

Writing matters

Jan Feld, Corinna Lines, Libby Ross^{*}

[March 15, 2022]

Abstract

For papers to have scientific impact, they need to impress our peers in their role as referees, journal editors, and members of conference committees. Does better writing help our papers make it past these gatekeepers? In this study, we estimate the effect of writing quality by comparing how 30 economists judge the quality of papers written by PhD students in economics. Each economist judged five papers in their original version and five different papers that had been language edited. No economist saw both versions of the same paper. Our results show that writing matters. Compared to the original versions, economists judge edited versions as higher quality; they are more likely to accept edited versions for a conference; and they believe that edited versions have a better chance of being accepted at a good journal.

Keywords: academic writing, writing quality, economics, randomized experiment

^{*} Jan Feld is Senior Lecturer in Economics at the School of Economics and Finance, Victoria University of Wellington, New Zealand. Corinna Lines and Libby Ross are plain language consultants at Write Limited in Wellington, New Zealand. We thank Sam Lentle-Keenan for her help at the beginning of this project, as well as Alexandra de Gendre, Nicolas Salamanca, and Ulf Zölitz for helpful comments. We also owe a big thank you to the PhD students, expert economists, and writing experts who participated in the experiment. For questions and comments about this paper, contact Jan at jan.feld@vuw.ac.nz

1. Introduction

Across the sciences, we have heard calls for better academic writing (e.g. Pinker 2015; Sword 2012; Salant 1969). In economics, McCloskey has argued since the 1980s that economists should write better, and demonstrates how (McCloskey 1992, 1987, 1985). Her book *Economical Writing* has since become a standard part of the curriculum for economics PhD students (McCloskey 2000).

Many economists and other scientists follow these calls and try to write better. They spend days toiling over their introduction, writing and rewriting, and spend their research budgets on language editing. Yet we do not know if our writing matters. If we write better, do our peers as referees, journal editors, and members of conference committees perceive our papers as better?

The answer to this question has implications for who gets to contribute to the scientific discourse. In most disciplines, written English has become the language of modern science. This dominance of English makes it harder for scientists from non-English-speaking countries to publish in leading academic journals — especially if they cannot afford language editing. As a result, we may be losing important contributions from scientists who struggle to write well in English.

Academic literature does not provide a satisfactory answer to how important writing quality is. Many papers have investigated the correlation between writing quality and scientific impact as measured, for example, with numbers of citations. Depending on the discipline, and measures of writing quality and scientific impact, these papers found that better-written papers have more impact, similar impact, or less impact (e.g. Didegah and Thelwall 2013; Dowling, Hammami, and Zreik 2018; Laband and Taylor 1992; Fages 2020; Hartley, Trueman, and Meadows 1988). It does not take a PhD in economics to recognize the limitations of these papers. Correlation is not causation. Papers that are well written likely differ on several dimensions from papers that are not. To find the causal effect of academic writing, we need to compare well-written papers with poorly written papers that are otherwise identical. This is what we do in this study.

We estimate the causal effect of writing quality by comparing how experts judge the quality of 30 papers originally written by PhD students in economics. We had two versions of each paper: one original and one that had been language-edited. The language editing was done

by two professional editors, who aimed to make the papers easier to read and understand. We then asked 18 writing experts and 30 economists to judge some of the original and edited papers. Each of these experts judged five papers in their original versions and five papers in their edited version, spending around 5 minutes per paper. None of the experts saw both versions of the same paper. None of the experts knew that some of the papers were edited. The writing experts judged the writing quality and the economists judged the academic quality of the papers. All economists in our sample have PhDs in economics and their academic positions range from postdoc to full professor; four of them are editors of academic journals; and all of them are regularly involved in judging the quality of academic papers as referees or members of conference committees. We estimate the effect of language editing on perceived writing quality and perceived academic paper quality by comparing the average judgement of original and edited papers.

Our results show that writing matters. Writing experts judged the edited papers as 0.6 standard deviations (SD) better written overall (1.22 points on an 11-point scale). They further judged the language-edited papers as allowing the reader to find the key message more easily (0.58 SD), having fewer mistakes (0.67 SD), being easier to read (0.53 SD), and being more concise (0.50 SD). These large improvements in writing quality translated into still substantial effects on economists' evaluations. Economists evaluated the edited versions as being 0.2 SD better overall (0.4 points on an 11-point scale). They were also 8.4 percentage points more likely to accept the paper for a conference, and were 4.1 percentage points more likely to believe that the paper would get published in a good economics journal. Our heterogeneity analysis shows that the effects of language editing on writing quality and perceived academic quality are particularly large if the original versions were poorly written.

The approach of manipulating text to estimate the effect of writing has been used in several contexts, for example, for legal documents (Mindlin 2005) and financial reports (Tan, Wang, and Yoo 2019). However, only one other paper has investigated how the writing quality of academic papers affects the evaluations of scientists. Armstrong (1980) altered the writing quality of the conclusion section in four management papers. In contrast to our findings, his results suggest that improving the writing causes experts to evaluate papers less favorably. We improve upon Armstrong's approach by having a larger sample, as well as a more rigorous study design and empirical analysis. For example, our study includes 30 papers (instead of four) and the language editing was done by professional editors (instead of Armstrong, who is a good

writer but not an expert editor). By asking scientists to evaluate whole papers instead of one individual section, the context of our study is also closer to how peer review is conducted in practice. Finally, since Armstrong conducted his study, attitudes towards writing have changed. We have seen the birth of the plain language movement, and several countries now require government and other agencies to write according to plain language principles. For example, the United States passed The Plain Writing Act of 2010, requiring all government agencies to write in plain language (Office of the Federal Register 2010). And the European Union requires companies to communicate clearly with customers about how they use their data (European Union 2022). Our study therefore provides the best answer of whether the quality of academic writing matters today.

2. Design of experiment and empirical strategy

To estimate the effect of writing quality on expert evaluations, we conducted an experiment which consisted of two stages. In the first stage, we collected and edited papers from economics PhD students. In the second stage, we asked writing experts and economists to evaluate the original version of some papers and the edited version of other papers. We received ethics approval for this experiment from Victoria University's Human Ethics Committee (application number 27561) and pre-registered the experiment at the AEA registry (Feld 2021).

2.1. Stage 1: Language-editing 30 economics papers

From May 2020 until May 2021, we collected and edited 30 papers from economics PhD students at New Zealand universities. To find these papers, we emailed PhD students and their supervisors from all eight New Zealand universities and offered them free language editing in exchange for filling in a short survey and allowing us to use their papers in our experiment (see Appendix A.1 for the text of the survey). Each PhD student could participate with up to two papers. In total, 22 PhD students accepted this offer: eight of them were female, 13 were male, and one did not disclose their gender. Only four PhD students were native English-speakers. By the time they submitted their papers to the experiment, they had worked on them for eight months on average; nine of the papers had been submitted to a conference; five papers had been submitted to a journal; none of the papers had been accepted for publication; 19 were about

microeconomics (18 empirical micro and 1 theoretical) and 11 were about macroeconomics (10 empirical and 1 theoretical).

