Full Report:
2019 ACL Survey on Reviewing

Graham Neubig
Marti Hearst

July 2019

Executive Summary

Background

The ACL survey on reviewing was conducted from May 6 - June 5, 2019 and received 422 complete responses. The survey was advertised twice on the ACL membership email distribution list as well as on social media.

Purpose: The purpose of this survey was to inform ACL leadership and the membership about the membership's opinions about several aspects of the reviewing process, based on members' experience as authors, reviewers, area chairs (AC), or program chairs (PC). The results of this survey will help guide decisions about these policies in the future.

Structure: The survey consisted of 26 questions organized as (i) respondents' experience as reviewers, authors, area chairs, and program co-chairs, and then opinions on: (ii) author response, (iii) author discussion, (iv) meta-review, (v) structured review forms, (vi) review transparency, (vii) acceptance rates, (viii) timing of conference dates and review release, (iv) public review, (x) open comments, and finally, (xi) demographics. Participants were allowed to skip questions, so results are reported in percentages throughout. An option of ``No opinion / not sure'' was offered on most questions.

Summary of Results

Demographics: The vast majority of participants (95%) are either current or previous members of the ACL, and as far as we can tell from available membership statistics, the sample appears to be representative with respect to demographic factors like geographical affiliation, gender, role and academic background.

The vast majority of participants have submitted an *ACL paper for review at least once (98%) or at least twice (95%) over the last 10 years. However, there were a significant number of respondents with little experience reviewing (13% never, 8% one time, 8% two times). Furthermore, most respondents have not had first-hand experience with higher levels of reviewing. 74% have not been an AC, and only 16% have been an AC two or more times in the last ten years. The respondent pool consisted of 7% of people who have been program co-chairs for *ACL in that time period. Thus, most respondents are relatively inexperienced in the reviewing process.

Opinions: This is a brief summary; details and charts appear below.

- **Author Response**: A significant majority of survey respondents were ***in favor of*** author response as a whole, with 61% in favor compared to 27% opposed. Very few were strongly opposed. Preference for author response was negatively correlated with reviewer experience.

- **Author Discussion**: Opinions were ***nearly evenly split*** on author discussion, with 37% in favor, 32% opposed, and 30% with no opinion. However, those who had experienced author discussion were more in favor of it than those who had not experienced it. Those with experience as an Area Chair were more opposed.

- **Meta-Reviews**: There was ***strong support*** for meta-reviews (only 9% opposed), with the majority (73%) agreeing that it is ok to have them only for borderline cases. However, former PC chairs were split 50%-35% on the latter question.

- **Review Form Structure**: A significant majority supported (65%) a ***minimal amount*** of structure in review forms, although a minority (24%) preferred ***more structure***. There was a clear trend that more experienced reviewers preferred less structured review forms.

- **Review Transparency**: A ***clear majority*** (54%) preferred that reviews be released only to authors, but a ***notable minority*** (36%) preferred public release of reviews. Those with no experience reviewing more often in favor of public release of reviews, while those with some experience reviewing were more often opposed. Public release of reviews was one area where there was a *significant gap between genders*, with only 26% of female respondents preferring reviews be publicly released, compared with about 40% for those who were male or preferred not to state. On the other hand, there were no regional differences in preferring public release, but a significant 20% of researchers from the Asia/Pacific region preferred that reviews not be released even to authors. Finally, those with program chair experience were even more strongly against public release of reviews, with only 14% agreeing. 18% were opposed to even releasing reviews to authors.

- **Post Review Discussion Transparency**: There was ***significant disagreement*** in whether post-review discussion should be released to authors, with about half of

those who responded with an opinion (47%) preferring that it not, and the remaining half split between release to authors (24%) and public release (20%). As seen in similar questions, those with greater amounts of review experience preferred that post-review discussion be kept private, while those who had never reviewed strongly preferred that discussion be available to authors or the public.

- **Meta-Review Transparency**: There was ***strong support in releasing*** meta-reviews to ***authors***, with 83% agreeing. 33% were also in favor of releasing to the public. Similar trends were seen as above for reviewer experience.

- **Acceptance Rates**: **A near majority** (47%) were in favor of keeping the status quo with respect to acceptance rates, while 32% were in favor of increasing, 9% were in favor of decreasing. A large majority of respondents (68%) preferred that acceptance rates not be decided before-hand, but rather post-hoc based on quality of papers. There was ***strong overall support*** (83%) for ***the status quo*** for conference publications remaining selective.

- **Timing of Review Release**: The ***great majority*** said that the timing between review release and the next conference deadlines was ***at least somewhat important***, (83%-14%). People would prefer at least 2-3 weeks, and many (39%) said at least a month was preferable. Female respondents slightly preferred having more time between review release and the next deadline, and those who preferred not to state their gender particularly preferred longer reviewing cycles.

- **Public Review**: Opinions about public review ***were mixed and tended to be strong***, with 42% opposed and 32% in favor. Of these, 37% held strong opinions. Those who had experience participating in conferences with public review tended to be in favor 50%-30%, while those who did not have such experience were opposed 27%-47%. In general, support for public review tended to be inversely correlated with reviewing experience. Female respondents were less likely to support public review than male respondents, 19%-55% vs. 36%-39% respectively. However, female respondents were only half as likely as male respondents to have participated in a conference with public review, and the effects of gender and experience with public review may be conflated in these results.

Summary of Open Comments: This is a summary of the main comments raised in the open comments section, as well as selected ideas put forward; the interested reader can review the full, organized list of comments:

- **Author Response:** Positive: ACs/PCs find it useful; reviewers find it useful; authors find it useful for clarifications, to repair poor reviews, and for fairness. Neutral: perhaps should be used only in some cases or only sent to ACs; need better indication of when response considered. Negative: too much effort for too little effect; a stressful time-sink with little feedback; too much time for reviewers and authors with little ultimate payoff; reviewer discussion is more valuable; ACL author response not as thorough/long as other conferences/journals.

- **Author Discussion:** Positive: If set up more like a journal model, then a good idea; the OpenReview style of discussion is a good model. Neutral: a good idea, but only with a mechanism for conditional acceptance; would need to reduce the number of papers per reviewer; maybe for limited cases only. Negative: too much work for reviewers/ACs/authors; not enough time between conferences; likely to be as ineffective as regular author response; only works for those with flexible schedules. *Idea: a revise-and-resubmit model for the next *ACL conference where reviewers have access to prior reviews.*

- **Meta-Reviews:** Positive: Helpful to authors and PCs when reviews conflict; ensures ACs review all reviews; imposes discipline on the reviewing process; catches errors. Negative: more work for AC and benefit not clear. *Idea: occasional message from the PC explaining decision.*

- **Review Form Structure:** Positive: Minimal structure (over no structure) is useful as an AC; minimal structure a good middle ground for experienced reviewers; ACL 2019's form received many positive comments; can reduce bias and increase fairness and clarity. Negative: frequent changes in format across conferences; too much structure increases reviewer workload and can reduce review coherence (NAACL 2018 in particular); minimal word counts are not a good idea.

- **Transparency (all three types):** Positive: Comments are similar to those about public review, summarized below. Negative: Public release of review discussion may make frank discussion difficult/unintentionally reveal reviewers' identities.

- **Acceptance Rates / Conference Selectivity:** Pro higher acceptance rates: Promote journals more and at the same time make conferences less selective; have different levels of selectivity within conferences; low acceptance rates lead to inefficiencies as papers are resubmitted multiple times, and quality problems since good papers are rejected and there is variance in quality. Neutral: current acceptance rates seem to be striking a good balance; the medical system is not a good one; our rate is good; other fields' are too low. Con higher acceptance rates: being selective at conferences is an important quality control; important for jobs; impractical logistically. *Idea: focus higher acceptance rates on specific special topics.*

- **Timing of Review Release:** Having more time encourages meaningful revisions; it is important to demotivate resubmissions "as is"; reviewers need time to rest between reviewing assignments; timing of major conferences needs to be better spaced out. Negative: people should not feel they have to submit to every conference. *Ideas: allow authors to see reviews before final accept/reject decision to give them more time to revise work; do not allow people to resubmit unaltered papers; prior reviews should travel with resubmitted papers.*

- **Public Review:** Positive: incentivizes better reviewing; a way to recruit reviewers; speeds up research; might improve review quality. Neutral: it is essential to retain

double-blind reviewing. Negative: may hurt diversity; may discourage participation; may hurt early careers; may be a popularity contest; no filter on quality of comments; lack of quality on posted; unreviewed papers; too easy to game; too time-consuming. *Ideas: support public discussion after publication; try public review once as an experiment.*

- **General Comments:** Concerns about reviewer preparation, review quality, and reviewing load; concerns about diversity of acceptable content; comments on the bidding process; concerns about over-publishing (just noticeable difference). ***I****dea: two week review cycle (suggested by Omer Levy).*

Implications for Policy

These results can be divided into three categories, broadly. The first is in terms of guidance for conference planning for ACL Exec and Chapter Boards. There is general agreement about keeping the status quo in terms of acceptance rates, keeping conference presentations selective, retaining gaps in time between conferences to allow for resubmission.

The second is to help inform Program Co-Chairs and conference organizing committees terms of details for reviewing procedure in terms of author response, meta-reviews, and their transparency. Program co-chairs can take this information into account when making decisions about reviewing, and the ACL Exec can use this information to help prioritize changes to reviewing software.

The third category is the potential for more radical changes to *ACL reviewing procedures. In particular, some people are interested in open review. This survey has canvassed the views of the membership about this topic. Currently, the views seem to be split, which suggests that any path to adoption most likely would need to first take place in some kind of laboratory setting, such as workshop reviewing or the like. Additionally, some respondents have suggested other changes to the reviewing process that were not explicitly asked about but which may be a useful way forward. These include: requiring authors to include prior reviews when re-submitting papers; allow authors to see reviews before final decisions are made to allow more time for improving their work, and providing a method for public discussion of papers after acceptance and publication.

This survey does not address other pressing issues relating to reviewing that remain before *ACL, including the need for better automation of detection of reviewing matches, recruiting of reviewers, detection of conflicts of interest, and improvement of reviewing quality.

Definitions

Terms were defined as followed in the survey.

Author Response: In recent years, some *ACL conferences have had author responses, which provide a chance for authors to respond to reviewers to answer questions and provide clarifications before final decisions are made.

Author Discussion: A more comprehensive version of author discussion, some conferences have a discussion period where authors can interact with reviewers over an extended period of time. After initial reviews are released, authors may respond to the reviews point-by-point, and then the reviewers or ACs can ask additional follow-up questions or clarifications until the author discussion period is over. All this can be done in an anonymous fashion, preserving double-blind review.

Meta-Review: A Meta-review is a review performed by ACs after the review process completes that summarizes the views of the reviewers, and also explains the reasoning of the ACs regarding why they reached their final decision. These meta-reviews potentially make the reasoning about why decisions were made more clear.

Structured Review Form: Review forms require various levels of structure ranging from a simple score and free-form text box with which to enter reviews, to multiple text boxes requiring information about different aspects of the review.

Transparency: The decision about by whom information should be seen: program chairs and reviewers only, the authors of the papers, the public at large.

Acceptance Rates: Currently, acceptance rates for the major *ACL conferences are around 20-25%, and papers must be accepted via a competitive review process in order to be presented. (Demonstrations, posters and some other methods of presenting materials tend to have a higher acceptance rate, but no questions were asked about these.)

Conference Selectivity: In other fields (e.g. medicine), most submitted papers or abstracts are given presentations at conferences, and other measures are used to indicate relative quality of papers (e.g. journal publications). This question asked if conference publications should remain selective, not be selective with journals taking the role, or not be selective but with public reviews released.

