CMPU 366 · Natural Language Processing

Project Peer Review and Getting Started

19 November 2026



Today

Peer review exercise

Getting started & data check work

Peer review

onlinelibrary.wiley.com/doi/epdf/10.1111/ssqu.12824

Wiley Online Library

Page 1 / 5

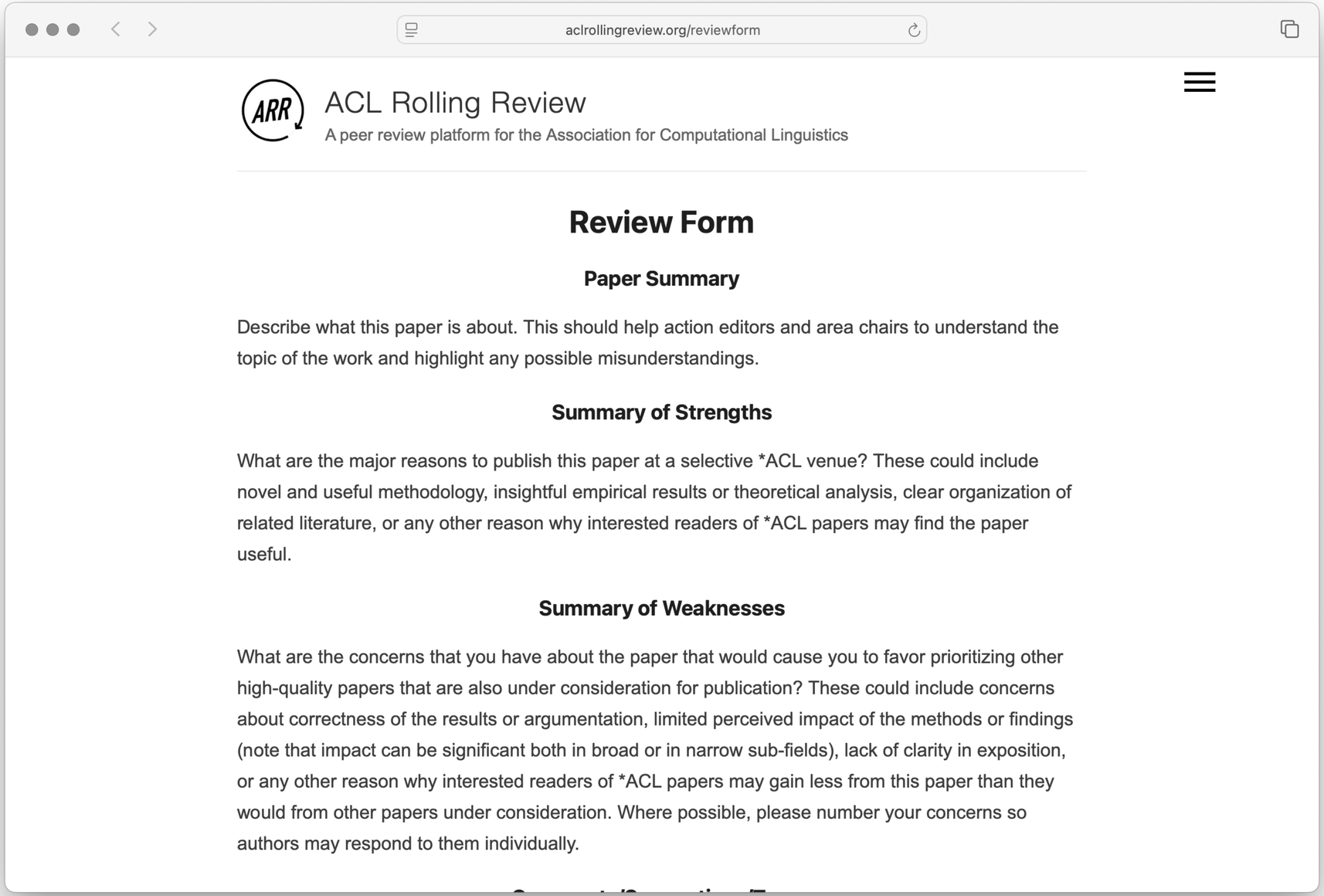
Dear Reviewer 2: Go F' Yourself

David A. M. Peterson, *Iowa State University*

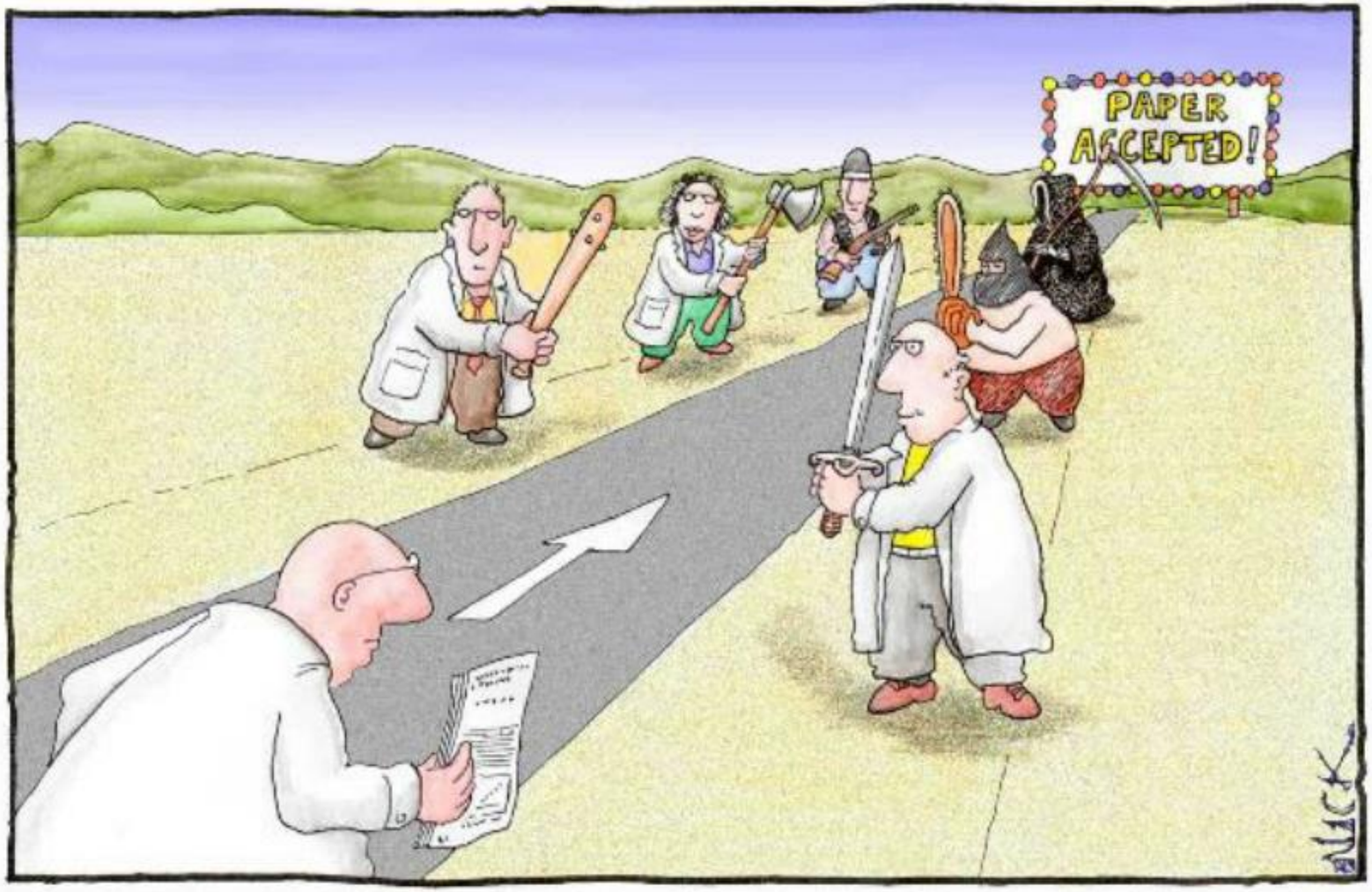
Objectives. The objective of this study was to empirically test the wide belief that Reviewer #2 is a uniquely poor reviewer. *Methods.* The test involved analyzing the reviewer database from *Political Behavior*. There are two main tests. First, the reviewer's categorical evaluation of the manuscript was compared by reviewer number. Second, the data were analyzed to test if Reviewer #2 was disproportionately likely to be more than one category below the mean of the other reviewers of the manuscript. *Results.* There is no evidence that Reviewer #2 is either more negative about the manuscript or out of line with the other reviewers. There is, however, evidence that Reviewer #3 is more likely to be more than one category below the other reviewers. *Conclusions.* Reviewer #2 is not the problem. Reviewer #3 is. In fact, he is such a bad actor that he even gets the unwitting Reviewer #2 blamed for his bad behavior.