Each of these papers was edited by two professional language editors. The first 18 papers were edited by Libby Ross and Sam Lentle-Keenan. The remaining 12 papers were edited by Libby Ross and Corinna Lines. At the time of editing the papers, all editors worked for Write Limited, a New Zealand-based consultancy which helps organizations make their writing easier to understand. Write Limited is not usually involved in editing academic papers.

The editing process worked as follows. One of the editors focused on improving the structure of the paper, for example, by rearranging parts of the introduction. After that, the second editor polished the writing at the sentence and word level. The editors then sent the paper back to the PhD student, asking them to answer any questions and to check whether the edited version had retained the original meaning. After receiving the revised version, the editors incorporated the changes suggested by the PhD students and asked the PhD students to check any substantial changes. On average, both editors spent a total of 6 hours per paper.

The editors broadly followed a guideline which Jan — the only academic economist among the co-authors — created with the help of Libby and Sam. This guideline states that the edits should aim to make the papers easier to understand and that edits should focus on the title, abstract, and introduction. It also contained information on academic writing in economics to help the editors understand the conventions of the discipline. For example, it detailed the key information to be included in the abstract and introduction. It also contained several writing tips taken from the editors' own experience, as well as articles and books about academic writing (Sword 2012; Pinker 2015; Cochrane 2005; Williams and Colomb 1995; McCloskey 2000). For example, the guideline stated, “Keep a short distance between nouns and their accompanying verb” and “Use simple, familiar words”. The complete guideline can be found in Appendix B.

2.2. Stage 2: Expert evaluation of original and edited versions of each paper

From May 2021 until October 2021, writing experts and academic economists filled in a short survey and evaluated the original and edited version of each paper (See Appendix A.2 and A.3 for the text of these surveys).

We recruited writing experts with the help of posts on the LinkedIn and Facebook profiles of plain language consultancy Write, and the Community Comms Collective. We also

found writing experts through emailing the network of Editors Aotearoa New Zealand. These posts and emails looked for help with a “research project on writing” and showed how to get in touch with us. Writing experts were offered vouchers worth NZD 50 for their participation. Our sample size was pre-registered and determined by our available funds. Twenty-six writing experts got in touch with us, allowing us to recruit the 18 writing experts as pre-registered. We invited economists from Australian universities and research institutions by email, offering them a voucher worth AUD 50 for their participation. We emailed 167 economists to achieve our pre-registered sample size of 30, which resulted in an 18 percent success rate.

Table 1 shows summary statistics for both kinds of experts. The participating writing experts had different professions related to writing, such as copywriter, technical writer, and communications manager. All of them stated in the survey that writing and reading were important in their job; 16 of them were women; 17 were native English-speakers. All participating economists are experts who are regularly involved in judging the quality of academic papers. All of them have a PhD in economics, which they received on average 11 years before participating in the experiment; their academic positions range from postdoc to full professor. All of them have refereed papers for a journal, with an average of six papers refereed per year; four are editors at academic journals; 25 have been involved in deciding which papers get accepted for a conference. As typical for economists, our sample was male-dominated and very international: 23 economists were men; 16 were non-native English-speakers.

We sent each participating writing expert and economist an online survey containing a few questions about themselves, links to ten papers, and five questions about each paper. Each expert was asked to spend around 8 minutes per paper. The median of the actual time spent on each paper was somewhat shorter: 5.7 minutes for writing experts and 4.5 minutes for economists. Our paper is therefore particularly informative for situations in which experts evaluate the quality of papers in a short period of time — as is often done, for example, when deciding whether to desk-reject a paper at a journal or whether to accept a paper for a conference.

Table 1: Summary statistics of economists and writing experts

	(1)	(2)	(3)	(4)	(5)
	N	mean	SD	min	max
Panel A: language experts					
Female	18	0.89	0.32	0	1
English first language	18	0.94	0.24	0	1
Writing important in job	18	1.00	0.00	1	1
Reading important in job	18	1.00	0.00	1	1
Panel B: Economists					
Female	30	0.23	0.43	0	1
Editor	30	0.13	0.35	0	1
English first language	30	0.47	0.51	0	1
Years since PhD	30	11.27	6.51	1	27
PhD in Economics	30	1.00	0.00	1	1
<i>Academic position</i>					
Postdoc	30	0.03	0.18	0	1
Lecturer	30	0.23	0.43	0	1
Senior lecturer	30	0.50	0.51	0	1
Associate professor	30	0.10	0.31	0	1
Full professor	30	0.13	0.35	0	1
Papers refereed per year	29	6.03	4.65	2	25
Has judged paper for conference	30	0.83	0.38	0	1
<i>Field of expertise</i>					
Empirical microeconomics	30	0.63	0.49	0	1
Theoretical microeconomics	30	0.20	0.41	0	1
Empirical macroeconomics	30	0.33	0.48	0	1
Theoretical macroeconomics	30	0.13	0.35	0	1

To make sure each expert evaluated the same number of original and edited papers, we randomly assigned papers to blocks which contained five original and five edited papers. No paper block contained both versions of the same paper. For economists, these block assignments considered the expertise of the economists and the topic of the paper. Macroeconomists were assigned to blocks containing 10 different macro papers. Since we had 11 macro papers in total, we moved one macro paper that could easily be judged by microeconomists to the pool of micro papers. Each microeconomist was therefore either assigned a block containing 10 different micro papers, or a block containing 9 different micro papers and 1 macro paper. Writing experts were

assigned blocks with a mixture of micro and macro papers. We assigned the experts to blocks in the order in which they agreed to participate in the experiment. We describe how we matched papers to experts in greater detail in Appendix C. In each survey, all papers within a block were presented in random order.

None of the experts knew that any of the papers were edited. Furthermore, only the writing experts were told that this was a project about writing. To gauge to what extent economists guessed the aim of the study, we asked them at the end of the survey to guess what this research was about. Only 6 out of 30 correctly guessed that the project was about writing quality.

Writing experts were asked to evaluate the writing quality of each paper by answering the following five questions:

- 1) “Overall, the quality of the paper is...”
- 2) “The paper allows me to easily find the key messages.”
- 3) “The paper is free of spelling and grammar mistakes.”
- 4) “The paper is easy to read.”
- 5) “The paper is concise.”

The answer scale for the first question ranged from 0 “very bad” to 10 “very good”. The answer scale for the remaining questions ranged from 0 “completely disagree” to 10 “completely agree”.

