

Reviewing Natural Language Processing Research

Kevin B. Cohen

Computational Bioscience Program
University of Colorado, USA

kevin.cohen@gmail.com

Margot Mieskes

University of Applied Sciences, Darmstadt
Germany

margot.mieskes@h-da.de

Karën Fort

Sorbonne Université, EA STIH, Paris
LORIA, Nancy

France

karen.fort@sorbonne-universite.fr

Aurélie Névéol

LIMSI, CNRS
Université Paris-Saclay

France

neveol@limsi.fr

1 Tutorial Content

This tutorial will cover the goals, processes, and evaluation of reviewing research in natural language processing. As has been pointed out for years by leading figures in our community (Webber, 2007), researchers in the ACL community face a heavy—and growing—reviewing burden. Initiatives to lower this burden have been discussed at the recent ACL general assembly in Florence (ACL 2019)¹. Simultaneously, notable “false negatives”—rejection by our conferences of work that was later shown to be tremendously important after acceptance by *other* conferences (Church, 2005)—has raised awareness of the fact that our reviewing practices leave something to be desired... and we do not often talk about “false positives” with respect to conference papers, but conversations in the hallways at *ACL meetings suggest that we have a publication bias towards papers that report high performance, with perhaps not much else of interest in them (Manning, 2015).

It need not be this way. There is good reason to think that reviewing is a learnable (and teachable) skill (Basford, 1990; Paice, 2001; Benos et al., 2003; Koike et al., 2009; Shukla, 2010; Tandon, 2014; Spyns and Vidal, 2015; Stahel and Moore, 2016; Kohnen, 2017; McFadden et al., 2017; Hill, 2018). To address the issues raised above, we propose this tutorial on reviewing natural language processing research, focusing on conference submissions and various review forms used in the NLP community. The extended part also covers journal submissions.

As the demand for reviewing grows, so must the pool of reviewers. As the survey presented by Graham Neubig at the 2019 ACL showed, a

considerable number of reviewers are junior researchers, who might lack the experience and expertise necessary for high-quality reviews. A tutorial on this topic might increase reviewers’ confidence, as well as the quality of the reviews. Given the importance of conferences in NLP, the reviewing standards should be as high as with journals in other fields.

2 Timetable

Table 1 shows an outline of the content discussed during the tutorial. Apart from a **general introduction** to the topic of peer reviewing and its role in the publishing circle, we will go into details on **reviewing for *ACL-venues**. All sections will include **exercises** and **practical** examples to get a better grasp for individual elements mentioned during the theoretical input. We will also take a look at **problems** with respect to peer reviewing and specific peer **reviewing models**, such as double-blind reviewing, which is the primary mode in *ACL-publication venues vs. single-blind and open reviewing. The **case study** will look at an actual example paper including reviews for that example.

3 Suggested Reading List

- John Bohannon. 2013. *Who’s afraid of peer review?* *Science*, 342(6154):60–65
- Kenneth Church. 2005. *Last words: Reviewing the reviewers.* *Computational Linguistics*, 31(4):575–578
- Button K. S., Bal L., Clark A., and Shipley T. 2016. *Preventing the ends from justifying the means: withholding results to address publication bias in peer-review.* *BMC Psychol.*, 4(1)
- Leif Engqvist and Joachim Frommen. 2008. *Double-blind peer review and gender publication bias.* *Animal Behaviour*, 76:e1e2

¹<http://www.livecongress.it/aol/indexSA.php?id=E2EAED7D&ticket=>

Section	Content
1	Role of peer review in scientific publishing
2	Approaches to reviewing and NLP-specific issues
3	Section-specific criteria (Materials & Methods, Results, etc.)
5	Ethics of reviewing
6	Case study: a paper to review

Table 1: Outline of the Tutorial.

- Michael J. Mahoney. 1977. [Publication prejudices: An experimental study of confirmatory bias in the peer review system](#). *Cognitive Therapy and Research*, 1(2):161–175
- Mark Peplow. 2014. [Peer review reviewed](#). *Nature*
- Mark Steedman. 2008. [Last words: On becoming a discipline](#). *Computational Linguistics*, 34(1):137–144
- Bonnie Webber. 2007. [Breaking news: Changing attitudes and practices](#). *Computational Linguistics*, 33(4):607–611
- Christine Weners. 1997. [Nepotism and sexism in peer-review](#). *Nature*, 387

4 Presenters (in alphabetical order)

Kevin Bretonnel Cohen has written, overseen, and received hundreds of reviews in his capacity as deputy editor-in-chief of a biomedical informatics journal, associate editor of five natural language processing or bioinformatics journals, special issue editor, workshop organizer, and author of 100+ publications in computational linguistics and natural language processing. His forthcoming book *Writing about data science research: With examples from machine and natural language processing* includes coverage of a number of aspects of the reviewing process. His current research focuses on issues of reproducibility.