Timing of Release: When several conferences are held in sequence back-to-back, it is sometimes the case that reviews are released only shortly before the next major conference submission deadline. Questions about timing refer to this limited time span for revisions before the next opportunity to submit.

Public Review: Some conferences have introduced a review mechanism called public review in which, in addition to a program committee, the public is allowed to view and publicly post comments about papers during the review period. Although in some conferences the authors' names are publicly exposed and the reviews remain even for rejected papers, it is possible to change the format and anonymize the submissions and remove rejected papers after the review period.

Abbreviated Table of Contents:

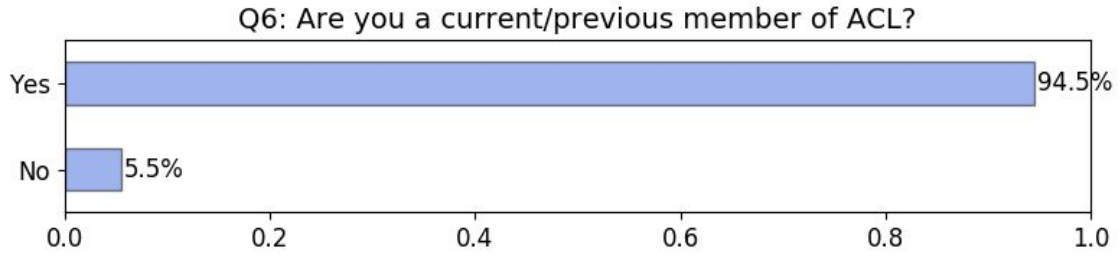
Detailed Results: Charts

The detailed results consist of charts on demographics and opinion statements. The open responses have been manually organized into categories for each question in the next section.

Demographics

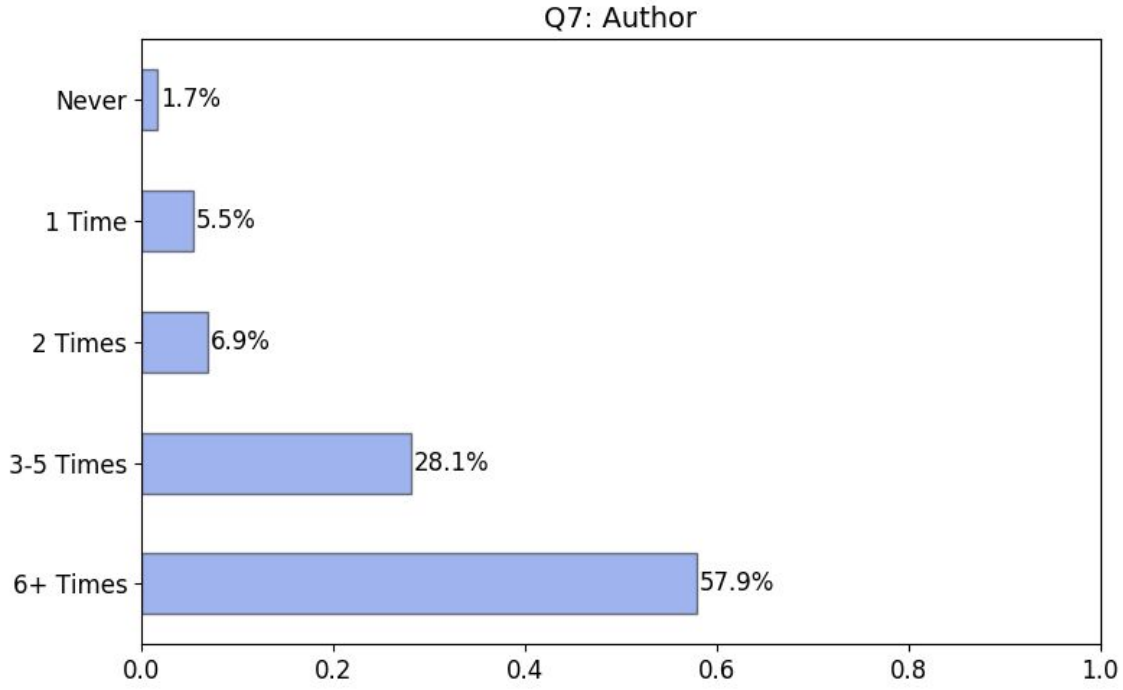
Q6: Are you a current/previous member of ACL?

The great majority of respondents were present/past members of the ACL.



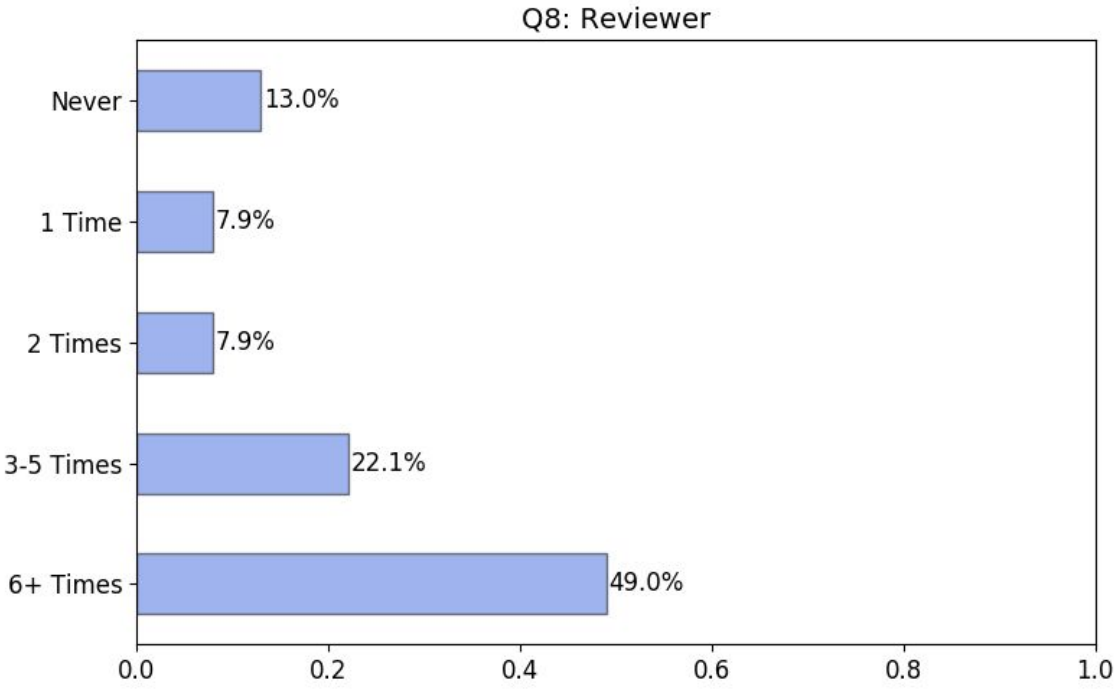
Q7: Author

Most respondents have experience as an author...

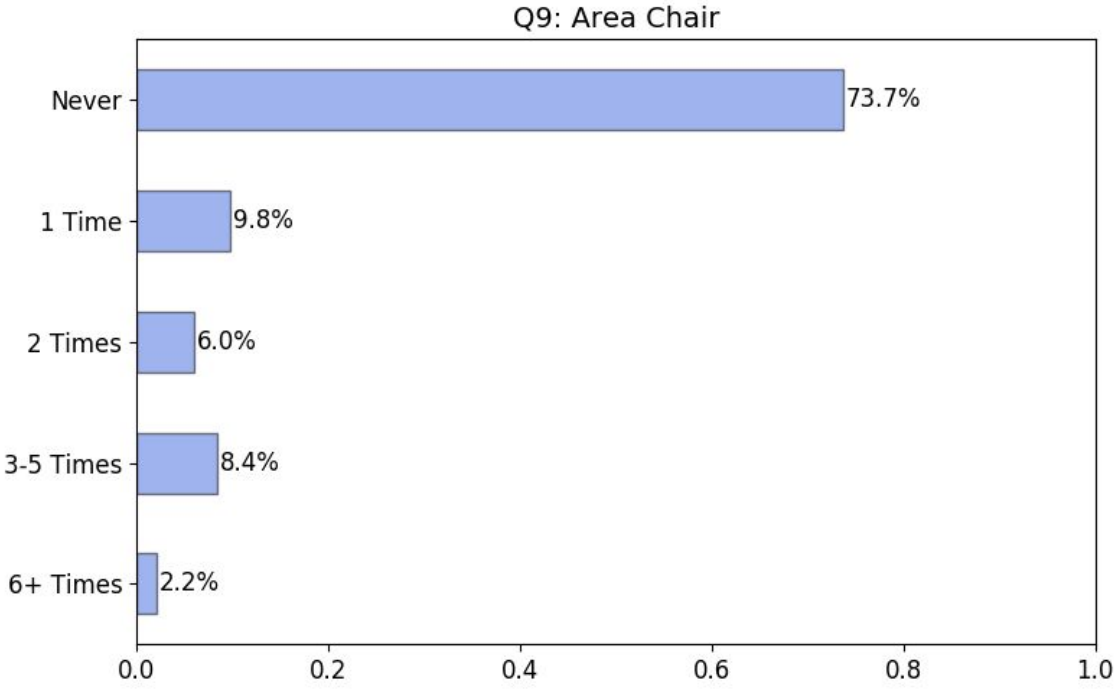


Q8: Reviewer

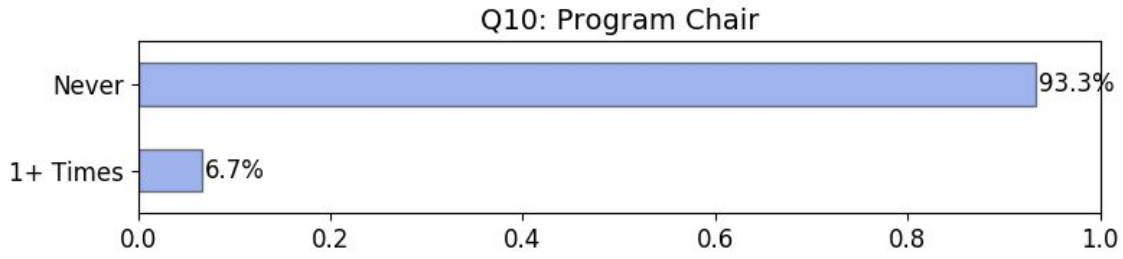
...but there were a significant number of respondents with little experience reviewing, and the majority had never been a PC/AC