Each economist was asked to evaluate the quality of each paper by answering the following five questions:

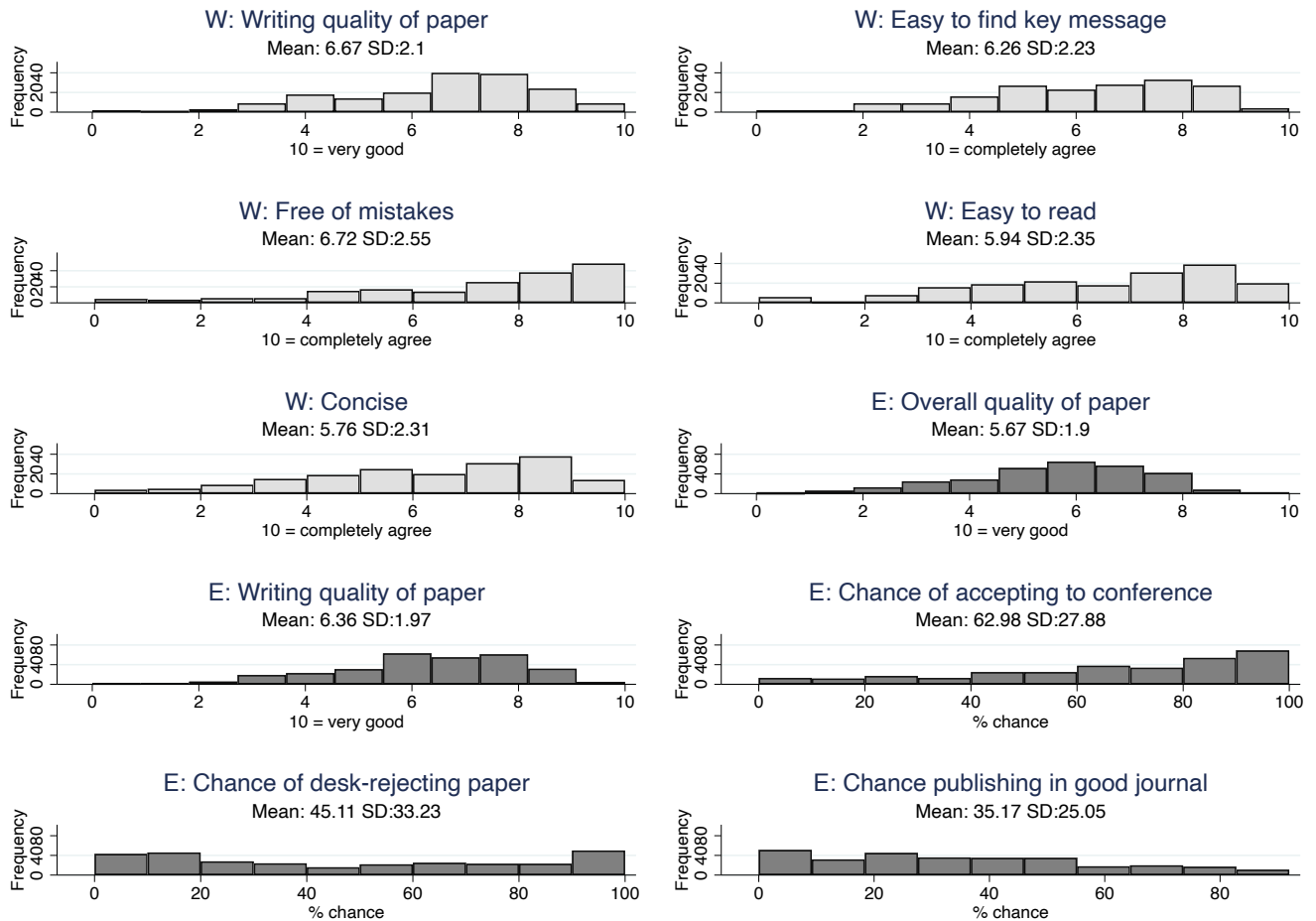
- 1) “Overall, the quality of the paper is...”
- 2) “How likely would you be to accept this paper at a general economics conference (such as the Australian Conference of Economists)?”
- 3) “Imagine you were an editor of a general economics journal that is an A journal on the ABDC list. How likely would you be to desk reject this paper?”
- 4) “How likely is it that this paper will get published in an A or A* journal on to the ABDC list?”
- 5) “Overall, the quality of the writing is...”

Questions 3) and 4) reference the Australian Business Deans Council (ABDC) journal ranking, which is widely used in Australia for hiring and promotion decisions (ABDC 2019). This ranking classifies journals in order of quality into A*, A, B, and C categories. As examples,

A* journals include the *American Economic Review* and the *Journal of Labor Economics*, and A journals include *Economics of Education Review* and *Southern Economic Journal*. Questions 1) and 5) could be answered on a scale ranging from 0 “very bad” to 10 “very good”. The remaining questions could be answered using a slider with answer options that ranged from “0 percent” to “100 percent”.

Figure 1 shows the means, standard deviations, and distributions for each of the outcome variables taken from writing experts (first five graphs) and economists (last five graphs). This figure shows that papers included in the study varied substantially in their writing and academic quality. On average, writing experts judged the overall writing quality with 6.7 points. They further agreed, with 6.3 points, that the paper made it easy to find the key message; with 6.7 points that the papers were free of mistakes; with 5.9 points that the papers were easy to read; and with 5.8 points that the papers were concise. On average, economists judged the quality of the papers with 5.7 points and the quality of the writing with 6.4 points. They stated an average chance of 63 percent of accepting the paper at a conference; a 45 percent chance of desk-rejecting the paper; and a 35 percent chance of the paper ending up in an A or A* journal.

Figure 1: Distributions of outcome variables



Note: This figure shows the distributions, means, and standard deviations of our main dependent variables. The first five graphs (in light grey) show the outcome variables for writing experts. The next five graphs (in dark grey) show the outcomes for economists. The average of the writing quality variable, taking into account the judgements of writing experts and economists, is 6.48 and its standard deviation is 2.03.

2.3. Empirical strategy

We estimate the effect of writing quality by comparing the average expert evaluations of the edited and original paper versions. If experts evaluate the edited versions of the papers more positively, this would show a causal effect of writing quality under two assumptions. First, the language editing must not have changed the technical content of the paper. Second, the language editing must have improved the writing quality.

The design of our experiment ensures that changes to the technical content are highly unlikely. The language editors were instructed to not change the content, and also do not have the expertise to do so. Jan, who has this expertise, was not involved in the editing process.

Furthermore, PhD students were instructed to not change the content of the paper when reviewing the edited versions. We checked and confirmed that they followed this instruction.

We test whether the edits improved the writing quality by comparing how writing experts evaluated the different versions of the papers. We also test whether original and edited versions differ on the Flesch-Kincaid grade level score (Kincaid et al. 1975). This score is a frequently used measure of writing quality which scores texts as easier to read if they have fewer words per sentence and words with fewer syllables. It is calculated with the following formula: $-15.59 + 0.39 * (\text{number of words/number of sentences}) + 11.80 * (\text{syllables/words})$. This formula gives texts that are deemed to be more readable a lower score. The scaling approximates the necessary grade level of reading ability to understand a text. For example, if a text has a Flesch-Kincaid grade level score of 8, the reader needs a grade 8 level of reading to understand it. We measured the Flesch-Kincaid grade level score for the introduction as this was the most-edited part of the papers.

