ACL Peer Review Committee Report

Report 2: Incentives for Improving Peer Review

Committee members:

* Roy Schwartz (co-chair)
* Anna Rogers (co-chair)
* Steven Bethard
* Roi Reichart

* Yuki Arase
* Vivian Chen
* Barbara Plank
* Noah A. Smith

Abstract

The ACL peer-review committee has been asked by the ACL exec, ACL Rolling Review and EMNLP 2025 program chairs to consider the problem of systematic incentives for peer review quality in the *ACL community. This report presents the results of a long discussion of the available options, their technical possibility and the current practices in other conferences. We provide recommendations that can be adopted starting from May 2025 (review cycle associated with EMNLP'25).

Background and Motivation

As the field and the volume of conference submissions grew, so did the volume of complaints about the quality of peer reviews. Some of the factors contributing to this issue are the following:

- The peer review process overall relies on volunteer service from researchers who are overall incentivized to devote time to their own research rather than service. The reviewer role bears little prestige and recognition.
- There are bad-faith actors who do not perform the duties they undertook without warning, or do not perform even the minimally expected amount of work (e.g. violating the policy on the use of generative language models, which is disallowed for the actual content of the review).
- As the field grows, it is harder to maintain the community norms about what constitutes a good review. WIth more and more new community members coming from different backgrounds, there is increasing need for various checklists and guidelines codifying the expectations and editorial policies, so that the authors could still have some idea of the venue identity and expectations for their papers. The iterative improvement of ARR also necessitates keeping up-to-date with its policies. While necessary, this increases the workload on chairs and reviewers (as the guidelines for each of these roles get progressively longer), while incentives for performing these roles are not improving.

The work of the peer review committee

The peer review committee collaborated with EMNLP'25 team and ACL Rolling Review to brainstorm the options for positive and negative incentives. We also considered the experience shared by [CVPR'25 chairs](#), who introduced for the first time the possibility of rejecting the papers of authors who show themselves highly irresponsible as reviewers.

Extent of the problem with bad reviews

According to analysis of ACL 2023 data[1], 12.9% reviews were flagged by authors as violating some norm expressed in reviewer guideline. Preliminary results from NAACL 2025 where review issue flagging was re-introduced suggest that the ratio of author-reported problematic reviews was in the same range overall, even slightly lower (9.79%). About 70% of NAACL 2025 review issue reports were fully upheld and taken into account by ACs.

These numbers indicate the ratio of reviews that are reported by authors as problematic in some way (e.g. not well-calibrated to the type of paper and contribution). While 10% of problematic reviews is of course 10% too many, this suggests that policy interventions should overall focus on reviewer training and incentives for providing good reviews, while providing some punitive options for truly egregious cases to be considered by chairs individually.

The 'problematic reviews' at ARR are not defined or originally intended to be used similarly to the CVPR notion of the 'highly irresponsible reviewers'[2]. Since only 19 out of 2,878 submissions recommended for submission were desk-rejected due to how the authors acted as reviewers, CVPR criteria seem to be more conservative and only aiming for the most egregious cases. Assuming each submission contributed 1 reviewer, this translates to less than 0.7% 'highly irresponsible' reviewers among the authors of successful papers.

Proposed solutions

Recommendation 1. Systematically provide incentives for great reviewers.

Based on the discussion, the committee considers the following financially and logistically feasible:

1. **Great reviewer/AC/SAC recognition and awards**. Consistently provide reviewer and AC/SAC awards akin to ACL outstanding paper awards. The 'great reviewers' will be identified based on AC recommendations, 'great ACs' will be identified by SACs, and 'great SACs' - by PCs. They will be listed on the ARR website for each cycle, and these lists will also be linked or reposted by the associated conferences, if

---

[1] See ACL'23 PC report, section 5.3: https://aclanthology.org/2023.acl-long.911/

[2] CVPR publicly posted the following criteria of 'highly irresponsible reviewers': *example cases of highly irresponsible behaviors are one-sentence reviews, reviews generated by Large Language Models (LLMs), reviews not relevant to the paper or that miss a substantial portion of the paper* https://cvpr.thecvf.com/Conferences/2025/CVPRChanges

any.

To increase the visibility of great service work during conferences, the reviewers/ACs/SACs, whose service was recognized either in the cycle directly associated with that conference, or a non-conference cycle preceding it, will be able to receive 'great reviewer/chair' stickers to use on their conference badges. Similarly to the CVPR initiative, the papers of 'great reviewers/chairs' (if any) can display a special 'great reviewer/chair' sticker in the poster sessions and the programs. A comparable design should be developed for virtual participation. The draft for proposed policy is included[3].