Q9: Area Chair

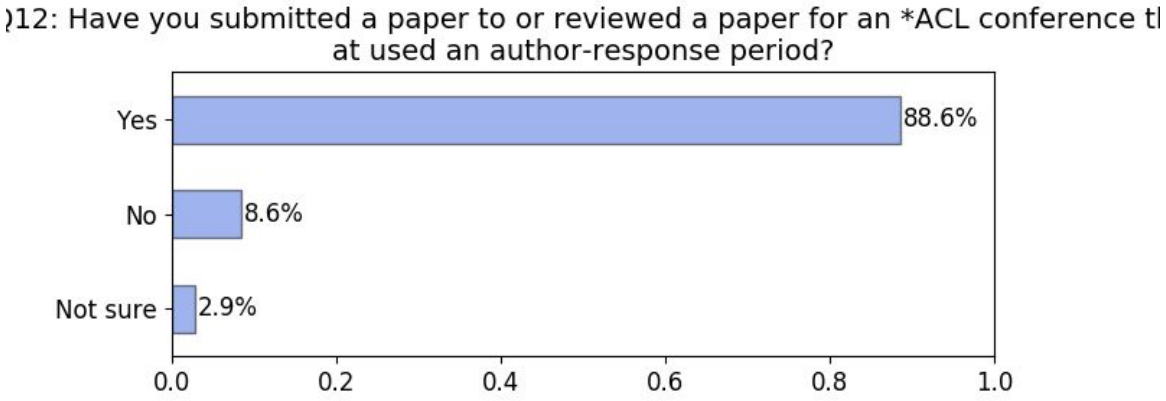


Q10: Program Chair



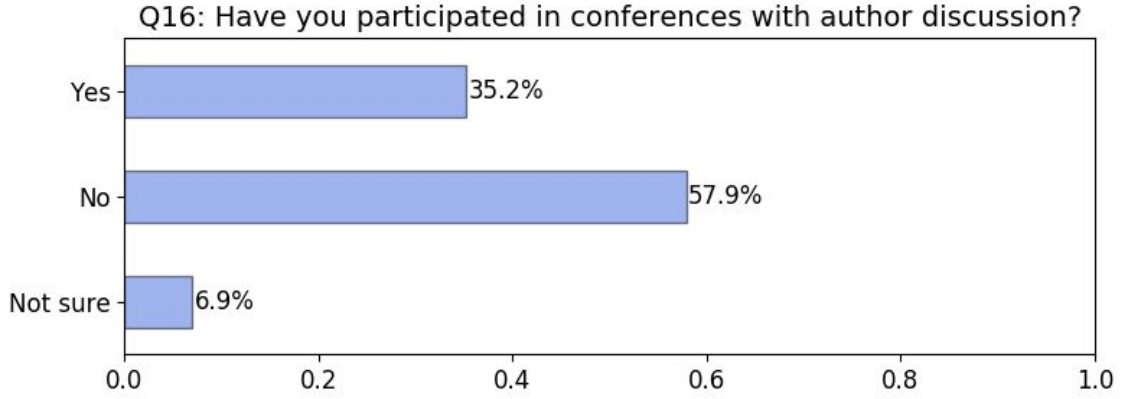
Q12: Have you submitted a paper to or reviewed a paper for an *ACL conference that used an author-response period?

The great majority of respondents have experience with author response...

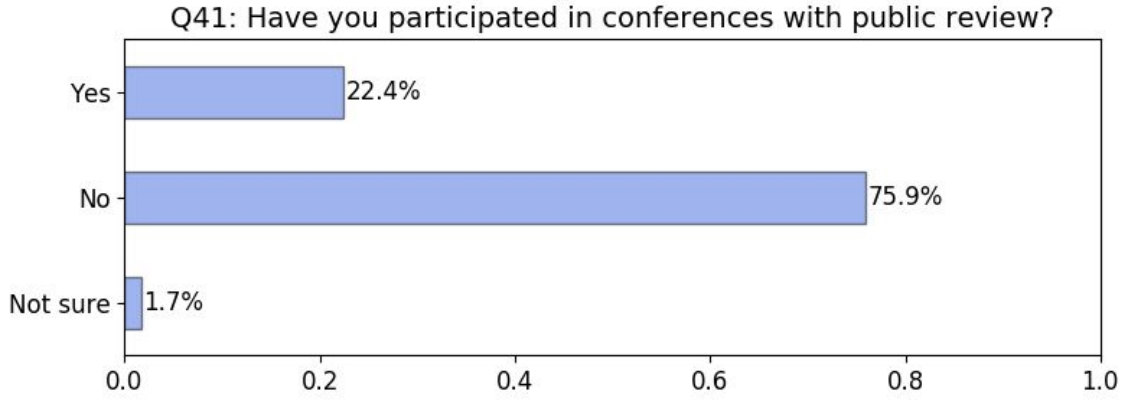


Q16: Have you participated in conferences with author discussion?

...but the majority had not participated in author discussion or in conferences with public review.

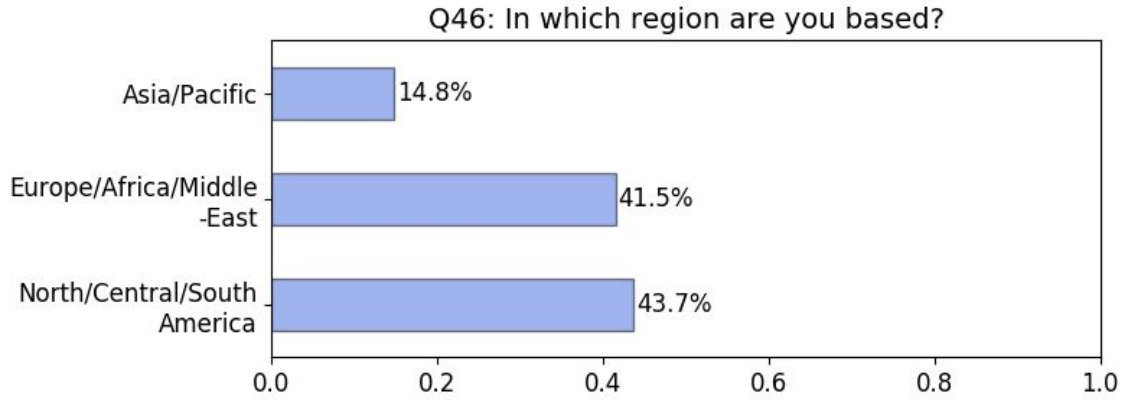


Q41: Have you participated in conferences with public review?



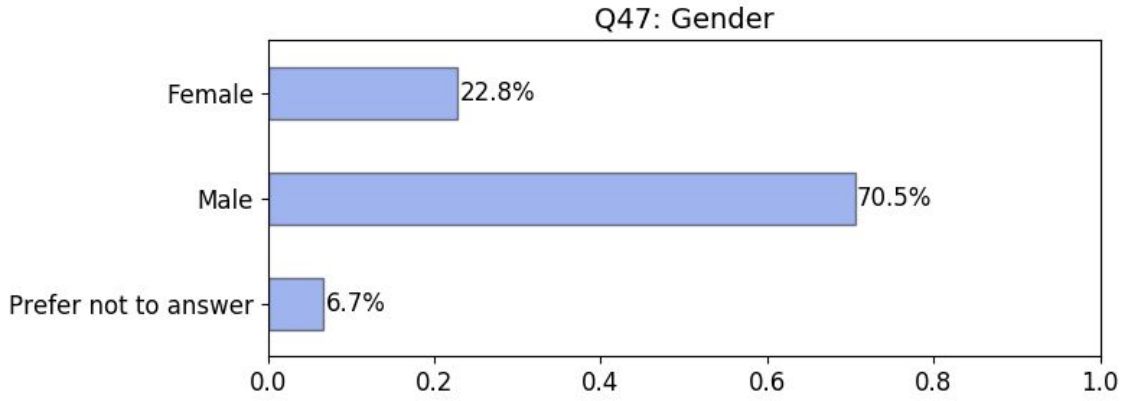
Q46: In which region are you based?

All regions of the world had some representation, but the Asia/Pacific region was somewhat underrepresented.



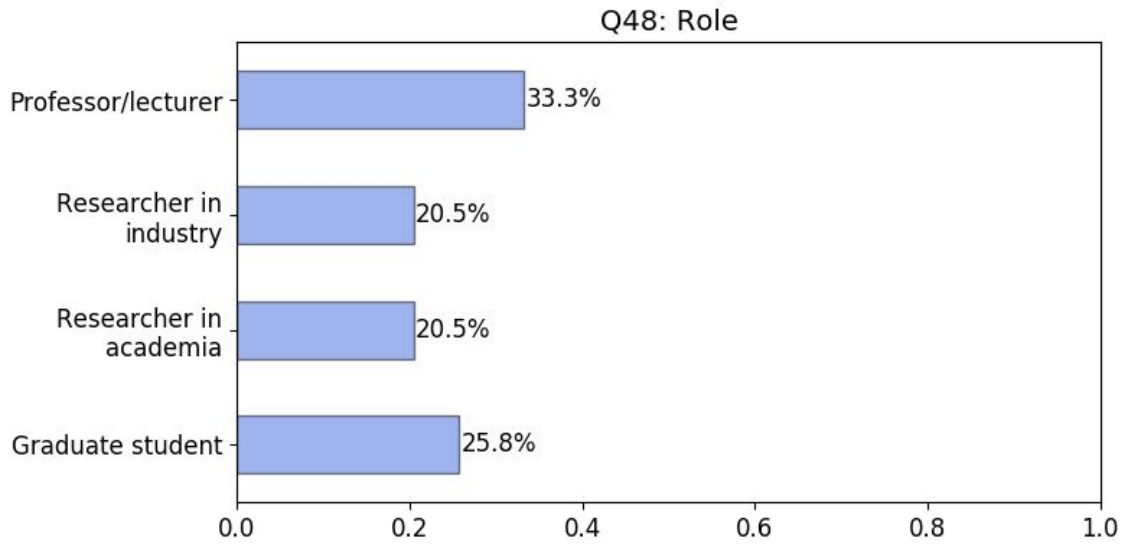
Q47: Gender

Female members of the community were underrepresented, with only about 1/4 of participants (although this may be roughly in line with ACL membership?).