Throughout the paper, we implement our comparisons with ordinary least squares (OLS) regressions of the outcome of interest (e.g. experts' evaluations) on a dummy variable which is equal to 1 if the evaluated paper was edited and 0 if it was in its original version. We also estimate the effect of language editing separately for papers that were poorly written and well written in their original version. To classify papers as poorly written, we use the average of writing experts' judgements of the overall writing quality of the original version. Each original version was judged by three different writing experts. We classify 15 papers as *poorly written* for which the average of these judgements was smaller or equal than 6 points, and the remaining 15 papers as *well written*. To estimate separate effects of editing for these kinds of papers, we re-do our analysis replacing the edited dummy with three variables: one dummy variable indicating whether the original paper was poorly written, one interaction term of the edited dummy and the poorly written dummy (*edited* \times *poorlywritten*), and one interaction term of the edited dummy and a well-written dummy (*edited* \times *wellwritten*). The coefficients of these interaction terms

show the average difference between original and edited versions for each of these two kinds of papers.¹

To increase the precision of our estimates, we also control for paper fixed effects and expert fixed effects. Neither of these sets of fixed effects affect our point estimates. Because each paper is evaluated multiple times, we cluster our standard errors at the paper level. We leave the dependent variables that show percentages and grades levels in their natural units. All other dependent variables are standardized to have means of 0 and standard deviations of 1. Note that one of our outcome variables — assessment of overall writing quality — is available for writing experts and economists. We standardize this variable over the whole estimation sample, so that coefficients of the estimated effect of editing on perceived writing quality by economists and by writing experts are directly comparable.

3. Results

3.1. Language editing improves writing quality

Panel A of Table 2 shows that language experts evaluated the edited versions as better written. All point estimates go in the expected direction and are statistically significant at the 1 percent level. Writing experts evaluated the edited papers as being 0.60 SD better written overall (1.22 points on the 11-point scale). They further evaluated the edited papers as making it 0.58 SD easier to find the key message, as having 0.67 SD fewer mistakes, as being 0.53 SD easier to read, and as being 0.50 SD more concise.

The language editing also affected the readability as measured by the Flesch-Kincaid grade level score. The introductions of edited papers have a readability score corresponding to grade level 14.7, compared to 15.3 of the introductions of original papers. This improvement of 0.6 grade-levels is statistically significant at the 1 percent level. For comparison, our introduction has a Flesch-Kincaid grade level score of 12.5. Abstracts of papers published in the *American*

¹ To see that this is the case, consider the following econometric model:

$$y = \beta_0 + \beta_1 \text{poorlywritten} + \beta_2 \text{edited} \times \text{poorlywritten} + \beta_3 \text{edited} \times \text{wellwritten} + u.$$

When we estimate this model with OLS, the average expert judgement of poorly written papers in their original version is $\widehat{\beta}_0 + \widehat{\beta}_1$ and the average judgement of poorly written papers in their edited version is $\widehat{\beta}_0 + \widehat{\beta}_1 + \widehat{\beta}_2$. The difference between these two averages is therefore captured by the coefficient on the interaction term $\widehat{\beta}_2$.

Analogously, the average judgement of well-written papers in their original version is $\widehat{\beta}_0$ and the average of well-written papers in the edited version is $\widehat{\beta}_0 + \widehat{\beta}_3$, which leaves $\widehat{\beta}_3$ as the difference between those two differences.

Economic Review, *Econometrica*, the *Journal of Political Economy*, and the *Quarterly Journal of Economics* have an average Flesch-Kincaid grade level score of around 13 (Hengle in press). The results of the expert evaluations and readability scores combined show a clear picture: edited papers are better written.

Table 2: Effect of language editing on experts' evaluation of papers

	(1) Std. Writing Quality	(2) Std. Key Message	(3) Std. No mistakes	(4) Std. Easy to read	(5) Std. Concise
<i>Panel A: Writing experts</i>					
Edited	0.603*** (0.138)	0.582*** (0.140)	0.666*** (0.150)	0.529*** (0.120)	0.496*** (0.119)
Observations	179	180	180	180	180
R-squared	0.587	0.589	0.595	0.591	0.595
Mean outcome of original versions	-.207	-.291	-.333	-.264	-.248
<i>Panel B: Economists</i>					
	Std. Paper Quality	Conference	Desk-reject	Publish	Std. Writing Quality
Edited	0.203** (0.095)	8.433*** (2.671)	-5.087 (3.385)	4.053** (1.976)	0.197* (0.104)
Observations	300	300	300	300	300
R-squared	0.652	0.551	0.548	0.653	0.611
Mean outcome of original versions	-.102	58.76	47.653	33.14	-.156

Note: This table reports results from OLS regressions of the dependent variables shown in the column headers on a dummy variable indicating whether the paper was edited as well as paper fixed effects and rater fixed effects. The first column in Panel A only has 179 observations because one of the experts did not rate the writing quality for one of the papers. Standard errors clustered at the paper level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

3.2. Economists evaluate edited papers more positively

Panel B of Table 2 shows that economists evaluate edited papers more positively. All five point-estimates go in the expected direction and three are statistically significant. Economists judge the overall paper quality 0.20 SD better (0.4 points on the 11-point scale). They are also 8.4 percentage points more likely to accept edited papers for a conference, and are 4.1 percentage points more likely to believe that edited papers will get published in an economics journal that is

classified as A* or A on the ABDC journal ranking. While not reaching statistical significance at conventional levels, the point estimates also suggest that economists would be 5 percentage points less likely to desk-reject edited papers, and judge edited papers as 0.20 SD better written.

The magnitudes of these effects are economically significant, especially considering that the language editors only spent 6 hours per paper. While the use of different outcome variables does not allow us to compare our results directly, we can see they are in the same ballpark as estimated effects of favoritism in the editorial process. At the *Journal of Human Resources*, the effect of having ever been a colleague of an editor increases the chances of making it past the editorial desk by 3.7 percentage points (Carrell, Figlio, and Lusher 2022). At the same journal, having ever been a colleague of a referee increases the chance of a positive evaluation by 3.7 percentage points. Relatedly, the estimated effect of having an all-male-authored paper compared to an all-female-authored paper on being accepted at an economics conference is 3.3 percentage points (Hospido and Sanz 2021).

3.3. Effect of language editing on poorly written and well-written papers

We would expect larger effects of language editing if the original papers were poorly written and left more room for improvement. The results shown in Table 3 suggest that this is indeed the case. Panel A shows that language editing particularly improved the writing quality of poorly written papers. All point estimates are statistically significant and large. For example, poorly written papers that have been language edited are judged 1.05 SD better written than poorly written papers in their original version. The estimated effects of editing on well-written papers are also positive for all five outcomes, but we only see one statistically significant effect: well-written papers that have been edited are judged to have 0.35 SD fewer mistakes.

We see similar results for the effect of language editing on the economists' evaluations. Panel B shows economists rate poorly written papers that were edited significantly more positively compared to poorly written papers in the original version for three outcomes. For poorly written papers, editing improved the overall quality by 0.30 SD, the chances of getting accepted at a conference by 11.11 percentage points, and the overall writing quality by 0.4 SD. For well-written papers, all point estimates are smaller, and none are statistically significant. Language editing appears to be particularly useful for poorly written papers.

