

# Open Knowledge for a Sustainable Future: Research, Ethics, and Wikipedia

Week 7 — Feedback on Academic Articles - Peer review and revision for publication

Francis Bond (**Academic**)   Pavel Bednařík (Wiki)

Palacký University Olomouc   |   Wikimedia ČR

2025-12-06

# Roadmap

- 1 Feedback and Revision
- 2 Peer Review, Trust, and Metrics
- 3 Problems: Bias & Inequality, (un)Reproducibility
- 4 Peer Review & Revision in Practice
- 5 References

# Writing is a process — feedback is essential

- First drafts are rarely clear, complete, or well-argued
- Feedback (from peers, tutors, self-reflection) helps identify:
  - ▶ unclear claims
  - ▶ logical gaps
  - ▶ missing evidence
  - ▶ stylistic issues
- Improving writing increases clarity, persuasiveness, and ethical responsibility — especially important when dealing with data, culture, or community topics

# What feedback gives to the writer

- Another reader's perspective — reveals assumptions or blind spots
- Fresh eyes on structure, coherence, and flow: does argument build logically?
- Opportunity to strengthen evidence and address counter-arguments
- Chance to improve style, tone, and accessibility depending on audience

# Feedback benefits the critical reader too

- Readers become more critical and informed — able to assess evidence, question assumptions, and spot weak arguments
- Improves collective standards: better writing leads to better discourse
- Feeds back into research culture: clarity and transparency help others replicate, critique, or build further work

# Feedback is ethical, not just cosmetic

- Writing about people, cultures, data — carries responsibility. Feedback helps catch problematic framing, bias, or unintended harm
- Encourages accountability: facts must be checked, sources verified, sensitive topics handled with care
- Discipline-wide benefit: careful writing builds trust, respect, and legitimacy for scholarship

# Roadmap

- 1 Feedback and Revision
- 2 Peer Review, Trust, and Metrics
  - Feedback Beyond Academia
  - Reviewing at the Association for Computational Linguistics
- 3 Problems: Bias & Inequality, (un)Reproducibility
- 4 Peer Review & Revision in Practice
- 5 References

# What peer review is

- Peer review is the process by which experts evaluate a manuscript before publication
- Its aims are to:
  - ▶ Check the **soundness** of methods and argument
  - ▶ Improve clarity, structure, and contribution
  - ▶ Build **trust** in published research
- Reviewers act as gatekeepers and mentors — ideally improving the quality of science and scholarship

*Peer review is not about perfection; it is about accountability and conversation.*



# Types of peer review

- There are several models of peer review, each balancing **transparency, fairness, and accountability** in different ways

**Single-blind:** Reviewers know the author's identity, but authors do not know the reviewers

- Common in science and social science
- Advantage: reviewers can assess expertise and potential conflicts
- Risk: unconscious bias (e.g., by gender, affiliation, language, or region)

**Double-blind:** Neither authors nor reviewers know each other's identity

- Aims to reduce bias and level the playing field
- Often used in humanities and linguistics conferences
- Risk: anonymity may limit accountability; reviewers can sometimes guess authors

**Open review:** Both sides know identities, and sometimes reviews are published

- Promotes transparency and recognition of reviewing labour
- Encourages constructive tone, but may discourage strong criticism
- Used in some open-science and data journals

# Peer review: strengths and weaknesses

## Strengths

- Enhances rigour and precision
- Provides expert feedback before publication
- Establishes a shared standard of credibility
- Encourages dialogue within a discipline

## Weaknesses

- Can be slow, inconsistent, or biased
- Overburdens a small reviewer pool
- Replication and transparency not always checked
- Negative or null results often unpublished ("file-drawer" effect)

*Peer review is essential but imperfect — a social process as much as a scientific one.*

# Evolving models of peer review

- **Post-publication review:**

- ▶ Articles or datasets are published first, then openly discussed
- ▶ Examples: [F1000Research](#), [PubPeer](#)

- **Community review:**

- ▶ Large-scale collaborative commenting before publication (e.g., open preprints, [GitHub issues](#), [Open Review at LangSci Press](#))
- ▶ Builds transparency and collective expertise

- **Preprint feedback:**

- ▶ Early sharing (e.g., [arXiv](#), [LingBuzz](#)) invites informal peer feedback before formal submission
- ▶ Speeds dissemination and reduces "scoop" anxiety