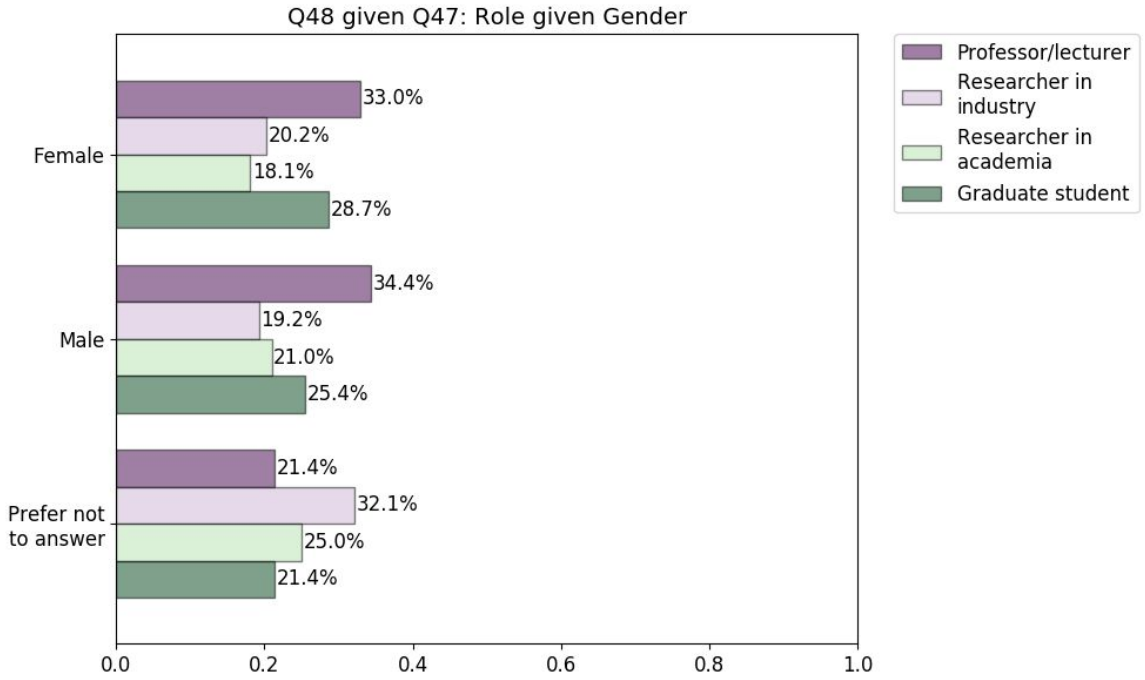


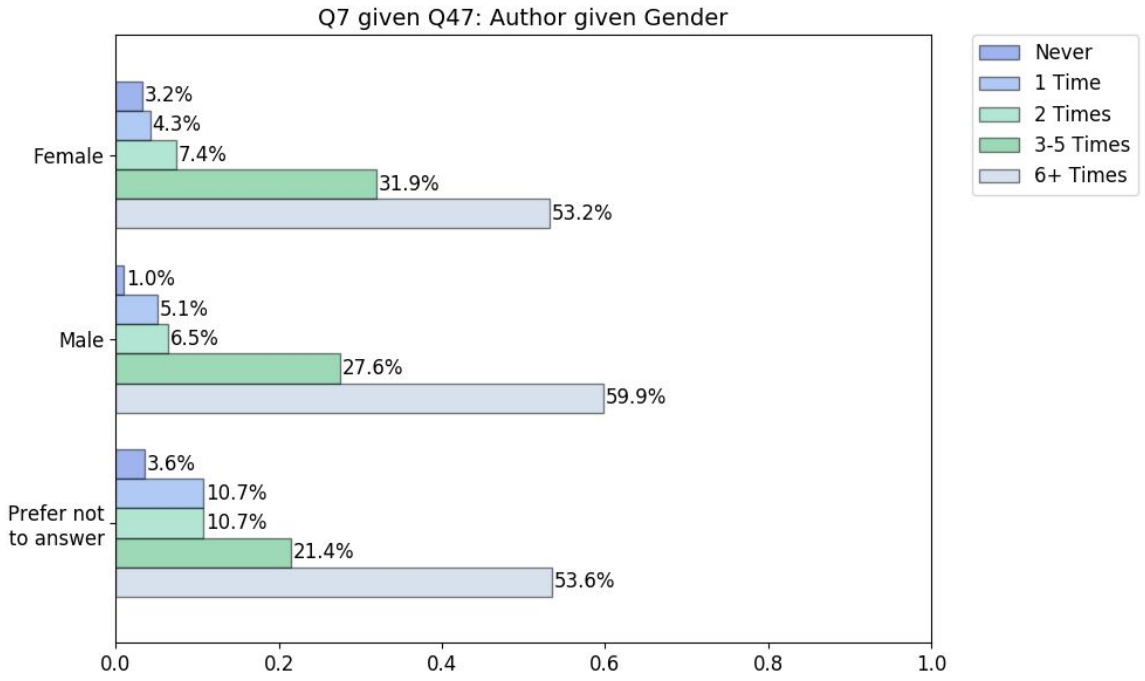
Q48: Role

There was fairly equal representation across many varieties of roles, with faculty members being slightly more represented than other groups.

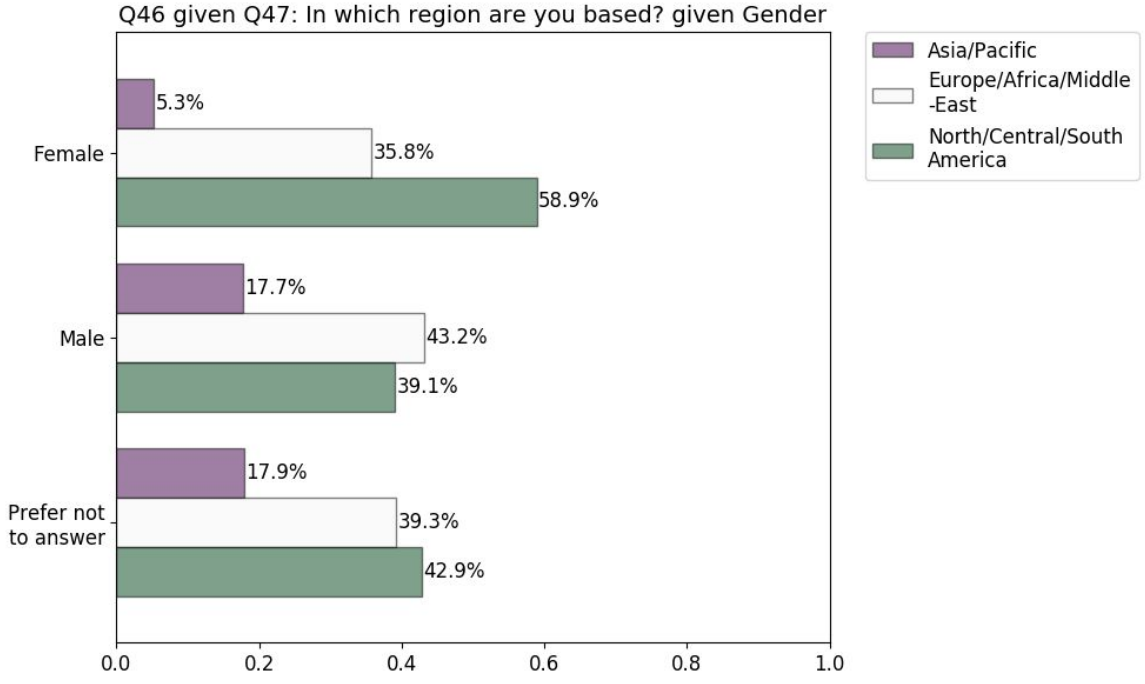


Notably, between respondents who identified as female or male there was not a significant difference in career roles, and only a small difference in reviewing experience. This will be important in later analysis, indicating that differences in opinion along gender lines are not simply due to correlated differences in experience or role. Out of those who answered "prefer not to answer" for the gender question, there were a lower proportion of professors/lecturers, and higher proportion of researchers in industry.

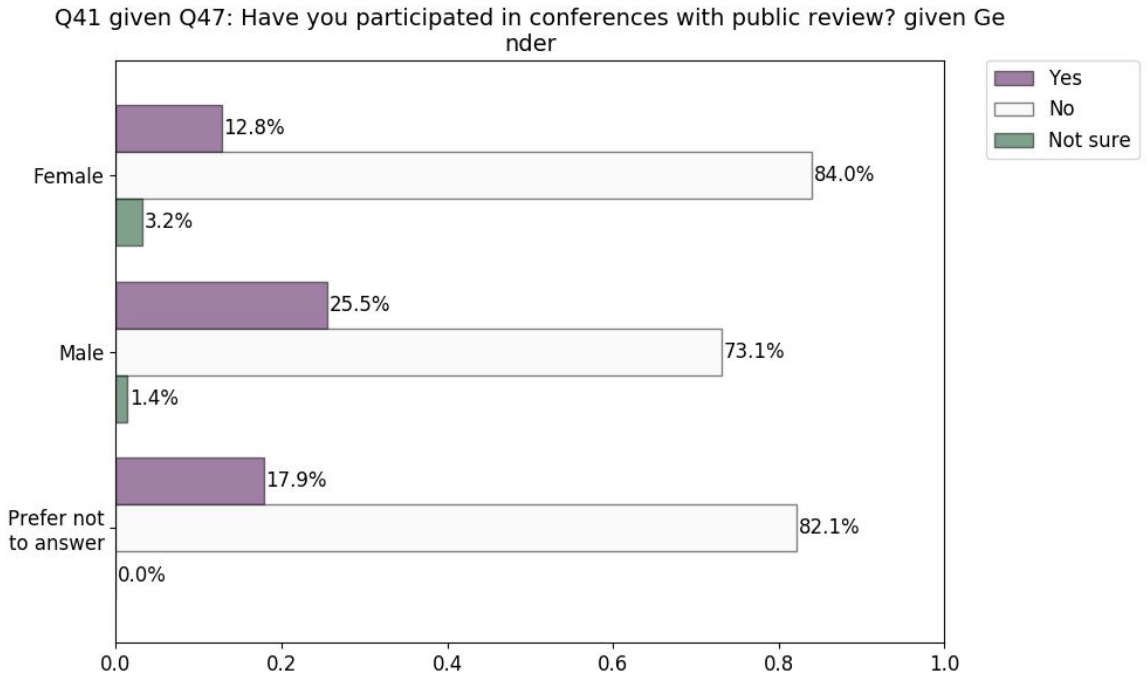




However, female respondents were much more likely to be from the Americas, and much less likely to be from the Asia/Pacific region, so preferences based on gender and preferences based on region may be conflated.



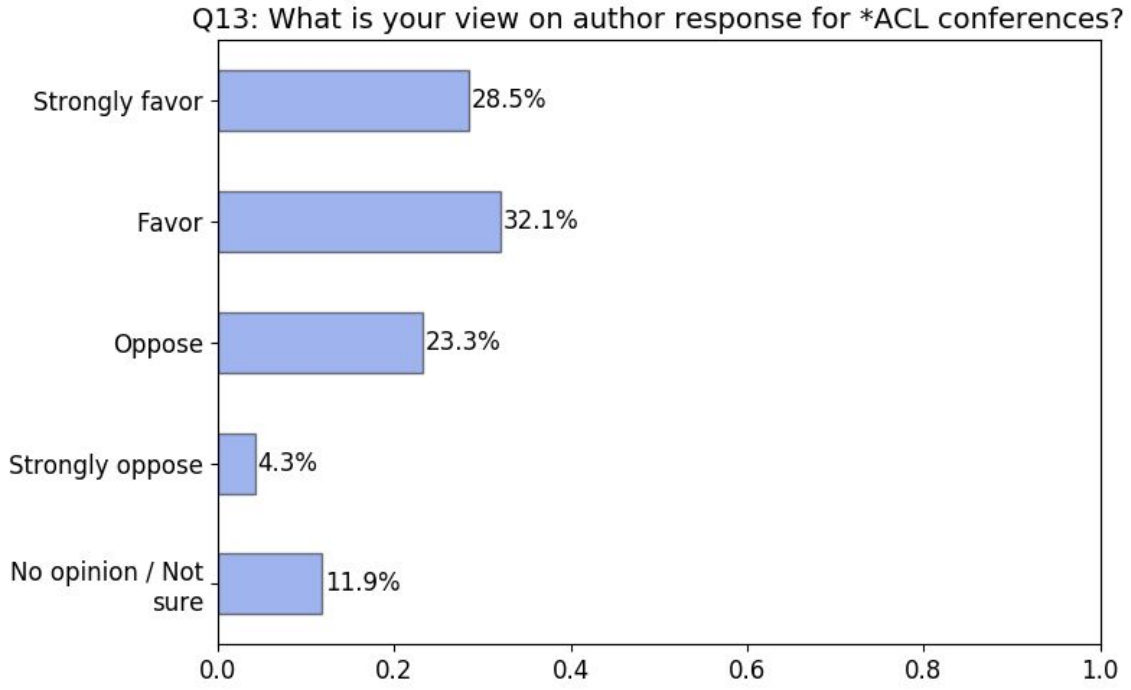
In addition, female respondents were much less likely to have participated in a conference with public reviews (13% vs. 26% for male respondents). This is perhaps because ICLR, the conference most heavily associated with public review, has a higher proportion of male participants than the average at *ACL conferences.