Anyone who has ever submitted a paper to a peer-reviewed outlet knows the reviewers can, occasionally, be unpleasant. While rejection always stings, the belief that a reviewer has either completely missed the point of the manuscript, been overtly hostile in his or her review, or simply held the author to an impossible standard is vexing. The source of this frustration has seemingly become personified in the identity of a single person—Reviewer 2. He (and it is always assumed to be a he) is embodiment of all that we hate about other scholars. Reviewer 2 is dismissive of other people's work, lazy, belligerent, and smug.

The purpose of this article is to test a very specific claim about Reviewer 2: he is the reviewer who holds us back. Using the database of reviewer responses from four years of *Political Behavior*, I empirically test if reviewers who are assigned number 2 are systematically more negative and more likely to be out of line from the other reviews a manuscript received. I assess this hypothesis in two ways. First, I compare the ordinal score each reviewer gives the manuscript. Second, I develop an original measure of "being *Reviewer 2*." In this specific case, being *Reviewer 2* is defined as giving a score to the manuscript that is more than one category below the mean ranking of the other two reviewers. The results



aclrollingreview.org/reviewform



Most scientists regarded the new streamlined peer-review process as 'quite an improvement.'

Writing a good literature review

We write a literature review (or “Related Work” section) to:

- Acquaint the reader with the field of study
- Establish relevance
- Reveal literature gaps

Good literature reviews

- Organize paragraphs around ideas
- Describe studies in your own voice
- Focus on a research question
- Contain clear topical sentences

A literature review needs to be clearly organized.

This is a bit like a jig-saw puzzle – when you read papers, you're finding the pieces, and then you need to figure out how they fit together.

If you have trouble finding some pieces, that may suggest that there is an important gap in the literature, which can be a great opportunity!

Common issue: No synthesis!

Do not:

- Write each paragraph about an individual study
- Try to report all the details from the paper

Rather, you're describing the literature to set the stage for your own work.

- Have people looked at this problem?
- This data?
- What did they try?
- How did it work out?

Poor literature reviews contain

- Lots of in-depth details about the studies
- Too much jargon
- Topic sentences that discuss specific authors/studies
- References without citations
- Very long paragraphs

Peer review exercise

Pair off with someone near you (who isn't in your group).

If you have a print copy of your literature review, give it to the other person.

If you have an electronic copy, share that.

Write your answers in a Google Doc shared with the author – before you talk to them.

Tonight, you'll paste them into Gradescope.

Before you talk to the author, write down:

1. In one sentence, the research question is…
2. The closest existing work is … and this project will differ from it by …
3. The evidence that will answer the question is … (What data? What metric? Compared to what?)
4. How is the related work organized (chronologically, by problem, by method, …)?
   Does any paragraph read like a book report on a single paper?
5. Where did you get most confused, or most want to know more?

Share your answers with the other person.

Where is their version different from what you meant? That's what to revise!

Write down: "The biggest thing I'll revise is…"

Getting started

Iterative development

Get something working end to end, e.g.,

data → *model* → *evaluation*,

then make it better.

Start with the simplest version of each step!

Baselines

- Tell you whether your numbers mean anything, e.g.,
  - Is 90% accuracy good?
  - Not if 92% of the examples are in one class!
- Give you a result to report even if the fancy approach doesn't work out
- Show that the rest of your pipeline works

Example baselines:

Classification:

Majority class

Bag-of-words logistic regression

Language modeling:

n-gram model (Asmt 2!)

Summarization:

The first few sentences of the document

Using an LLM:

Zero-shot with a plain prompt (before any prompt engineering or fine-tuning)

Analysis projects:

Simple counts and frequencies

Aim to have a baseline running by the end of this week.

Or, at the very latest, before Thanksgiving.

For the rest of today:

Load your data

Split off a test set and set it aside, if your data doesn't already come split (train/dev/test).

Look at your training data. How many examples? What's the distribution of labels or lengths?

Read 20 random training examples.

Run the simplest baseline you can, and evaluate it on your dev set.

No data yet? That's your top job today – talk to me about it.

On Gradescope: “Project work: Peer review and data check”

- Your review (paste from your Google Doc)
- The biggest thing you’ll revise
- Where your data and baseline stand

Due tonight (11:59 p.m.)

Acknowledgments

The class incorporates material from:

Ruth Szpunar