Karën Fort is an associate professor at Sorbonne Université. Besides being a reviewer for most major NLP conferences, she has been editor in chief for a *Traitement automatique des langues* journal special issue on ethics and acted as Area Chair for ACL in 2017 and 2018 (as senior AC). Her main research interests are ethics, and the construction of language resources for natural language processing. She co-authored the report on the EMNLP reviewer survey (Névéol et al., 2017).

Margot Mieskes is a professor at the Darmstadt University of Applied Sciences and as such has a lot experience teaching, also in culturally diverse settings, which are prevalent in German Universities of Applied Sciences. Additionally, she has

written and received a number of reviews in conferences as well as journals. She is a member of the ACL Professional Conduct Committee and an active member of the Widening NLP efforts. Her research interests are in summarization and summarization evaluation, replicability, repeatability and transparency of NLP experiments in general. **Aurélie Névéol** is a permanent researcher at LIMSI CNRS and Université Paris Saclay. She has been involved in reviewing natural language processing papers at many stages of the reviewing process, including: reviewer, associate editor for three journals, area chair for *ACL and bioinformatics conferences, workshop organizer. Her research focuses on biomedical natural language processing as well as ethics issues in NLP research. She co-authored the report on EMNLP reviewer survey (Névéol et al., 2017).

References

- P Basford. 1990. How to... review an article. *Nursing times*, 86(40):61–61.
- Dale J Benos, Kevin L Kirk, and John E Hall. 2003. How to review a paper. *Advances in physiology education*, 27(2):47–52.
- John Bohannon. 2013. [Who’s afraid of peer review?](#) *Science*, 342(6154):60–65.
- Kenneth Church. 2005. [Last words: Reviewing the reviewers.](#) *Computational Linguistics*, 31(4):575–578.
- Leif Engqvist and Joachim Frommen. 2008. [Double-blind peer review and gender publication bias.](#) *Animal Behaviour*, 76:e1e2.
- Michael D Hill. 2018. How to review a clinical research paper. *Stroke*, 49(5):e204–e206.
- Thomas Kohnen. 2017. How to write a good peer review. *Journal of Cataract & Refractive Surgery*, 43(10):1243–1244.
- Kaoru Koike, Luca Ansaloni, Fausto Catena, and Ernest E Moore. 2009. WJES: how to review a clinical paper. *World Journal of Emergency Surgery*, 4(1):8.
- Michael J. Mahoney. 1977. [Publication prejudices: An experimental study of confirmatory bias in the peer review system.](#) *Cognitive Therapy and Research*, 1(2):161–175.
- Christopher D Manning. 2015. Computational linguistics and deep learning. *Computational Linguistics*, 41(4):701–707.
- David McFadden, Scott LeMaire, Michael Sarr, and Kevin Behrns. 2017. How to review a paper: Suggestions from the editors of surgery and the journal of surgical research. *Surgery*, 162(1):1–6.
- Aurélie Névéol, Karën Fort, and Rebecca Hwa. 2017. [Report on EMNLP Reviewer Survey.](#) Technical report, Association for Computational Linguistics.
- Elisabeth Paice. 2001. How to write a peer review. *Hospital Medicine*, 62(3):172–175.
- Mark Peplow. 2014. [Peer review reviewed.](#) *Nature*.
- Button K. S., Bal L., Clark A., and Shipley T. 2016. [Preventing the ends from justifying the means: withholding results to address publication bias in peer-review.](#) *BMC Psychol.*, 4(1).
- Satish K Shukla. 2010. How to review an article. *Indian Journal of Surgery*, 72(2):93–96.
- Peter Spyns and María-Esther Vidal. 2015. *Scientific Peer Reviewing: Practical Hints and Best Practices.* Springer.
- Philip F Stahel and Ernest E Moore. 2016. How to review a surgical paper: a guide for junior referees. *BMC medicine*, 14(1):29.
- Mark Steedman. 2008. [Last words: On becoming a discipline.](#) *Computational Linguistics*, 34(1):137–144.
- Rajiv Tandon. 2014. How to review a scientific paper. *Asian journal of psychiatry*, 11:124–127.
- Bonnie Webber. 2007. [Breaking news: Changing attitudes and practices.](#) *Computational Linguistics*, 33(4):607–611.
- Christine Wenners. 1997. [Nepotism and sexism in peer-review.](#) *Nature*, 387.