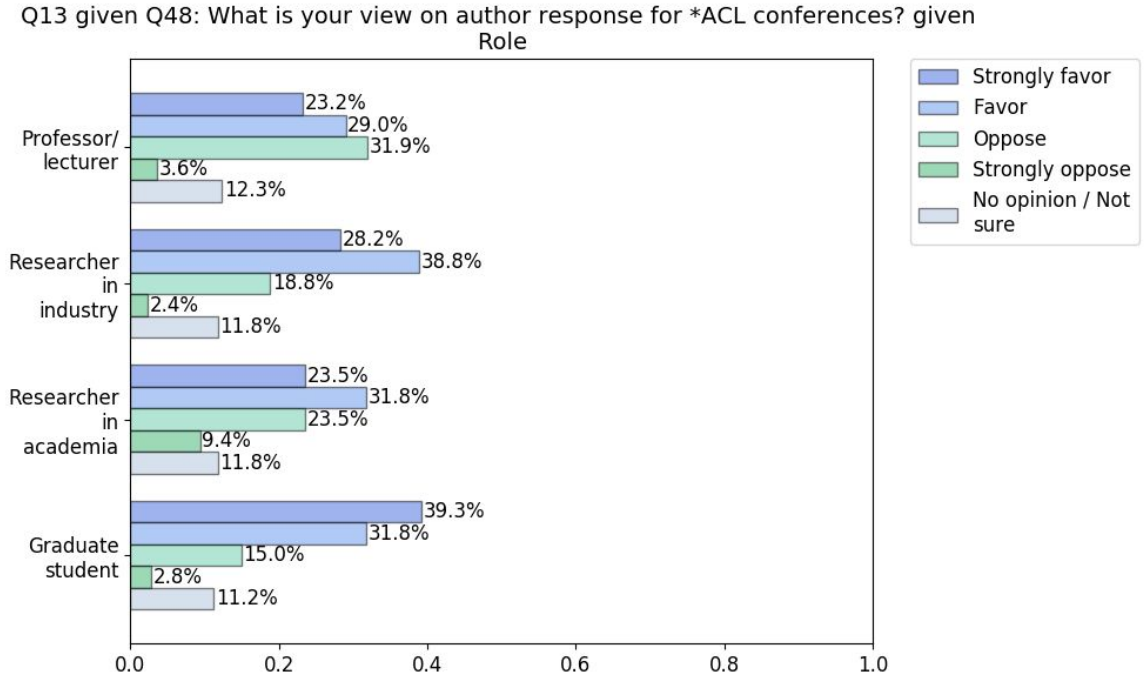
Author Response

Q13: What is your view on author response for *ACL conferences?

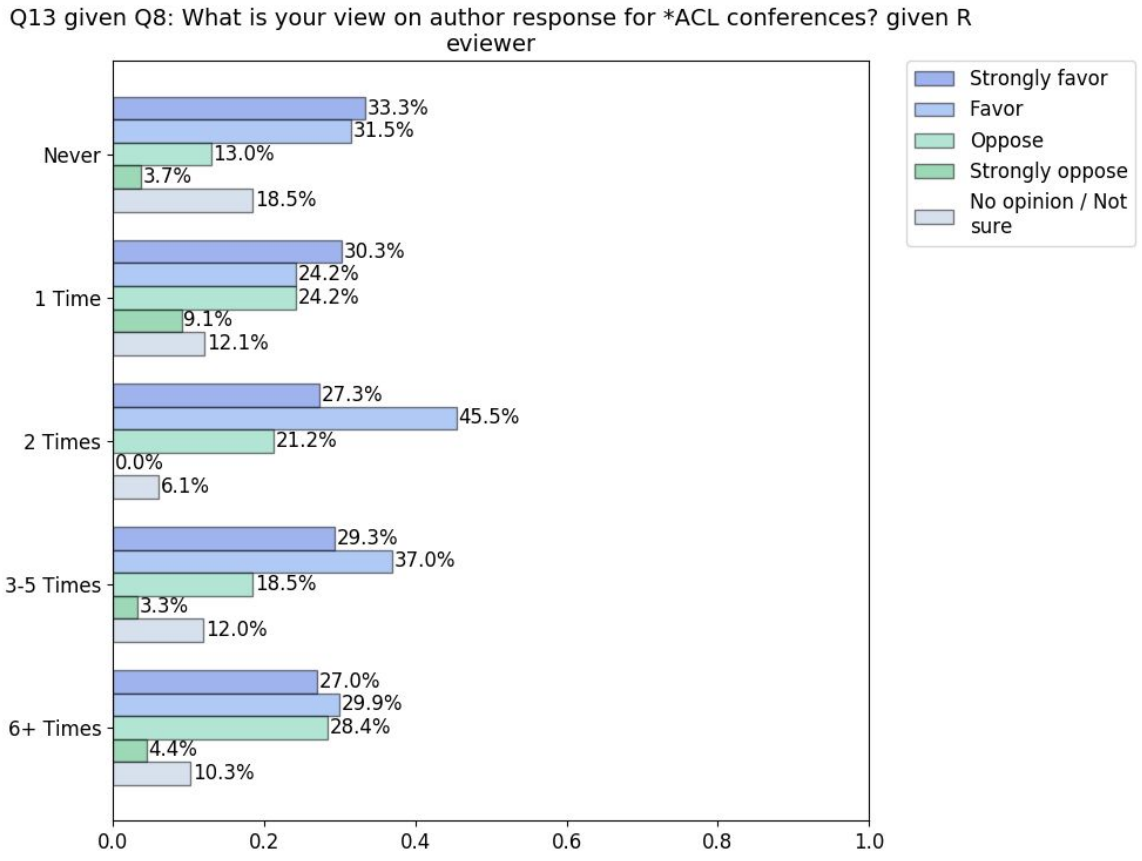
Survey respondents were in favor of author response as a whole, with 61% in favor compared to 27% opposed. Very few were strongly opposed.



The majority supported author response across all career roles, but role had a significant impact on preference for author response, with professors/lectures being least in favor at 52%-36%, and graduate students being most in favor 71%-18%



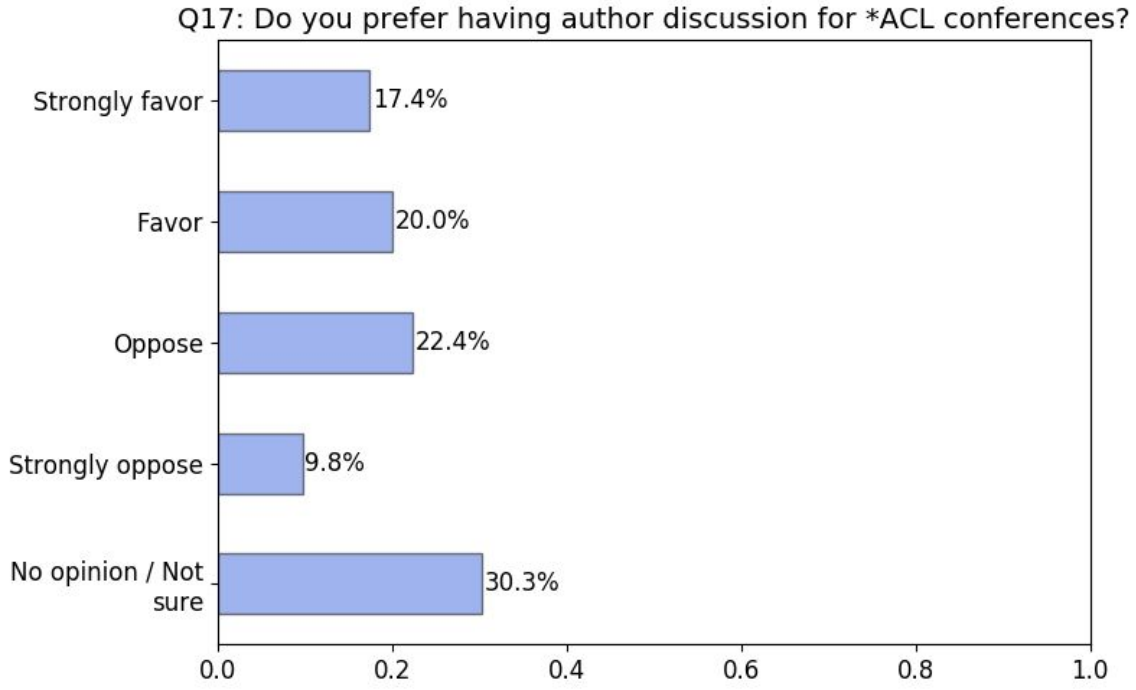
On a related note, preference for author response was negatively correlated with review experience.



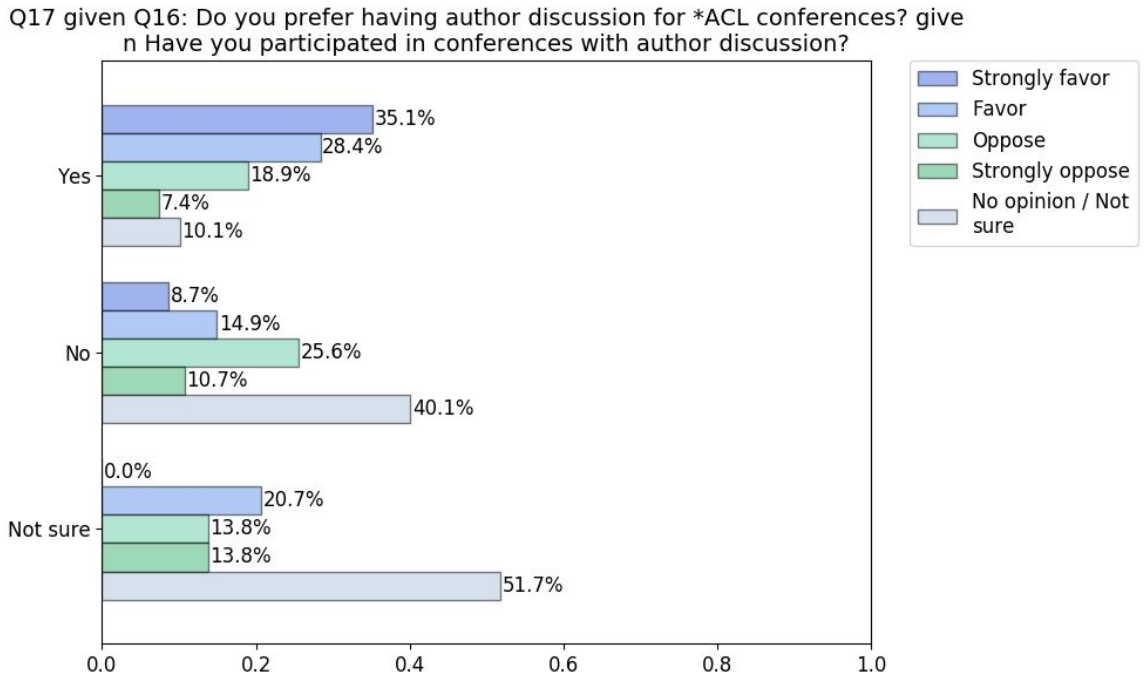
Author Discussion

Q17: Do you prefer having author discussion for *ACL conferences?

