.25
22	*Academy of Management Journal*	Management	-7.39
23	*Journal of International Business Studies*	Management	-7.44
24	*Academy of Management Review*	Management	-7.70
25	*Organization Science*	Management	-7.79
26	*Information Systems Research*	Operations and Information	-8.33
27	*Journal of Operations Management*	Operations and Information	-8.71

Table 3
Correlation Matrix

This table presents Pearson correlation coefficients across our main independent and dependent variables. The sample includes 2,618 scientific journal articles published in the *Journal of Finance*, the *Journal of Financial Economics* and *The Review of Financial Studies* from 2005 through 2014. Correlations that are significant at the 5% level are in bold.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) *Citations*	1.000									
(2) *Readability*	**0.056**	1.000								
(3) *Years since Publication*	**0.353**	**-0.045**	1.000							
(4) *Number of Authors*	0.019	**-0.041**	**-0.047**	1.000						
(5) *Affiliation Ranking*	**0.151**	0.031	**0.039**	**-0.100**	1.000					
(6) *Number of Presentations*	**0.059**	0.004	**-0.177**	**0.080**	**0.132**	1.000				
(7) *Award Paper*	**0.149**	0.039	0.023	-0.023	**0.135**	**0.072**	1.000			
(8) *Theory Paper*	**-0.083**	-0.034	0.035	**-0.114**	**0.062**	**0.086**	-0.015	1.000		
(9) *Length of Title*	-0.024	**-0.070**	**0.127**	**0.062**	**-0.049**	**-0.083**	**-0.050**	**-0.141**	1.000	
(10) *Number of JEL Codes*	**-0.071**	0.001	-0.007	**0.099**	-0.023	-0.036	**-0.079**	**-0.044**	**0.150**	1.000

Table 4
Experimental Evidence on of the Validity and the Effectiveness of our Readability Measure

This table presents survey responses from Finance PhD students that are pertinent to the readability of scientific journal articles. We conduct the following experiment: We sort introduction sections of papers based on *Readability*. We randomly select twenty papers from the top quartile ("High Readability") and twenty papers from the bottom quartile ("Low Readability"). We assign these introductions to twenty-one PhD students and ask the following questions:

(1) "How easy to read was the introduction?" The scales range from 7 ("Very easy") to 1 ("Not at all easy").
(2) "How do you feel right now (1/2)?" The scales range from 7 ("Calm") to 1 ("Bothered").
(3) "How do you feel right now (2/2)?" The scales range from 7 ("Relaxed") to 1 ("Tense").
(4) "How interesting do you find the paper?" The scales range from 7 ("Very interesting") to 1 ("Not at all interesting").

Each introduction is read by three students, yielding a total of sixty observations in each of the two cells. We report the average score given by the students for the "High Readability" articles and the "Low Readability" articles. *T*-statistics, reported in parentheses, account for heteroscedasticity. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

	(1) High Readability Papers	(2) Low Readability Papers	(3) Δ High- and Low Readability Papers
(1) "How easy to read was the introduction?" - Scale: 7 ("Very easy") to 1 ("Not at all easy")	5.38	4.70	0.68*** (2.70)
(2) "How do you feel right now (1/2)?" - Scale: 7 ("Calm") to 1 ("Bothered")	5.13	4.80	0.33 (1.23)
(3) "How do you feel right now (2/2)?" - Scale: 7 ("Relaxed") to 1 ("Tense")	5.00	4.85	0.15 (0.59)
(4) "How interesting do you find the paper?" - Scale: 7 ("Very interesting") to 1 ("Not at all interesting")	4.57	4.32	0.25 (0.91)

Table 5
Effect of Readability: Regression-Based Evidence

This table presents coefficient estimates from regressions of the natural logarithm of *Citations* on various article characteristics. The sample includes 2,618 scientific journal articles published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014. All variables are as described in Table 1. *T*-statistics are reported in parentheses and are based on standard errors adjusted for heteroscedasticity and clustered by year and the first JEL code. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)
Readability	0.063***	0.084***	0.072***
	(3.83)	(5.35)	(4.98)
Years since Publication		0.192***	0.210***
		(8.56)	(12.12)
Number of Authors			0.117***
			(6.03)
Affiliation Ranking			0.009***
			(9.05)
Number of Presentations			0.025***
			(7.46)
Award Paper			0.462***
			(2.69)
Theory Paper			-0.454***
			(-11.10)
Length of Title			-0.024***
			(-5.07)
Number of JEL Codes			0.010
			(0.62)
Journal FE	Yes	Yes	Yes
1st JEL Code FE	Yes	Yes	Yes
R-squares	0.148	0.339	0.412
Number of observations	2,618	2,618	2,618

Table 6
Effect of Readability: Portfolio-Based Evidence

This table presents the average *Citations* per year across articles sorted independently each year by *Readability* and other article characteristics. The sample includes 2,618 scientific journal articles published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014. Column (1) represents articles of "Highest Function yet Bad Form". An article is of "Highest Function" if the article won a best-paper award, if the article is in the top decile with respect to *Number of Presentations*, if at least one of the authors is from a Top 25 Institutions as per the *Finance Research Ranking*, or if the article is in the top decile with respect to *Number of Authors*. "Bad Form" is defined as being in the bottom decile with respect to *Readability*. Column (2) represents articles of "Regular Function yet Good Form". An article is of "Regular Function" if the article did not win a best-paper award, if the article is in the bottom decile with respect to *Number of Presentations*, if none of the authors are from a Top 25 Institutions as per the *Finance Research Ranking*, or if the article is in the bottom decile with respect to *Number of Authors*. "Good Form" is defined as being in the top decile with respect to *Readability*.

"Function" Criterion	(1) "Highest Function yet Bad Form"	(2) "Regular Function yet Good Form"
Award-winning paper?	38.15	36.81
High number of presentations prior to publication?	27.63	39.59
Author from Top 25 school?	27.91	29.32
High number of authors?	27.79	50.64