- These models shift peer review from an evaluation to a conversation

# Ethical responsibilities of peer reviewers

- Reviewers are entrusted with **privileged access** to unpublished work
- Ethical responsibilities include:
  - ▶ **Confidentiality**: do not share, quote, or use ideas from the manuscript
  - ▶ **Impartiality**: declare conflicts of interest; avoid personal or institutional bias
  - ▶ **Constructive criticism**: focus on improving the work, not attacking the author
  - ▶ **Timeliness**: return reviews promptly; long delays can harm authors' careers
  - ▶ **Accuracy**: evaluate methods, evidence, logic, and citations fairly
  - ▶ **Respect**: use professional tone; assume good intent; avoid dismissive language
- As Rockwell (2005) emphasises, access to unpublished research creates a **power imbalance**: reviewers must not exploit this advantage

*A good reviewer protects confidentiality, improves clarity, and strengthens trust in scholarship.*

# Feedback in industry: the practice of "giving notes"

- In many industries (film, publishing, games, advertising), feedback is formalised as "**notes**" (Catmull & Wallace, 2014)
- Notes come from editors, producers, managers, clients — often in a hierarchical chain
- The goal is similar to academic review:
  - ▶ improve clarity, impact, and quality
  - ▶ ensure the work meets its purpose and audience
- But constraints differ:
  - ▶ strict deadlines, budgets, brand guidelines, legal concerns
  - ▶ competing visions between writer/creator, editor, and client

Video: [Ed Catmull on the Pixar Braintrust](#) (Stanford, 8 min)

# How industry notes differ from academic peer review

- **Hierarchical:** notes often come from someone with authority over the project (editor, producer, manager)
- **Goal-driven:** feedback must serve the needs of the project, brand, or product — not the advancement of knowledge
- **Collaborative but constrained:** creators negotiate changes, but ultimate decisions may be external
- **Iterative cycles:** multiple rounds of notes → revisions → more notes → final approval
- **Success = alignment** with project goals, not independent scholarly rigour

# Best practices from industry feedback culture

- **Clarity first:** Notes should define the problem before proposing solutions (Pixar rule)
- **Respect creative intent:** Improve the work without imposing a personal style
- **Prioritise high-impact changes:** Not all notes can be addressed; choose the ones that serve the project
- **Negotiate professionally:** Creators can ask for clarification and push back when necessary
- **Shared responsibility:** Good notes identify issues; creators choose how to solve them

# What students can learn from industry feedback

- You will encounter feedback in many forms: editorial, managerial, client-oriented, or technical
- Key transferable skills:
  - ▶ receiving feedback without defensiveness
  - ▶ interpreting conflicting goals
  - ▶ deciding which suggestions to adopt, adapt, or reject
  - ▶ communicating clearly about revisions
- Ethical themes remain:
  - ▶ respect, clarity, collaboration, responsibility

*Good feedback — academic or industrial — improves the work and strengthens the team.*



# Feedback cultures beyond academic peer review

- **Journalism:** Editors and fact-checkers verify claims against primary sources before publication — different from academic review, which often assumes sources are cited correctly
- **Wikipedia:** Articles pass through multiple stages of community review; good/featured article criteria require verifiability and neutral point of view
- **Code review:** In software development, peers review code for bugs, style, and maintainability — feedback is typically public and tied to version control (e.g., GitHub pull requests)

*Feedback aims to improve quality through structured critique before the results are made public*

# Peer review in Computational Linguistics

- Peer review is so central to scholarship that whole communities invest in improving it
- The Association for Computational Linguistics (ACL) publishes:
  - ▶ reviewer training guides (ACL 2023 Program Committee, 2023)
  - ▶ best-practice recommendations (ACL Rolling Review, 2025)
  - ▶ annual reports analysing what works and what fails (ACL Committee on Peer Review, 2023)
- These offer a model for responsible, constructive, ethical reviewing
- ACL 2025 followed the following structure:
  - ▶ 8,300 submissions
  - ▶ 1,700 accepted as full, 1,392 as findings
  - ▶ 169 Senior Area Chairs, 1,937 Area Chairs, 11,720 reviewers!
  - ▶ Each paper reviewed by three reviewers, the authors respond, the reviewers update their reviews, the area chairs suggest which papers are accepted, the senior area chairs make the final decisions, ...