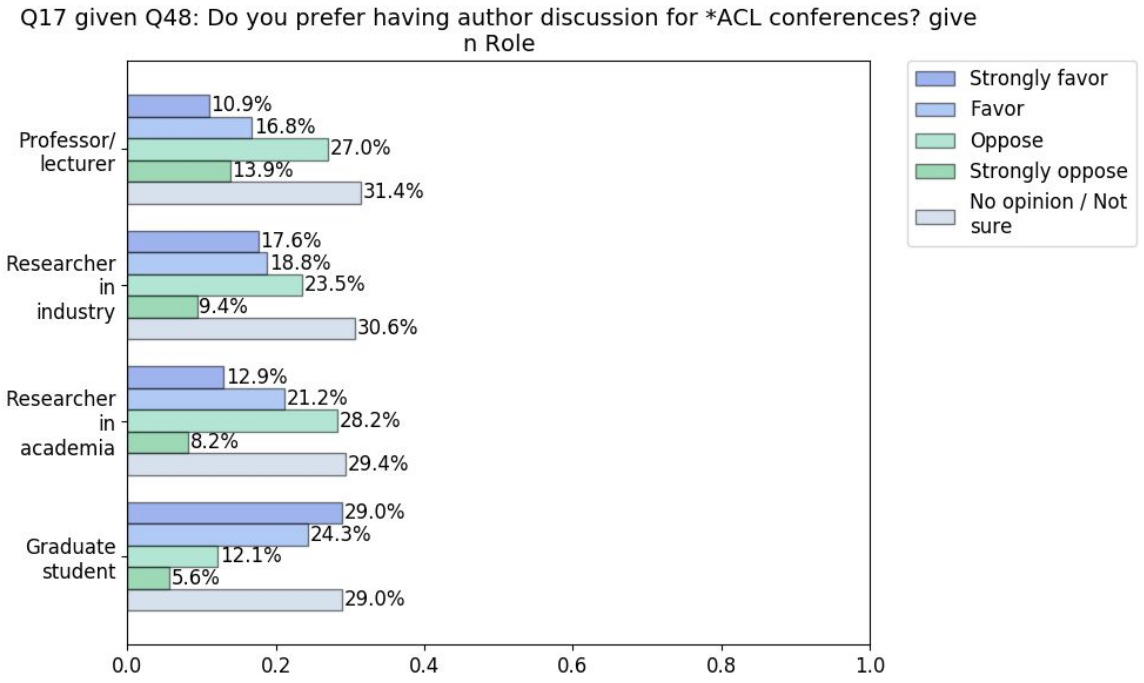
Opinions were nearly evenly split on author discussion, with 37% in favor, 32% opposed, and 30% with no opinion.



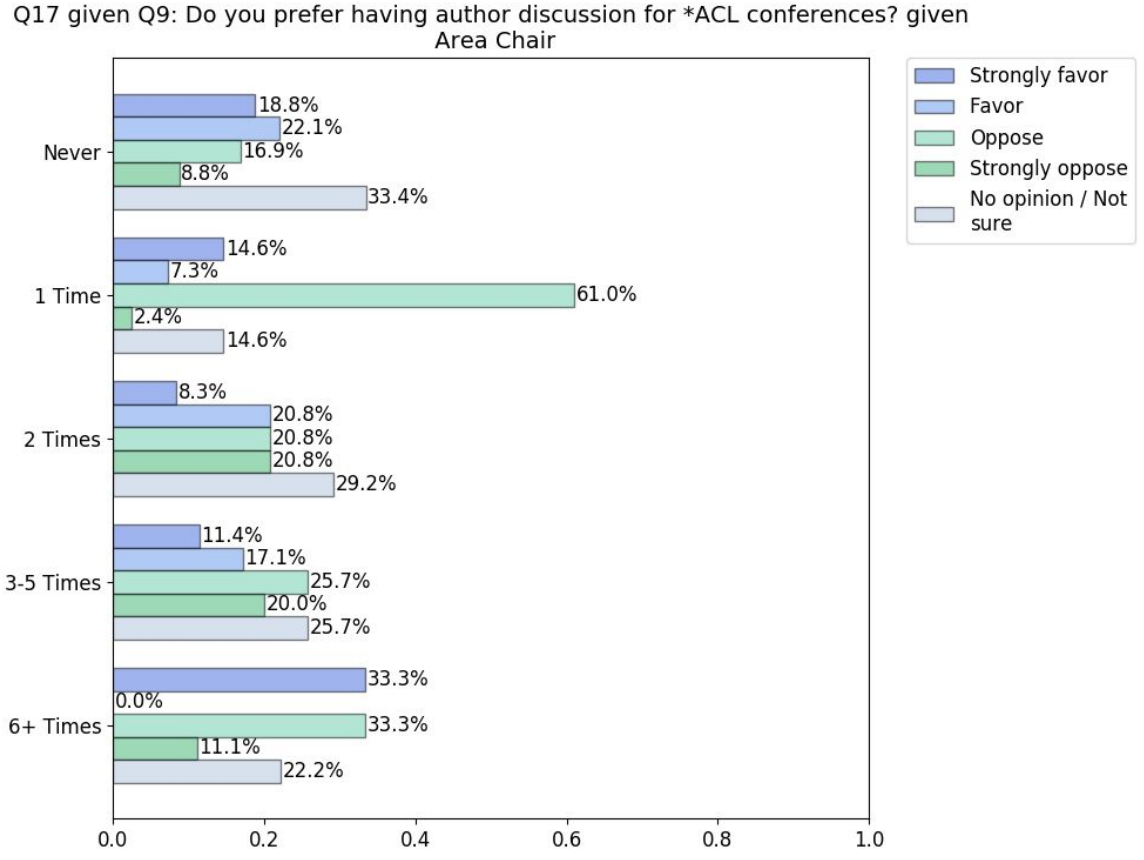
However, there was a strong dichotomy between respondents who have submitted to a conference with author discussion. Those who have experience with the format were largely in favor 63%-26%, while those who did not have experience were more opposed 24%-37% (with 40% marking no opinion/not sure).



Similarly to opinions about author response, faculty were less in favor of author discussion and graduate students were more so, with researchers in academia and industry falling in the middle.



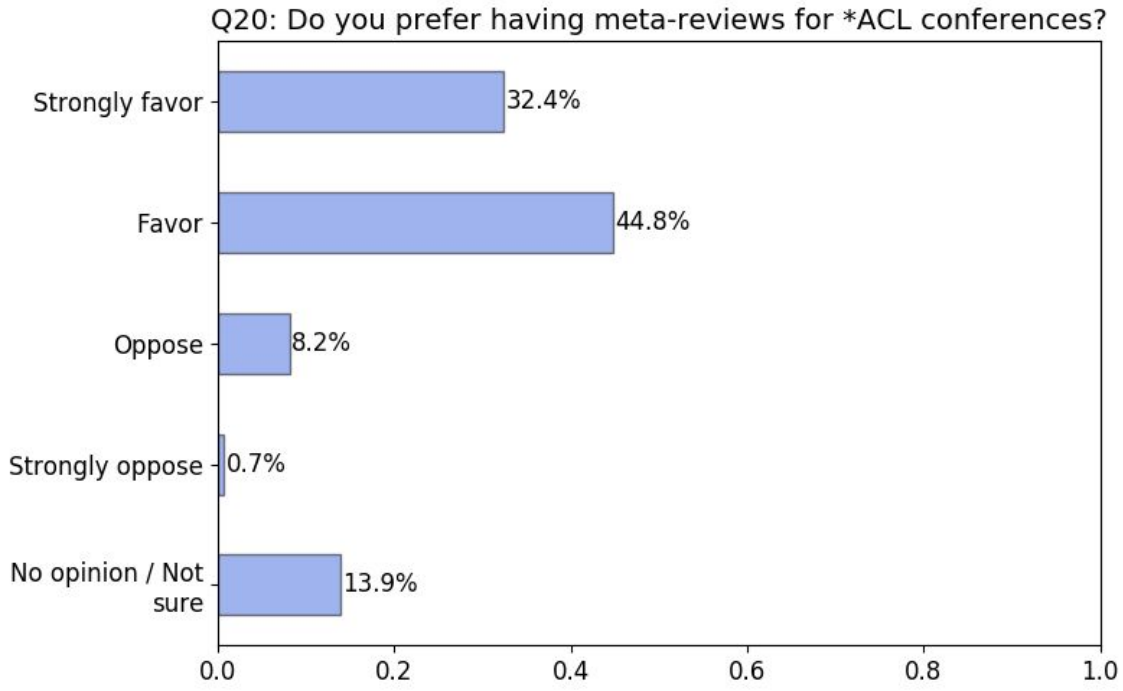
Support for author discussion also differed based on experience as an area chair, with those with no experience as an area chair being more in favor, and those with experience as an area chair being more opposed.



Meta-Reviews

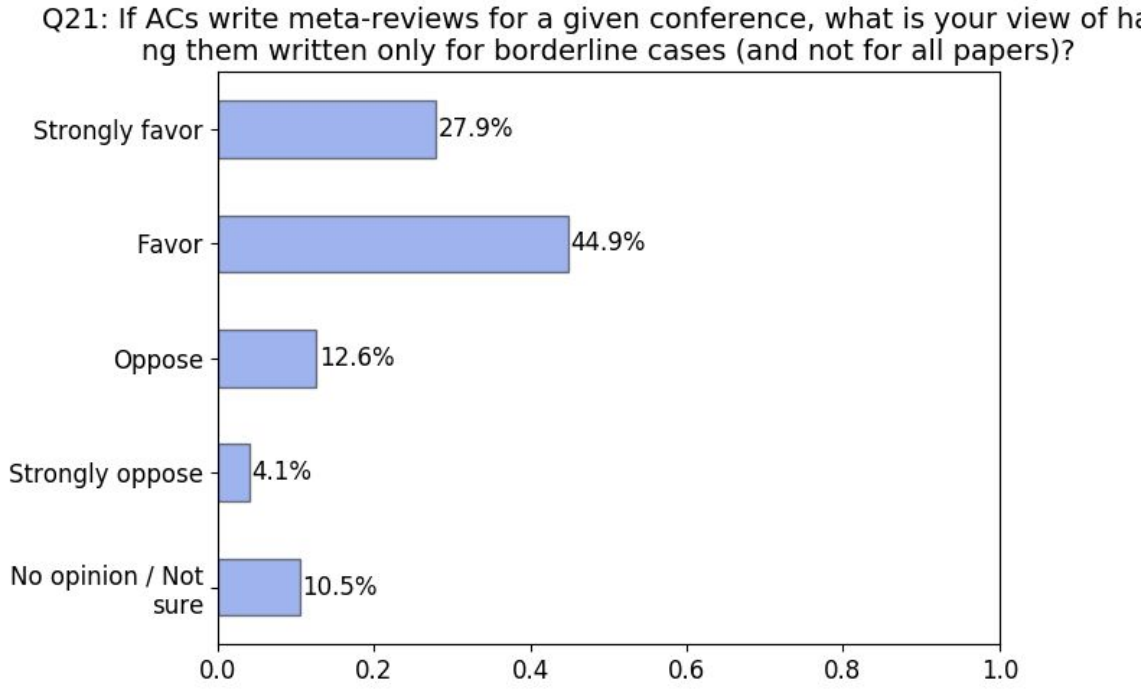
Q20: Do you prefer having meta-reviews for *ACL conferences?

There was strong across-the-board support for writing meta-reviews...

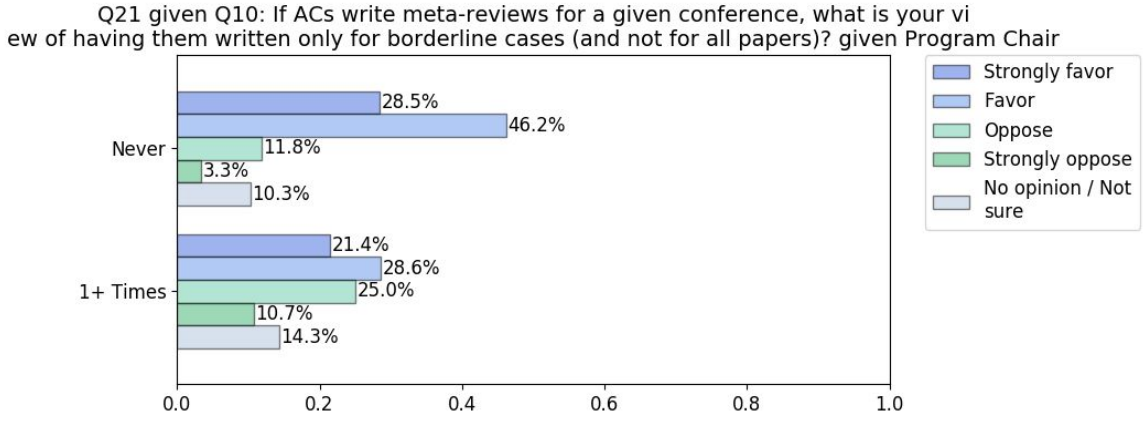


Q21: If ACs write meta-reviews for a given conference, what is your view of having them written only for borderline cases (and not for all papers)?