Table 3: Effect of language editing on poorly written and well-written papers

	(1) Std. Writing Quality	(2) Std. Key Message	(3) Std. No mistakes	(4) Std. Easy to read	(5) Std. Concise
<i>Panel A: Effect on Writing experts' evaluations</i>					
Edited × Poorly written papers	1.041*** (0.151)	0.773*** (0.190)	0.981*** (0.255)	0.790*** (0.166)	0.820*** (0.126)
Edited × Well-written papers	0.173 (0.139)	0.391* (0.194)	0.350** (0.138)	0.267* (0.137)	0.171 (0.161)
Observations	179	180	180	180	180
R-squared	0.629	0.598	0.619	0.608	0.620
Mean outcome of poorly written originals	.16	.291	.398	.222	.171
Mean outcome of well-written originals	-.053	-.097	-.133	-.074	-.057
p-value of F-test for equality of coefficients	.0002	.1729	.0498	.0216	.0047
<i>Panel B: Effect on Economists' evaluations</i>					
	Std. Paper Quality	Conference	Desk- reject	Publish	Std. Writing Quality
Edited × Poorly written papers	0.291** (0.135)	11.113** (4.105)	-5.962 (4.528)	5.079* (2.735)	0.402*** (0.140)
Edited × Well-written papers	0.116 (0.144)	5.754 (3.570)	-4.211 (5.399)	3.028 (3.052)	-0.007 (0.143)
Observations	300	300	300	300	300
R-squared	0.654	0.554	0.548	0.653	0.622
Mean outcome of poorly written originals	.077	66.8	45.787	36.56	.002
Mean outcome of well-written originals	-.026	61.702	44.884	34.702	-.077
p-value of F-test for equality of coefficients	.4069	.3494	.8123	.6332	.0595

Note: This shows the estimated effect of language editing on writing experts' evaluations (Panel A) and economists' evaluations (Panel B) separately for poorly written papers and well-written papers. More specifically, both panels show results from OLS regressions on the dependent variables shown in the column headers. The independent variables in these regressions are: one dummy variable indicating whether the paper's original version was poorly written; one interaction term of the edit dummy and an indicator of the paper being classified as poorly written; one interaction term of the edit dummy and an indicator of the paper being classified as well written; as well as paper and rater fixed effects. The definition of poorly written and well-written papers is explained in Section 2.3. The p-values reported in the last row stem from F-tests of equality of the two main interaction terms. Standard errors clustered at the paper level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

4. Conclusion

We have shown that writing matters. Making the writing easier to understand causes economists to evaluate academic papers more positively. This finding has obvious implications. To improve your chances of publishing well, you can work on your writing by, for example, spending time polishing your paper and paying for language editing. These efforts are likely to be particularly valuable if you find writing challenging.

Our finding also has implications for who contributes to the scientific discourse. Writing quality is not random. Because English is the language of most of modern science, non-native speakers are at a disadvantage. Similarly, researchers who cannot afford language editing will find it more difficult to get papers past the gatekeepers of science. This is not only a problem of lack of representation, but we likely also miss out on important ideas. After all, the answer to life, the universe, and everything might be hidden in a poorly written paper.

Should the gatekeepers ignore the writing quality? That is difficult to do and may not be desirable. Writing is an important aspect of the quality of an academic paper. Better writing makes our ideas easier to understand for our fellow scientists, policy makers, and the general public. A more promising solution would be to help struggling scientists with their writing. For example, the American Economic Association could offer free online courses on academic writing. Or the Bill and Melinda Gates Foundation could offer free language editing to researchers in developing countries. Alongside those initiatives, we are all responsible for writing well. We know this can be hard, but those efforts matter.

References

- ABDC. 2019. "ABDC Journal Quality List." accessed January 21, 2022.
<https://abdc.edu.au/research/abdc-journal-quality-list/>.
- Armstrong, J Scott. 1980. "Unintelligible management research and academic prestige."
Interfaces 10 (2):80-86.
- Carrell, Scott E, David N Figlio, and Lester R Lusher. 2022. Clubs and Networks in Economics Reviewing. National Bureau of Economic Research.
- Cochrane, John H. 2005. *Writing tips for Ph. D. students*: University of Chicago Chicago, IL.
- Didegah, Fereshteh, and Mike Thelwall. 2013. "Which factors help authors produce the highest impact research? Collaboration, journal and document properties." *Journal of Informetrics* 7 (4):861-873.

- Dowling, Michael, Helmi Hammami, and Ousayna Zreik. 2018. "Easy to read, easy to cite?" *Economics Letters* 173:100-103. doi: 10.1016/j.econlet.2018.09.023.
- European Union. 2022. "Writing a GDPR-compliant privacy notice (template included)." accessed February 5. <https://gdpr.eu/privacy-notice/>.
- Fages, Diego Marino. 2020. "Write better, publish better." *Scientometrics* 122 (3):1671-1681.
- Feld, Jan. 2021. "Paper Quality." *AEA RCT Registry*. doi: <https://doi.org/10.1257/rct.5620-2.0>.
- Hartley, James, Mark Trueman, and AJ Meadows. 1988. "Readability and prestige in scientific journals." *Journal of Information Science* 14 (2):69-75.
- Hengle, Erin. in press. "Publishing while female." *Economic Journal*
- Hospido, Laura, and Carlos Sanz. 2021. "Gender gaps in the evaluation of research: evidence from submissions to economics conferences." *Oxford Bulletin of Economics and Statistics* 83 (3):590-618.
- Kincaid, J Peter, Robert P Fishburne Jr, Richard L Rogers, and Brad S Chissom. 1975. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. Naval Technical Training Command Millington TN Research Branch.
- Laband, David N, and Christopher N Taylor. 1992. "The impact of bad writing in economics." *Economic Inquiry* 30 (4):673-688.
- McCloskey, Deirdre N. 2000. *Economical writing*. 2nd ed. Prospect Heights, Ill.: Waveland Press.
- McCloskey, Donald N. 1985. "Economical writing." *Economic Inquiry* 23 (2):187-222.
- McCloskey, Donald N. 1987. "Reply to Jack High." *Economic Inquiry* 25 (3):547.
- McCloskey, Donald N. 1992. "Writing as a responsibility of science: a reply to Laband and Taylor." *Economic Inquiry* 30 (4):689-695.
- Mindlin, Maria. 2005. "Is Plain Language Better-A Comparative Readability Study of Court Forms." *Scribes J. Leg. Writing* 10:55.
- Office of the Federal Register, National Archives and Records Administration. 2010. Plain Writing Act.
- Pinker, Steven. 2015. *The sense of style: The thinking person's guide to writing in the 21st century*: Penguin Books.
- Salant, Walter S. 1969. "Writing and reading in economics." *Journal of Political Economy* 77 (4, Part 1):545-558.
- Sword, Helen. 2012. *Stylish academic writing*: Harvard University Press.
- Tan, Hun-Tong, Elaine Ying Wang, and G-Song Yoo. 2019. "Who likes jargon? The joint effect of jargon type and industry knowledge on investors' judgments." *Journal of Accounting and Economics* 67 (2-3):416-437.
- Williams, J, and G Colomb. 1995. *Style: Toward Clarity and Grace*. Chicago guides to writing, editing, and publishing. University of Chicago Press.

Appendix A – Text of surveys

A.1: Survey sent to PhD students

What is your name?

What is your gender?

- Male
- Female
- Other
- Prefer not to say

Is English your first language?