# Accept reviews only when appropriate (ACL guidelines)

- Before agreeing to review, check:
  - ▶ **Expertise:** Can you fairly assess the methods and claims?
  - ▶ **Conflicts of interest:** shared institution, collaborators, competitors
  - ▶ **Capacity:** Do you actually have time to read carefully?
- Accepting when you are unqualified or overloaded leads to unfair or rushed reviews
- Slightly undermined by the fact that if you submit a paper you must agree to review at least 5 others, ...
  - ▶ ACL has a 15-20% acceptance rate, so for one paper accepted 4 are rejected.

# Read carefully and generously (ACL guidelines)

- First read: understand the research question and contribution
- Second read: evaluate evidence, structure, and clarity
- Look for what the author is *trying* to do
- Avoid snap judgments based on style, formatting, or first impressions

# Write constructive, actionable feedback (ACL guidelines)

- Every review should include both:
  - ▶ **Strengths** — what works well
  - ▶ **Weaknesses** — what needs improvement
- Make criticisms specific:
  - ▶ point to a sentence or section
  - ▶ explain why it is unclear
  - ▶ suggest how to fix it
- Vague comments (*paper is confusing*) are unhelpful

# Check evidence, ethics, and reproducibility (ACL guidelines)

- A reviewer is responsible for more than novelty
- Consider:
  - ▶ Are claims supported by evidence?
  - ▶ Are methods and data described clearly?
  - ▶ Are ethical implications acknowledged?
  - ▶ Can results be verified or reproduced?
- Responsible peer review strengthens the entire field

# Professionalism and respect (ACL guidelines)

- Use a constructive and collegial tone
- Avoid dismissive language:
  - ▶ "This is trivial"
  - ▶ "obviously wrong"
  - ▶ "the authors don't understand"
- Remember: reviewers and authors are colleagues
- Treat the paper the way you would want yours to be treated

# Reviewer power and responsibility (ACL COPR)

- Reviewers gain **access to unpublished work** before the community
- This creates responsibility:
  - ▶ Never use unpublished ideas in your own work
  - ▶ Never delay a review to gain priority
  - ▶ Never punish papers that challenge your favourite theory
- Peer review is a trust-based system; unethical reviewing harms entire fields



# What this means for you (student reviewers)

- The same principles apply in the classroom:
  - ▶ read generously
  - ▶ comment specifically
  - ▶ focus on clarity, evidence, and logic
  - ▶ be respectful and supportive
- Practising good peer review improves your own writing:
  - ▶ you learn what makes an argument work
  - ▶ you see how readers interpret drafts
  - ▶ you build editorial judgment

Based on ACL reviewer training materials (ACL 2023 Program Committee, 2023; ACL Committee on Peer Review, 2023; ACL Rolling Review, 2025).

# ACL-inspired peer review checklist

- Am I qualified and conflict-free?
- Have I read the paper carefully (twice)?
- Did I identify strengths and weaknesses?
- Are my comments specific and actionable?
- Did I evaluate ethics, evidence, and reproducibility?
- Is my tone respectful and professional?
- Have I upheld confidentiality and integrity?

# Roadmap

- 1 Feedback and Revision
- 2 Peer Review, Trust, and Metrics
- 3 Problems: Bias & Inequality, (un)Reproducibility
  - Punktoza
  - Bias and Inequality in Peer Review
  - The Reproducibility Crisis
  - AI and Peer Review
- 4 Peer Review & Revision in Practice
- 5 References

# Publish or Perish

- Academics are under pressure to publish
- This is a major part of how they are evaluated
- This should lead to more, higher quality publications
- But can lead to many more, lower quality publications
  - ▶ Or even fraud!
- And means we have to review more and more publications