...and the majority agreed that it was likely sufficient to write them in only borderline cases.



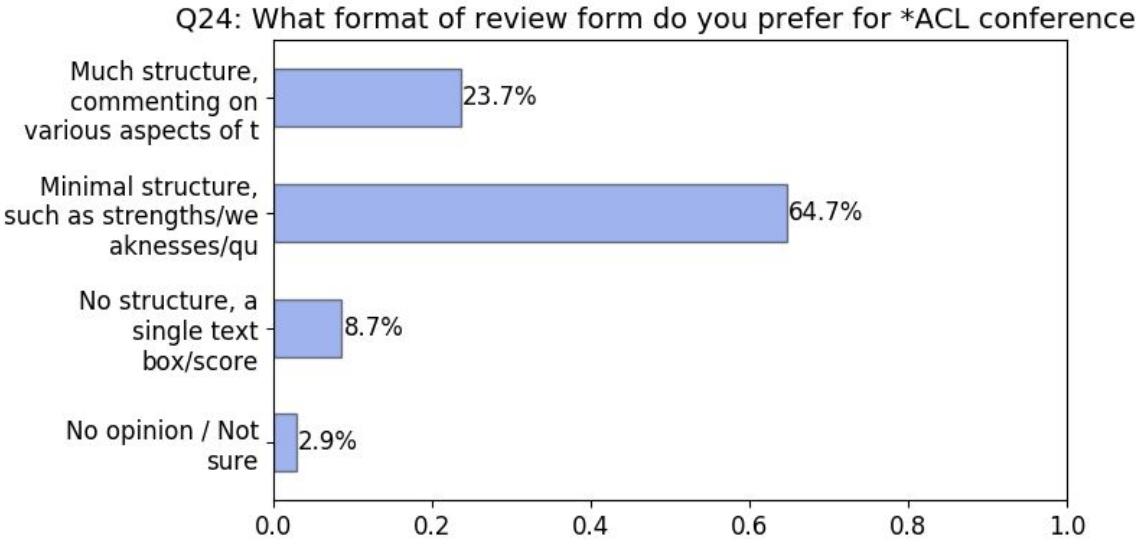
However, those with experience as a program chair much more strongly preferred having meta-reviews be written in all cases, although opinions were still split 50%-35%.



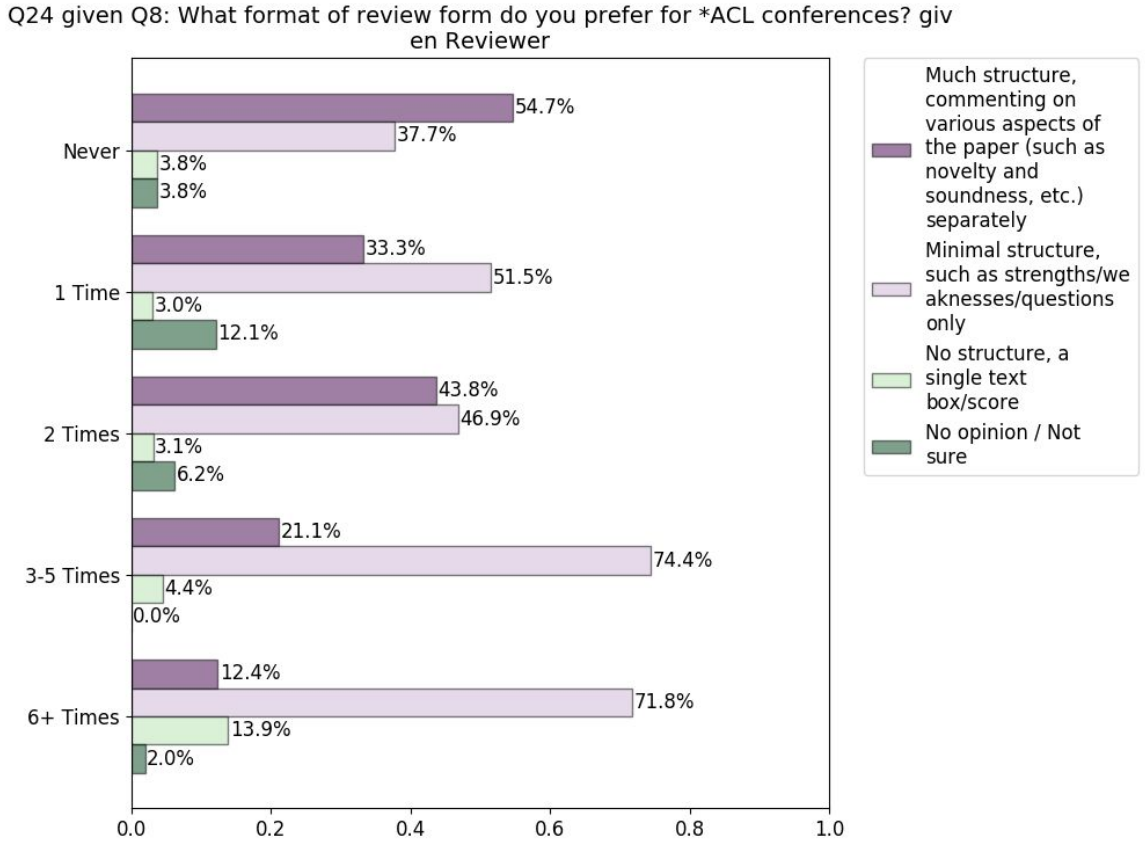
Structured Review Forms

Q24: What format of review form do you prefer for *ACL conferences?

A significant majority of 65% supported a minimal amount of structure in review forms, although some preferred more structure.



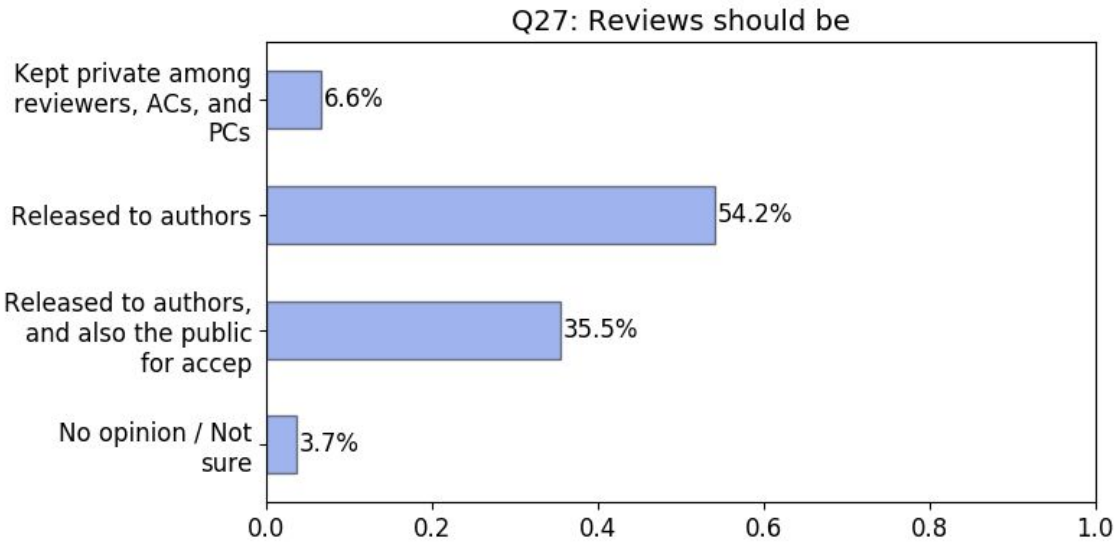
There was a clear trend that more experienced reviewers preferred less structured review forms, and a similar trend held with respect to experience as an author or area chair.



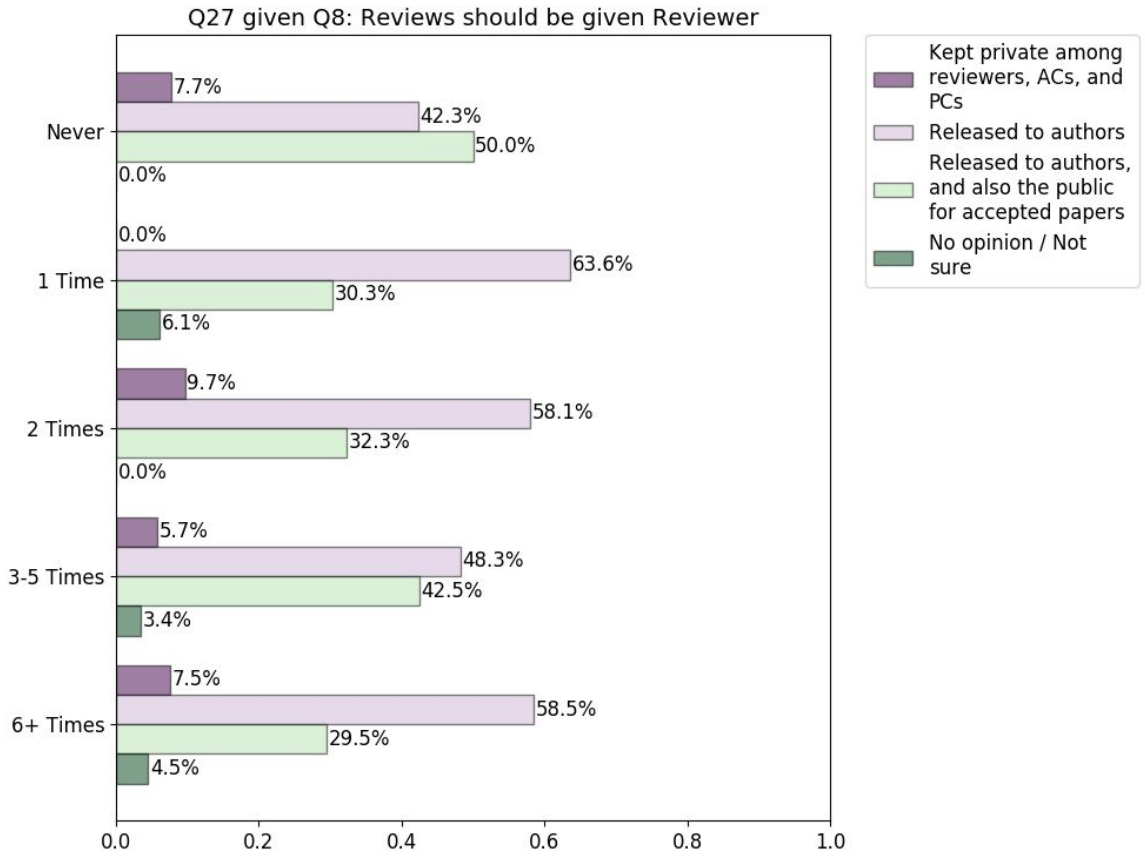
Transparency (of Reviews, Review Discussion, Meta-Reviews)