- Yes
- No

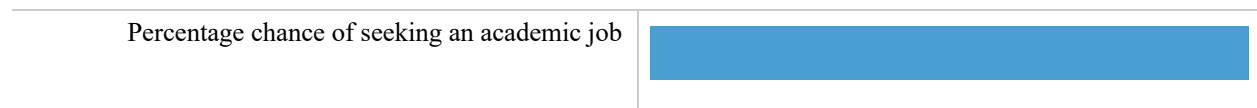
Are you currently studying towards a PhD in economics?

- Yes
- No, I have already completed my PhD in economics
- No, I am studying towards a PhD in another discipline
- No, other. Please explain _____

In which year did you start your PhD?

After finishing your PhD, what is the probability that you will seek an academic job?

0 10 20 30 40 50 60 70 80 90 100



Please state how much you agree with the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I find academic writing difficult	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I enjoy academic writing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think my papers are well-written	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Have you ever taken a course on academic writing?

Yes

No

Have you ever read a book on academic writing?

Yes

No

Have you learned about academic writing in any other way?

Yes. Please explain _____

No

What is the title of your paper? (if you participate in this experiment with two papers, just state the title of one of them)

Please briefly describe for non-economists what this paper is about.

Which topic describes the content of this paper best?

- Empirical microeconomics
- Theoretical microeconomics
- Empirical macroeconomics
- Theoretical macroeconomics
- Other, please specify: _____

How many months have you spent working on this paper?

Have you submitted this paper to a conference?

- Yes
- No

Have you submitted this paper to a journal for publication?

- Yes
- No

Has this paper been accepted for publication in a journal?

- Yes
- No

How would you judge the quality of this paper. Overall, the quality of the paper is

- Very bad (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- Very good (10)

How would you judge the overall quality of the writing of this paper. Overall, the quality of the writing is

- Very bad (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- Very good (10)

Do participate with a second paper in this experiment?

- Yes
- No

What is the title of your second paper?

Please briefly describe for non-economists what this paper is about.

Which topic describes the content of this paper best?

- Empirical microeconomics
- Theoretical microeconomics
- Empirical macroeconomics
- Theoretical macroeconomics
- Other, please specify: _____

How many months have you spent working on this paper?

Have you submitted this paper to a conference?

- Yes
- No

Have you submitted this paper to a journal for publication?

- Yes
- No

Has this paper been accepted for publication in a journal?

- Yes
- No

How would you judge the quality of this paper. Overall, the quality of the paper is

- Very bad (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- Very good (10)

How would you judge the overall quality of this paper. Overall, the quality of the writing is

- Very bad (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

9

Very good (10)

Thank you for filling in this questionnaire! Is there anything we should know? Please tell us below.

Would you like to receive a first draft of this research project once it is available?

Yes

No

A.2: Survey sent to economists

What is your gender?

- Male
- Female
- Other
- Prefer not to say

Do you have a PhD in economics?

- Yes
- No, I have a PhD in another discipline
- No, I don't have a PhD

Is English your first language?

- Yes
- No

In which year did you receive your PhD?

In which country did you receive your PhD?

▼ Afghanistan ... Zimbabwe

What is your current position?

- Postdoc
- Lecturer
- Senior lecturer
- Associate professor
- Full professor
- Other, please specify: _____

Have you been involved in deciding which papers get accepted for a conference?

- Yes
- No

On average, how many papers do you referee per year?

Are you an editor of an academic journal?

- Yes
- No

What are your fields of expertise?

- Empirical microeconomics
- Theoretical microeconomics
- Empirical macroeconomics
- Theoretical macroeconomics
- Other, please specify: _____

Instructions (part 1)

Thanks for agreeing to judge the quality of ten academic papers. On top of each page is a link to a paper. Please download this paper and review it. Below the link to the paper are five questions. Please answer all of them. I suggest you take 8 minutes per paper. To be able to review a paper quickly, you can read the abstract and introduction, and then skim read the rest of the paper.

Instructions (part 2) One of these questions asks about the overall quality of the paper. Please judge the quality of the paper as you would judge other papers when refereeing or deciding which paper to accept for a conference. Two of these questions reference the Australian Business Deans Council (ABDC) journal quality list (also called ABDC list). Are you familiar with this list?

Yes

No

Instructions (part 2 continued) This ABDC list ranks journals in terms of their quality using the letters A*, A, B and C

The 2019 ABDC list endorses 2,682 journals with the following classifications:

A* 7.41% (199)

A 24.27% (651)

B 31.69% (850)

C 36.61% (982)

Here are some examples of economics journals that are classified as A*, A, B, and C journals

A*: American Economic Review, Economic Journal, Journal of Economic Behavior and Organization

A: Economic Inquiry, Applied Economics, Economics of Education Review

B: Australian Economic Papers, New Zealand Economic Papers, Applied Economics Letters

C: Asian Economics, Bulletin of Applied Economics, Indian Economic Review

The ABDC list is used in Australian and New Zealand universities for judging academic performance, for example, in promotion decisions.

For more information, see: <https://abdc.edu.au/research/abdc-journal-quality-list/>

Instructions (part 3)

If you have any questions about these instructions, send an email to jan.feld@vuw.ac.nz.

The next page shows the first paper.

[At this stage, the survey included 10 pages, each with a link to a different paper and the same 5 questions. For brevity, we only show one of these pages]

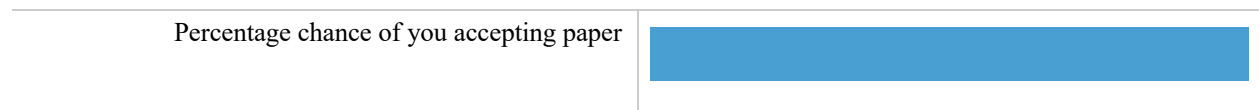
Please download and read this paper: Paper#[paper code]

Overall, the quality of the paper is

- Very bad (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- Very good (10)

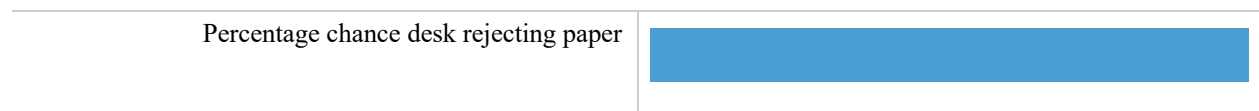
How likely would you be to accept this paper at a general economics conference (such as the Australian Conference of Economists)?

0 10 20 30 40 50 60 70 80 90 100



Imagine you were an editor of a general economics journal that is an A journal on the ABDC list. How likely would you be to desk reject this paper?

0 10 20 30 40 50 60 70 80 90 100



How likely is it that this paper will get published in an A or A* journal on to the ABDC list?

0 10 20 30 40 50 60 70 80 90 100

Percentage chance of this paper getting published in an A or A* journal

Overall, the quality of the writing is

- Very bad (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- Very good (10)

Thank you for evaluating the quality of 10 papers. What do you think this research project is about?

Would you like to receive a first draft of this research project once it is available?

Yes

No

Please write your email address so that I can send you your \$50 voucher

Do you have any feedback about the survey? Please let me know below.

A.3: Survey sent to writing experts

What is your gender?

- Male
- Female
- Other
- Prefer not to say

Is English your first language?

- Yes
- No

What is your job?

Is writing an important task in your job?

- Yes
- No

Is reading an important task in your job?