# Metrics: why they were introduced

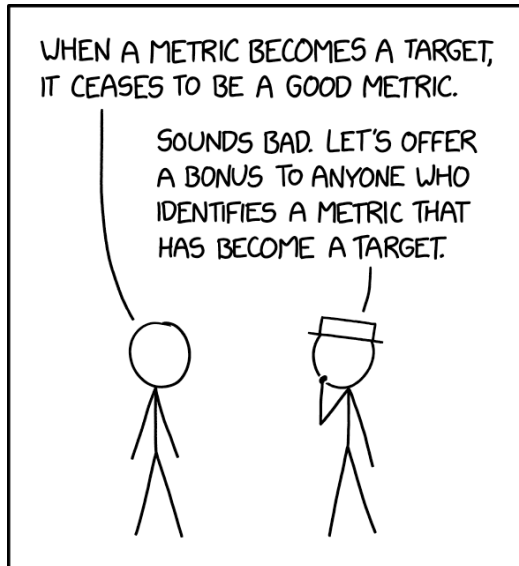
- Growing research ecosystems needed **standardised indicators** to compare:
  - ▶ journals (Impact Factor)
  - ▶ researchers (h-index)
  - ▶ institutions (citation counts, productivity)
- **Impact Factor (IF)**: average number of citations to articles in a journal over a two-year period
  - ▶ Designed to help libraries choose journals and to signal visibility
- **h-index**: a researcher has index  $h$  if they have  $h$  papers each cited at least  $h$  times
  - ▶ Intended to combine productivity + influence in a single number
- Metrics aimed to make evaluation **transparent, comparable, and data-driven**
- They were introduced to help — not to replace — expert judgment

# Metrics: when measures become targets

- When hiring, promotion, or funding is tied to metrics:
  - ▶ Researchers optimise for what is counted, not what is meaningful
  - ▶ Citation counts become currency; novelty outranks replication or rigour
- **Goodhart's Law:**  
*When a measure becomes a target, it ceases to be a good measure.*
  - ▶ IF and h-index lose validity once people actively game them
- **Punktoza ("points fever"):** publishing strategically to accumulate evaluation points rather than to advance knowledge (Kulczycki, 2017a, 2017b)
  - ▶ Encourages salami-slicing, low-risk incremental work, quantity over substance
  - ▶ A systemic response to metric-driven national evaluation frameworks
- Metrics intended to guide evaluation end up **shaping behaviour**, sometimes harming scholarship

*Quantifying trust can erode it when numbers replace judgment.*

# Goodhart's Law [xkcd 2899](#)



# Peer review is not neutral — known biases

- Peer review should be meritocratic — but social biases can shape outcomes
- Empirical studies document systematic bias based on gender, institutional prestige, and affiliation
- This can reproduce inequalities in academic visibility and access



# Gender bias and underrepresentation

- A large study of 1.7 million authors and 740,000 referees found that manuscripts by women were not universally penalized — yet women remain under-represented among peer reviewers and editors (Helmer et al., 2017)
- Another review finds that women are less likely to appear as first/last authors, especially in high-impact journals, and are underrepresented as reviewers/editors (Murray et al., 2020)
- In grant-review contexts, women have also been shown to receive lower scores despite similar qualifications (Wennerås & Wold, 1997)

# Institutional – affiliation and prestige bias

- Recent evidence shows "affiliation bias": papers from prestigious or well-known institutions are more likely to get favourable review or acceptance (Tomkins et al., 2017)
- This can disadvantage authors from lesser-known institutions, even when the manuscript quality is comparable
- Result: citation networks and publication records amplify institutional inequality — making "the rich institutions" richer

# Why these biases happen — and what can we do?

- Why these biases happen
  - ▶ Reviewers (and editors) may rely on stereotypes, familiarity, or "status cues" (institution name, authors' reputation) rather than content
  - ▶ Reviewing is human work; unconscious bias — on gender, prestige, nationality, career stage — is common
  - ▶ Peer-review bias is not just about individual reviewers, but structural patterns: who gets invited to review, whose papers get visibility, who gets accepted
- What this means for us
  - ▶ As reviewers: stay alert to bias. Evaluate based on content, not identity or institution
  - ▶ As authors: be aware that results and visibility may be influenced by social factors beyond quality
  - ▶ As a community: encourage diversity in reviewers/editors, anonymised review where possible, transparency in editorial decisions
  - ▶ Use blind review where possible

# What is reproducibility / replicability?

- **Reproducibility:** re-analyse the same data/code to get the same result
- **Replicability:** collect new data under same conditions and test whether the effect holds
- Both are fundamental for reliable knowledge — without them, published findings remain provisional (Goodman et al., 2016)

# What the crisis looks like — some sobering numbers

- In a 2015 project replicating 100 psychology studies, only 36% of results were successfully replicated; effect sizes dropped by about 50% (Open Science Collaboration, 2015)
- A 2016 survey in biomedical research found over 70% of researchers had tried and failed to reproduce another scientist's results — and 52% judged that there is a significant crisis (Baker, 2016)
- Fields like oncology (e.g. cancer-biology) show particularly low reproducibility: external audits of many "landmark" papers found  $< 1/3$  of results reliably reproduced (Errington et al., 2021)

## Why this matters

If published findings routinely fail replication, the scientific literature becomes unreliable — which undermines trust, wastes resources, and misleads follow-up research.