To also create a more tangible incentive, up to 2.5% great reviewers/chairs will be selected by lottery to receive a free virtual attendance at an *ACL conference or a discount on the on-site participation (to be used within a year of award). We provide separately a more detailed awards policy proposal, coordinated with the ACL events director.

2. **System for promoting reviewers.** The ARR reviewer pool is already shared with ACL journals. The people nominated as great reviewers/chairs should be eligible (if they so wish) for offers to move on to a more senior role within ARR, or to join the pool of a journal. The decision to recruit a given candidate remains with the editorial board of the venue, but the likelihood for such a 'promotion' increases.

3. **Revisit this policy after OpenReview reviewer history is implemented.** OpenReview is currently working on adding statistics to reviewer profiles, but it is not currently clear what exactly it will be possible to track (at the very least number of reviews and number of days late to submit reviews), and to what extent it will be customizable for the review quality measures already available at ARR. Hence, the current recommendation is to revisit this policy to enhance review effort visibility after the Reviewer history feature is implemented, and consider what statistics is possible to make public per reviewer, what can be visible internally, and how the consent of the reviewers/chairs will be collected.

Recommendation 2. Create a clear protocol for handling deeply problematic reviewers.

1. **Penalty for 'highly irresponsible' reviewers/chairs.** The reviewers and area chairs who show themselves as 'highly irresponsible' (subject to case-by-case consideration by the program chairs) will be subject to the following penalty. They will be ineligible to (re)submit any paper for which they are an author to the next cycle of ARR, or commit any paper for which they are an author to any *CL conference whose submission deadline falls within the next cycle of ARR.

The goal of this policy is to encourage the authors to make sure that everybody on their teams submits their reviews on time and in accordance with guidelines. The policy is not meant to force the irresponsible reviewers out of authorship positions (which would be counter to ACL publication ethics). Instead, the teams of such

---

[3]

https://docs.google.com/document/d/1rsKMfyrIoJ3cBUJtFj_qlk6gDMNzJRwKj7h0YQvMurM/edit?tab=t.0#heading=h.vseb4395tr5b

papers would have to sit out one cycle or go elsewhere.

2. **Mechanism, process and criteria for reporting 'highly irresponsible' reviewers/chairs.** For both regular [reviewers](#) and [area chairs](#), ARR authors can already submit issue reports for various problems, including rude, too-short, possibly LLM-generated (meta-)reviews. The chairs at the next level are supposed to consider these complaints.

   They may also wish to flag reviews not flagged by authors. To enable this, a new button will be added to reviews/metareviews, visible to the ACs/SACs/PCs (for reviews) and SACs/PCs (for meta-reviews). This form will enable reporting either great or highly irresponsible reviews/meta-reviews.
   - Where the author reports a review issue but AC fully or partly disagrees, the reviewer is considered cleared.
   - Where the author reports a review issue and the AC agrees with the author, the AC has the option to further flag the case to be considered by SACs/PCs as potentially 'highly irresponsible', which could merit the highest penalty (see (1) above).
   - Where the author reports a review issue, but AC does not report taking an action, the SACs/PCs consider the case and decide on the case as well as AC responsiveness.
   - Cases where reviewers or ACs disappear without giving a notice will be detected automatically, and confirmed by SACs/PCs, based on the non-submissions of the new form in ARR that allows [declaring personal emergencies](#) that necessitate reassignments. People who went missing in action without notice will be considered 'highly irresponsible'.
   - Where AC/SAC flag reviews/metareviews that had not been reported by authors, all cases are considered by SACs/PCs.
   - The specific criteria for which the EMNLP 2025 team deems reviewers/ACs as 'highly irresponsible', as well as a high-level description of the decision process and statistics for penalties, will be publicly available in the PC report.

3. **Maintain and develop this policy in the future ARR cycles.** The definition of a 'highly irresponsible reviewer' may need to be adjusted in the future, based on the real cases observed in the conference work, and the process may need to be revisited based on how much workload the above process would generate for PCs/SACs.

4. **Publicly announce this policy and any subsequent changes to it.** The initial criteria for the 'highly irresponsible' reviewers and any further changes to it will be publicly announced and integrated in ARR documentation. Reminders about this policy should be included in the reviewer/AC registration forms and late reviewer/AC chasing emails. Ideally, we wouldn't ever need to implement it in practice.

5. **Create and publicly announce an appeal mechanism.** The reviewers/ACs who want to dispute the Program Chair decisions in their case may appeal to the publication ethics committee[4], which typically handles cases on the timeline of 2-3 months.

---

[4] https://www.aclweb.org/adminwiki/index.php/Process_for_ACL_Publication_Ethics_Review