- Yes
- No

Thanks for agreeing to judge the **writing quality** of ten academic papers.

For this questionnaire, think of well-written papers as papers that are easy to read and easy to understand. Well-written papers convey complex information efficiently. They do this by presenting information well on many levels. For example, a well-written paper may have a structure which makes it easy to follow the content and it may contain sentences that are easy to read and comprehend.

Instructions I'll show you ten different pages — one for each paper. On top of each page is a link to a paper. Please download this paper and review it. Below the link to the paper are five questions. Please answer all of them. I suggest you take 8 minutes per paper. To be able to review a paper quickly, read the abstract and introduction, and then skim read the rest of the paper. This is the process many academics use when deciding if they want to accept a

paper for a conference.

[At this stage, the survey included 10 pages, each with a link to a different paper and the same 5 questions. For brevity, we only show one of these pages]

Please download and read this paper:

[paper code]

Overall, the quality of the writing is

- Very bad (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- Very good (10)

Please state how much you agree with the following statements.

The paper allows me to easily find the key messages.

Completely disagree (0)

1

2

3

4

5

6

7

8

9

Completely agree (10)

The paper is free of spelling and grammar mistakes.

Completely disagree (0)

1

2

3

4

5

6

7

8

9

Completely agree (10)

The paper is easy to read.

Completely disagree (0)

1

2

3

4

5

6

7

8

9

Completely agree (10)

The paper is concise.

- Completely disagree (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- Completely agree (10)

The paper is concise.

- Completely disagree (0)
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- Completely agree (10)

Thank you for evaluating the writing quality of 10 papers.

Would you like to receive a first draft of this research project once it is available?

- Yes
- No

Please write your email address so that I can send you your \$50 voucher

Do you have any feedback about the survey? Please let me know below.

Appendix B: Language Editing Guidelines for Experiment

This version [1 April, 2020]

An outline of the general approach the language editors will take

The goal is to edit the paper so that an expert who has 10 minutes to evaluate the paper will understand it more easily. We'll focus on improving the title, abstract, and introduction using these guidelines. For the rest of the paper, we'll focus on making the paper easier to skim read. We'll prioritise what we edit based on what (in our experience) will best help the reader easily understand the paper while skim reading. We'll edit for readers who can comfortably read academic documents, but who aren't familiar with the discipline of economics.

If the writing is very poor, we'll rewrite some parts instead of editing.

There are many good reasons we'll deviate from these guidelines: if it ain't broke, don't fix it. If we're unsure, we'll ask the writer.

To save time, we may outsource some tasks to the writers. For example, we might ask them to round all numbers in a table to 2 decimal points.

Heavy edit of the title, abstract, and introduction only

Editing or rewriting the abstract and introduction will likely take the most amount of time.

Improving the structure

We'll start by making sure the structure is clear.

The title should explain what the paper is about

We'll make sure the title is clear.

The abstract should contain all the necessary parts

We'll check the abstract is one paragraph that contains:

- the research question
- an explanation of how this question is answered
- the main findings.

The introduction should contain all the necessary parts

We'll edit the introduction so that it has one paragraph for each of the following parts:

- the motivation for the research
- what the paper does (this paragraph often starts with "In this paper,...")
- results
- the related literature (unless there is a separate literature section)
- contribution to the literature.

Avoid roadmap paragraphs

A good structure and informative section titles will do the trick in most cases.

No need to have a roadmap paragraph at the end of the introduction or a part in the beginning previewing what the writer will do. For example, paragraphs don't need to start with something like, "In this section we first explain model 1 and then explain model 2".

Signposting for the reader

We'll make sure the information flows well and is clear for the reader.

Remove road map phrases used to connect paragraphs

If possible, link to previous paragraphs by using "we" and what the writer and the reader saw/learned/explored.

- "As we have seen, children are getting fatter" is better than "As discussed in the previous paragraph, children are getting fatter."
- "Now that we have explored" is better than "the previous section analysed".

Keep track of what the reader knows

Most international readers will not know what KiwiSaver is. Writers can use the term, but need to explain it the first time they mention it.

Keep track of what the reader expects

Readers naturally form expectations, and it's important to be aware of them.

We can identify gaps in information, for example mentioning information sooner. Or, when the writer says "There are two theories..." and the reader expects to hear about both of them. Writers should give the reader what they expect with sentences that start with "The first theory..." and "The second theory..."

If the writer asks three questions, we'll check that they're answered in the order the writer posed them.

This doesn't mean that writers cannot surprise the reader

In 99% of the cases writers will want to fulfil the expectations. However, in 1% of the cases writers may want to break expectations for effect. For example, "I had a normal childhood growing up in Germany, full of anger and disappointment."

Keep paragraphs focused and clear

Make sure paragraphs are focused and only discuss one idea. For example, have separate paragraphs for describing the results and for discussing the related literature.

Basically, we want the writer to write focused paragraphs. For example, this paragraph is about our results, this paragraph summarises the contributions to the literature.

Make sentences clear

Make sure the subject of the sentence is short and concise

The secret to a clear and readable style is in the first five or six words of every sentence. At the beginning of every sentence, locate the reader in familiar territory. The writing needs to have a clear flow of logic that is easy for the reader to follow — don't frame information in a way that breaks the flow.

Make the subject of the sentence the “actor”

Find the actor of the sentence and the actions they perform. If the actors are not the subjects and the actions are not verbs, we’ll revise so that they are.

Choose verbs that describe the crucial actions of those actors

Bad example:

Our lack of knowledge about local conditions...

Good example:

We *knew* nothing about local conditions...

Revise abstract nouns into verbs expressing action

Bad example:

A need exists for greater candidate selection efficiency.

Good example:

We *must* select candidates more efficiently.

Keep a short distance between nouns and their accompanying verb

Bad example:

The knowledge that criminalisation of marijuana use can lead to a wide variety of other social ills, including an increased risk of addiction to more dangerous and expensive drugs such as heroin and cocaine, has prevented law makers...

Avoid breaking the reading flow by inserting clauses

For example: Jan, economist by trade, loves fishing.

We’ll edit this where we can and make a comment when it’s hard to do, or if we run out of time.

Make sure the subject–verb combination makes sense

For example

“The first specification controls for the pre-trend and assumes that the...”

Can specifications control or assume?

Words

Use simple, familiar words

Use “use” instead of “utilize”. Use “people” instead of “individuals”.

Delete unnecessary words or clauses

For example, in “We are the first to introduce a novel method”, there is no need to mention both “first” and “novel”.

In Section 2, we explain ~~to the reader~~ how our results are estimated.

Many introductory clauses that end with “that” can be deleted. Everything before the “that” should be deleted from a sentence.

~~It is usually the case that most good writers find that...~~

~~It should be noted that~~ writing is an art and a science.

Avoid nominalizations

Nominalizations are nouns derived from a verb or an adjective. As a general rule, avoid them.

Bad example:

We conducted an investigation on it.

Good example:

We investigated it.

Exceptions: Some nominalizations are useful

Some nominalizations are subjects which allow us to better link sentences into a more cohesive flow.

Example:

The president argued that the Federal Reserve should lower its interest rate. His **arguments** all depended on a single unproven claim.