Video: [Veritasium: "Is Most Published Research Wrong?"](#) (12 min)

# Causes: why reproducibility fails

- Low statistical power (too few subjects/samples) — leads to unstable, non-robust effects (Ioannidis, 2005)
- Publication bias & questionable research practices (QRPs): only positive or novel results tend to get published; null results stay hidden
- Poor documentation, missing code/data, hidden methods — even honest researchers make replication hard when they do not share full materials
- Time & incentive structures: scientists rewarded for novelty, impact, and speed — not for replication or careful method sharing

# How FAIR + CARE + Good Writing helps fix reproducibility

- Proper metadata, shared code/data, and clear methods (FAIR) make re-analysis possible
  - Ethical responsibility (CARE) encourages transparent, honest reporting — including limitations, context, and reproducibility risks
  - Careful writing and documentation help readers understand exactly how a study was done
  - Together — transparent data practices, open sharing + honest reporting
- = more reliable science

# Citation as trust

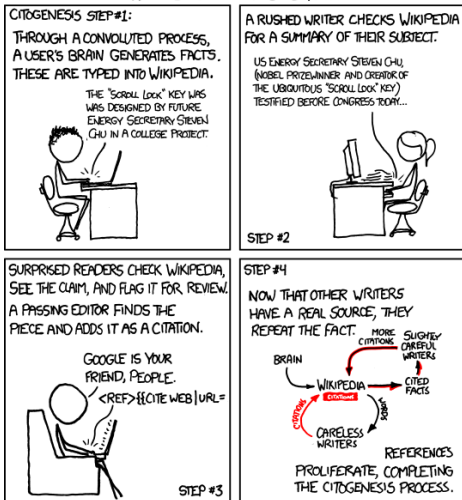
- A citation is a **link of trust**: it says “I have read and found this work reliable enough to build upon”
- Over time, citations form **networks of credibility** — maps of intellectual influence
- Not all citations are equal:
  - ▶ Some endorse, others critique or merely mention
  - ▶ “Citation cascades” can perpetuate unverified claims
- Responsible citation means reading, understanding, and acknowledging context — not just counting links

*Citation is conversation: each reference extends a thread of shared trust.*



# Citogenesis xkcd 978

## WHERE CITATIONS COME FROM:



# AI in peer review: emerging concerns

- Large language models (LLMs) like ChatGPT are increasingly used in academic writing — and in peer review
- A 2024 study estimated that 6.5–17% of peer reviews at major AI conferences may have been substantially written or modified by LLMs (Liang et al., 2024)
- Telltale signs: unusual adjectives (“commendable”, “meticulous”, “intricate”) spiked 10–35× in frequency
- Major publishers (Springer Nature, Taylor & Francis, Sage) now prohibit or restrict AI use in peer review

# Why AI-generated reviews are problematic

- **Confidentiality breach:** uploading a manuscript to an LLM shares it with a third party
- **Accountability:** reviewers are responsible for accuracy and judgment — LLMs cannot be held accountable
- **Bias and hallucination:** LLMs can perpetuate biases and produce plausible-sounding but incorrect critiques
- **Superficiality:** AI-generated reviews tend to be generic and miss domain-specific issues
- **Trust erosion:** if reviews are automated, peer review loses its meaning as expert evaluation

*Peer review depends on human expertise, judgment, and responsibility — qualities LLMs cannot replicate.*

# Current guidance on AI in peer review

- Most journals and conferences now require:
  - ▶ Disclosure of any AI assistance in writing or reviewing
  - ▶ Human accountability for all submitted content
  - ▶ No uploading of confidential manuscripts to AI tools
- Acceptable uses (with disclosure):
  - ▶ Grammar and spell-checking
  - ▶ Improving clarity of your own writing
- Unacceptable uses:
  - ▶ Generating review text
  - ▶ Summarising or analysing manuscripts you are reviewing
  - ▶ Submitting AI-generated content as your own expert judgment

# Roadmap

- 1 Feedback and Revision
- 2 Peer Review, Trust, and Metrics
- 3 Problems: Bias & Inequality, (un)Reproducibility
- 4 Peer Review & Revision in Practice**
- 5 References

# From theory to practice

- Peer review, reproducibility, and citation ethics all serve one purpose: maintaining credibility
- Today we practise this on a smaller scale:
  - ▶ Read your partner's paper carefully
  - ▶ Give specific, constructive, respectful feedback
  - ▶ Remember: your review contributes to their improvement — and your own understanding of good writing

# Why revision matters (again)

- Good writing is *rewriting*: clarity, structure, and claims improve with feedback
- Revision is not cosmetic editing; it is **rethinking** your argument and evidence
- Peer review helps you see what your reader actually understands
- Goal today: find **actionable changes** you can complete before final submission

# What peer review is (and isn't)

- A reader's report on **clarity, evidence, logic, and structure**
  - Specific, constructive, and respectful; focused on **the writing**, not the writer
  - Not copyediting every comma; not "gotcha"
- ⇒ Prioritise the few changes with the biggest impact



# How to give useful feedback

- Be **specific**: point to a sentence/paragraph and say what is unclear and why
- Be **diagnostic**: name the issue (claim lacks evidence; paragraph lacks topic sentence)
- Be **actionable**: propose a next step (add a source; move paragraph; define a term)
- Balance: 1–2 strengths + 1–2 high-value improvements

# How to respond to reviews

- **Read carefully:** understand what the reviewer is actually asking — not what you fear they mean
- **Respond to everything:** address each comment, even if briefly (“We have revised X as suggested” or “We respectfully disagree because Y”)
- **Be specific in your response:** quote the comment, then explain exactly what you changed (with page/line numbers)
- **Don’t be defensive:** assume good intent; reviewers are trying to improve your work
- **Push back professionally:** if you disagree, explain why with evidence — editors respect reasoned disagreement
- **Thank your reviewers:** they volunteered their time

# Reviewer comment decoder

## What reviewers write

"The motivation is unclear"

"This needs more context"

"The contribution is incremental"

"Consider reorganising"

"Minor revisions needed"

"Major revisions needed"

"This is interesting but..."

"The authors may wish to consider..."

## What they often mean

Your introduction doesn't explain why anyone should care

You assumed background knowledge I don't have

I don't see what's new here

I got lost; the structure doesn't work

Fix these specific things and you're done

Good potential, but significant work required

Brace yourself

You really should do this

See also: [Shit My Reviewers Say](#) and [Reviewer 2 Must Be Stopped](#)

# Today's workflow (40 min)

- 1 Swap drafts and read silently (8–10 min)
- 2 Complete the peer review form (10–15 min)
- 3 Discuss feedback with your partner (5–10 min)
- 4 Each writer drafts a **revision plan** (5 min): top 3 changes

**Name your review.** This is non-blind, collegial feedback.

## 1) Content & Argument

- What is the main thesis / research question?
- Is the argument clear and sustained throughout?
- What claims need more evidence or better support?

## 2) Structure & Clarity

- Are abstract, introduction, body, conclusion present and effective?
- Does the paper flow logically? Are paragraphs coherent?

# Peer review form (2/2)

## 3) Use of Sources

- At least 5 reliable sources?
- Are sources integrated and evaluated (not just dropped in)?
- Are citations/references consistent (APA)?

## 4) Language & Style

- Tone appropriate for an academic audience?
- Any recurring grammar/style issues affecting clarity?

## 5) Suggestions for Improvement

- Strongest part of the paper:
- One concrete change before final submission:

## 6) Overall Assessment — Perfect / Minor Revision / Major Revision / Reject

# Revision memo

This should be around 1 page per review, due with the revised draft.