Some nominalizations allow us to make sentences more compact by naming what would be the object of its verb.

Example:

“I do not understand his intention” is more compact than “I do not understand what he means”

Some nominalizations refer to an often-repeated concept. Rather than repeatedly spelling out what they mean, these abstractions become virtual actors.

Example:

Few issues have so divided American as **abortion on demand**.

Avoid abbreviations and acronyms

Use them only if they help the reader and choose the ones that sound good. OECD is fine, but use Facebook instead of FB. Write New Zealand, not NZ.

Avoid negations

Avoid negations when the statement you are negating is plausible.

Example:

“A whale is not a fish” is difficult to understand because many believe whales are fishes.

“A whale is a mammal” is easier to understand.

If you have to negate a plausible statement, do it in two steps.

Example:

“You might think a whale is a fish. But no, it is a mammal.”

Remove hedging

We’ll remove any hedging statements that seem unnecessary. Writers don’t need to always say “all else equal”, “fairly”, “I would argue”. However, sometimes they need to qualify the statements to avoid people getting it wrong.

Avoid naked this or that in the beginning of the sentence

We’ll add more information where needed, such as “this regression shows” instead of “this shows”.

Remove noun strings

Where possible we'll break up strings of words (usually used to name something) that create ambiguity for the reader.

Light structural edit to make the paper easier to skim read

We'll do a light edit of the rest of the paper so that skim readers can find information more easily.

Use informative section titles

We'll add information to section titles and keep them short and concise. For example, "The credit market in New Zealand" is better than "Background".

We'll also make sure the section titles clearly stand out from the body text, using **bold font**.

Use self-explanatory titles of tables and figures

"The effect of peer gender on educational outcomes" is better than "Main Results".

Include tables and figures in the body of the document and not at the end

Doing this allows the reader to see the table and figure the text describes.

Make sure the results section starts with the main results

Many readers will skim read the paper. Showing the main results at the beginning of the results section will help them find the most important findings.

The main results are typically the results that answer the research question.

Optional: Make sure the conclusion answers the key question posed in the introduction

If we have time, we'll make sure the conclusion restates the key results and shows us how they contribute to the solving of the larger problem shown in the introduction.

Light word-edit of the paper (if we have time)

We'll focus on improving the writing in the rest of the paper to make it easier to skim read. The time we have left will determine how much of this we can do. For each document, we'll make the changes that make the document easiest to skim read. For example, if a document uses lots of ambiguous noun-strings we're more likely to prioritise editing these.

Changes we'll make to the text

Most of these changes are already covered in detail in the above sections.

- Fix problems with long sentences (with StyleWriter)
- Fix problems with passive voice (with StyleWriter)
- Fix problems with nominalizations (with StyleWriter)
- Untangle noun strings
- Delete unnecessary words and clauses
- Use more personal pronouns (like "we" and "our") where possible

Check the paper is in the present tense

We'll make changes to the tense where possible. If we need to make a lot of changes to the paper, we'll make a comment to the writer where sections need the tense made consistent.

We'll quickly proofread each paper after the student has looked at it

We'll fix any proofreading errors we notice as we edit. When we receive each edited paper back from the student we'll do a quick proofread and fix any errors we notice.

Specifically we'll look at spelling, punctuation and grammar. We'll make sure each paper uses a consistent style.

Check a consistent style is used throughout

If an aspect of style is inconsistent in a paper, follow the advice in the Write Style Guide. The sorts of things we'll check for are that:

- the paper is in present tense
- text is justified
- line spacing is 1.5
- headings and subheadings clearly stand out from the text: for example, using **bold** font
- words and symbols are consistent: % or percent, Feld and Salamanca (2016) or Feld & Salamanca (2016)
- numbers are consistent: 10 or ten
- commas separate a thousand: 1,345

Appendix C – Details on assignment of papers to experts

To avoid any systematic matching of experts to judge particular kinds of papers, we randomly assigned papers to blocks and assigned experts to rate papers in a given block in the order in which they agreed to participate in the experiment. We describe our approach in more detail below.

C1: Assignment of papers to paper blocks for economists

We created two groups of papers: “micro papers” and “macro papers”. The micro papers group consist of 19 microeconomics papers and one macro paper which can easily be judged by microeconomists. The macro papers group consist of 10 macro papers.

We randomly assigned each paper within a group one “paper number”, so that papers in the micro group were assigned numbers 1-20 and papers in the macro group were assigned numbers 21-30.

After that, we created the following six paper blocks, each consisting of 5 original versions and 5 edited versions.

1. Papers 1-10 (micro), papers 1-5 original version, papers 6-10 edited version
2. Papers 1-10 (micro), papers 1-5 edited version, papers 6-10 original version
3. Papers 11-20 (micro), papers 11-15 original version, papers 16-20 edited version
4. Papers 11-20 (micro), papers 11-15 edited version, papers 16-20 original version
5. Papers 21-30 (macro), papers 21-25 original version, papers 26-30 edited version
6. Papers 21-30 (macro), papers 21-25 edited version, papers 26-30 original version

C2: Assignment of papers to paper blocks for witting experts

We randomly assigned each paper one “writing paper number”. This was a different randomization, the “writing paper numbers” is not the same as the “paper numbers”.

After that, we create the following six blocks, each block consisting of 5 original versions and 5 edited versions.

1. Papers 1-10, papers 1-5 original version, papers 6-10 edited version
2. Papers 1-10, papers 1-5 edited version, papers 6-10 original version
3. Papers 11-20, papers 11-15 original version, papers 16-20 edited version
4. Papers 11-20, papers 11-15 edited version, papers 16-20 original version
5. Papers 21-30, papers 21-25 original version, papers 26-30 edited version
6. Papers 21-30, papers 21-25 edited version, papers 26-30 original version

C3: Assignment of economists to paper blocks

We created two lists of academic economists from Australian universities and research institutes: one for micro economists and one for macro economists. We invited economists in both groups via email to evaluate 10 academic papers in their discipline.

Those who agreed to evaluate ten papers were assigned to a paper block, and were sent a survey containing links to the relevant papers. More specifically, micro-economists were assigned paper blocks 1-4 (which contain micro papers) and macro-economists were assigned paper blocks 5-6 (which contain macro papers).

The order of the paper block assignment was determined by the order the economists agreed to participate. For example, the first micro-economist was assigned to paper block 1, the second micro-economist was assigned to paper block 2, etc. Similarly, the first macro-economist was assigned to paper block 5 and the second macro-economist was assigned paper block 6.

We deviated from this assignment-procedure to avoid that economists from the same institution are asked to judge different versions of the same paper within a short time. Economists from the same institution are more likely to talk to each other about the task and therefore might realize that we have included different versions of the same paper in the experiment. This might have

raised suspicions, which we wanted to avoid. We solved this problem by swapping the assignment to paper blocks with economists from other institution.

C4: Assignment of writing experts to paper blocks

We invited writing experts to evaluate 10 academic papers. Once they agreed, we sent them a survey containing links to the relevant papers.

The order of the paper block assignment will be determined by the order the writing experts agree to participate. The expert who first agrees to participate was assigned paper block 1, the second expert was assigned to paper block 2, etc.