- **What you changed** (structure, argument, evidence, style) and why
- **How peer feedback helped** (cite specific comments)
- **What you did not change** (and why), if applicable
- **Remaining questions** or risks you want the instructor to note

*Treat the memo as a guided tour of your revisions.*

# Writer's pre-submission checklist (final paper)

- Thesis is specific and arguable; each section advances it
- Every claim with evidence; sources are credible and cited correctly
- Paragraphs have clear topic sentences; transitions guide the reader
- Figures/tables (if any) are labeled and discussed in text
- Title, abstract, and conclusion match the actual contribution



# Peer review in the classroom

- The same principles that guide journals can strengthen student writing:
  - ▶ **Critical reading:** noticing logic, evidence, and structure in someone else's paper sharpens your own awareness
  - ▶ **Audience awareness:** seeing how readers interpret your draft teaches you what is clear—or not
  - ▶ **Constructive feedback:** learning to explain why something works or fails builds editorial skill
  - ▶ **Collaboration:** mutual review builds a community of practice, not competition
- Effective peer review is generous but honest:
  - ▶ Praise what works; suggest specific improvements
  - ▶ Think of yourself as a temporary editor, helping the author achieve their intention

*Peer review trains both empathy and precision—the two skills every writer needs.*

# Acknowledgements

- I have had over 200 papers reviewed (thanks reviewers!)
- I have peer reviewed over 2,000 conference papers, 100 journal papers, 10 PhDs, 2 Habilitations, and many grants
- OpenAI (2025) was used to format the references, and generate a first draft of the slides

# Roadmap

- 1 Feedback and Revision
- 2 Peer Review, Trust, and Metrics
- 3 Problems: Bias & Inequality, (un)Reproducibility
- 4 Peer Review & Revision in Practice
- 5 References

# References I

- ACL 2023 Program Committee. (2023). Review basics: Guidance for reviewers (acl 2023) [Accessed 2025-03-10]. <https://2023.aclweb.org/blog/review-basics/>
- ACL Committee on Peer Review. (2023). Report of the committee on peer review (cycle 5) [Accessed 2025-03-10]. <https://www.aclweb.org/adminwiki/images/9/9e/COPR-5-cycle-report-to-publish.pdf>
- ACL Rolling Review. (2025). Reviewer guidelines [Accessed 2025-03-10]. <https://aclrollingreview.org/reviewerguidelines>
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533, 452–454. <https://doi.org/10.1038/533452a>
- Catmull, E., & Wallace, A. (2014). *Creativity, inc.: Overcoming the unseen forces that stand in the way of true inspiration* [Contains Pixar's principles on feedback, candor, and iterative notes sessions]. New York, Random House.
- Errington, T. M. Et al. (2021). Investigating the replicability of preclinical cancer biology. *eLife*. <https://doi.org/10.7554/eLife.71601>

## References II

- Goodman, S. N., Fanelli, D., & Ioannidis, J. P. A. (2016). What does research reproducibility mean? In *Meta-research: Methods and reporting*. National Academies Press.  
<https://www.ncbi.nlm.nih.gov/books/NBK547531/>
- Helmer, M., Schottdorf, M., Neef, A., & Battaglia, D. (2017). Gender bias in scholarly peer review. *eLife*, 6, e21718. <https://doi.org/10.7554/eLife.21718>
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Kulczycki, E. (2017a). Assessing publications through a bibliometric indicator: The case of comprehensive evaluation of scientific units in Poland. *Research Evaluation*, 26(1), 41–52. <https://doi.org/10.1093/reseval/rvw023>
- Kulczycki, E. (2017b). Punktoza jako strategia w grze parametrycznej w Polsce. *Nauka i Szkolnictwo Wyższe*, 49(1), 63–78. <https://doi.org/10.14746/nisw.2017.1.4>

## References III

- Liang, W. Et al. (2024). Monitoring AI-modified content at scale: A case study on the impact of ChatGPT on AI conference peer reviews. In *International conference on machine learning*. <https://doi.org/10.48550/arXiv.2403.07183>
- Murray, D., Siler, K., Lariviere, V., Chan, W., Collings, A., Raymond, J., & Sugimoto, C. R. (2020). Author–reviewer gender mismatch in peer review. *eLife*, 9, e57184. <https://doi.org/10.7554/eLife.57184>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
- OpenAI. (2025). *Chatgpt (mar 14, 2025 version)* (Large language model). Retrieved October 14, 2025, from <https://chat.openai.com/>
- Rockwell, S. (2005). Ethics of peer review: A guide for manuscript reviewers [Accessed 2025-12-06].

## References IV

- Tomkins, A., Zhang, M., & Heavlin, W. D. (2017). Reviewer bias in single- versus double-blind peer review. *Proceedings of the National Academy of Sciences*, 114(48), 12708–12713. <https://doi.org/10.1073/pnas.1707323114>
- Wennerås, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, 387(6631), 341–343. <https://doi.org/10.1038/387341a0>