Q27: Reviews should be

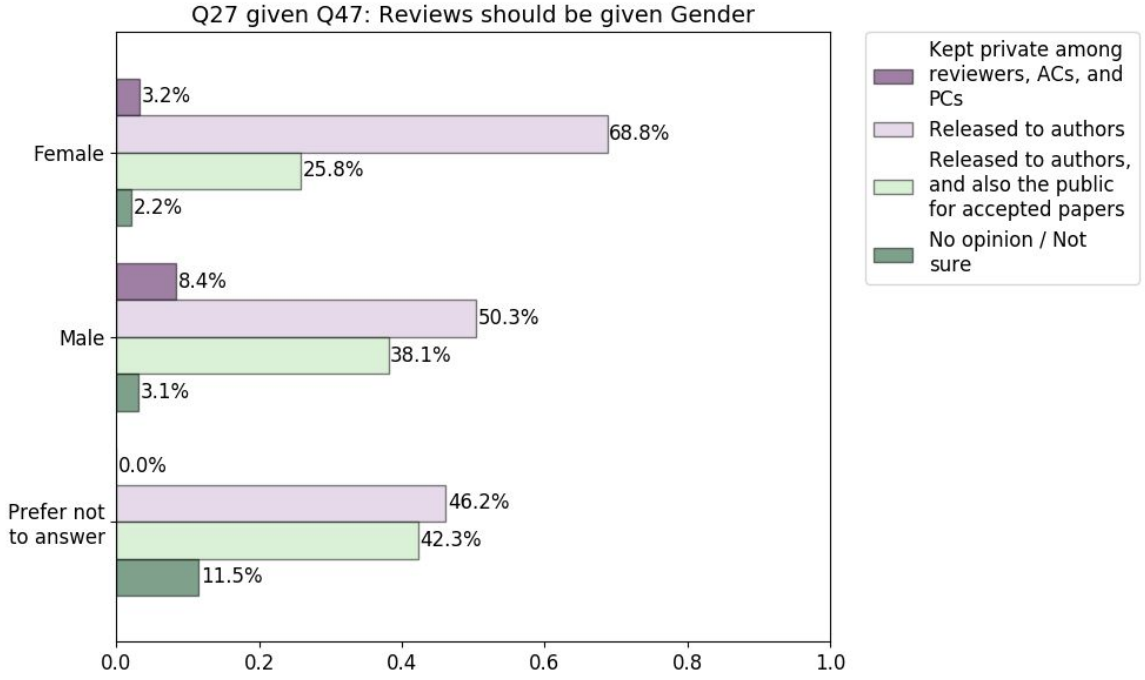
A majority of 54% preferred that reviews be released only to authors, but there was a significant minority of 36% who preferred public release of reviews.



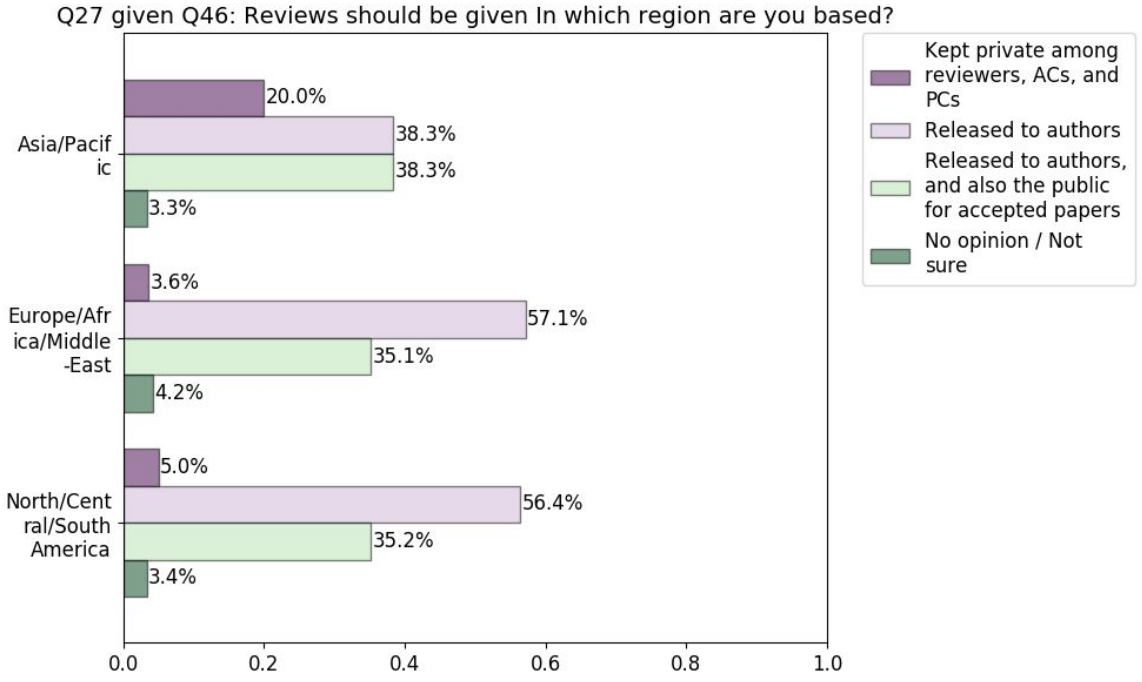
In general, those with no experience reviewing preferred public release of reviews, while those with some experience reviewing were against public release.



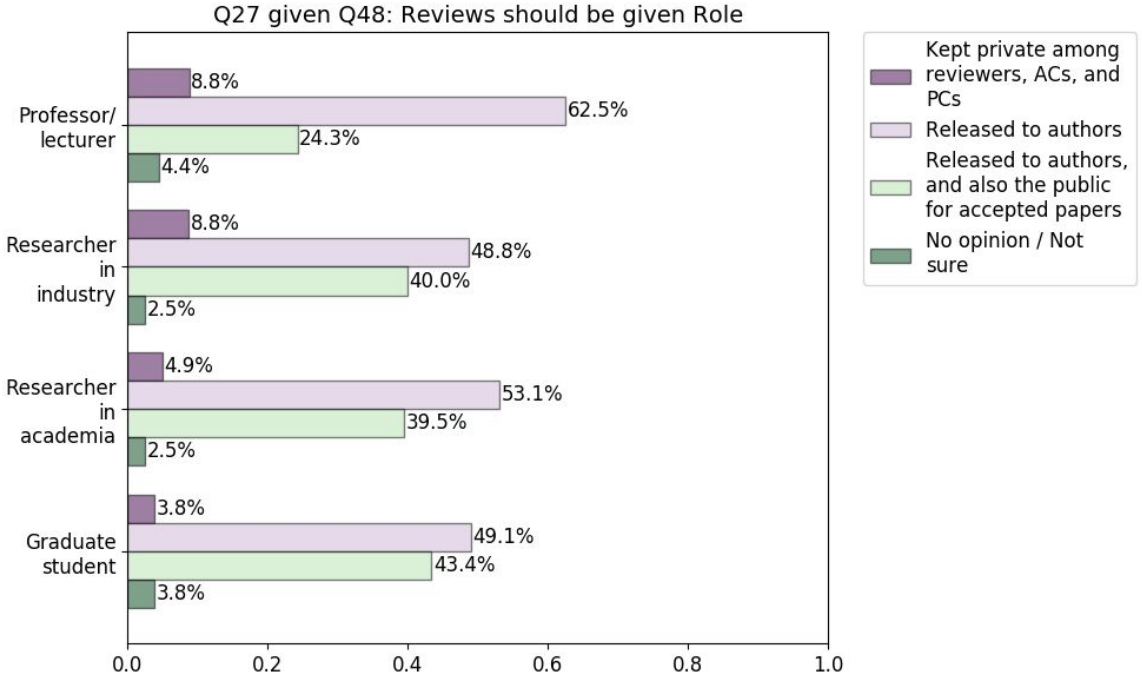
Public release of reviews was one area where there was a significant gap between genders, with female only 26% of female respondents preferring reviews be publicly released, compared with about 40% for those who were male and preferred not to state their gender.



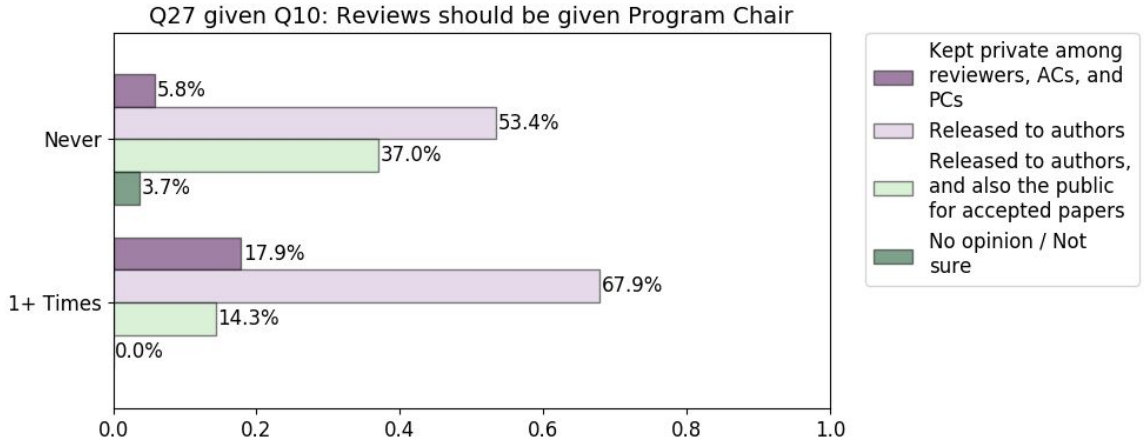
On the other hand, there were no regional differences in preferring public release, but a significant 20% of researchers from the Asia/Pacific region preferred that reviews not be released even to authors.



Researchers and graduates students were nearly evenly split on the question of whether to publicly release reviews, but professors/lecturers were strongly for releasing only to authors, 62%-24%.

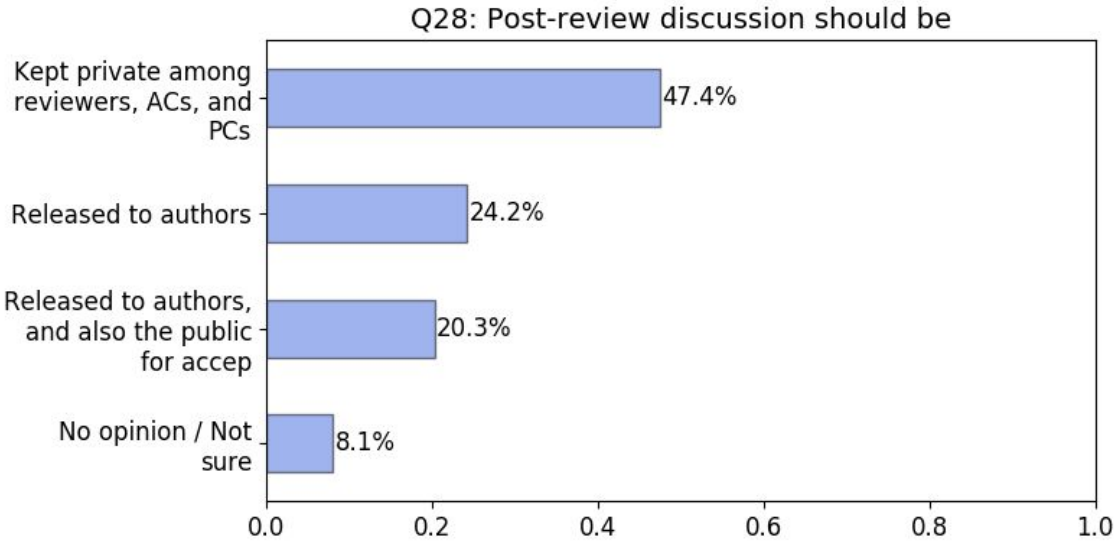


Finally, those with program chair experience were even more strongly against public release of reviews, with only 14% agreeing. 18% were against even releasing reviews to authors.

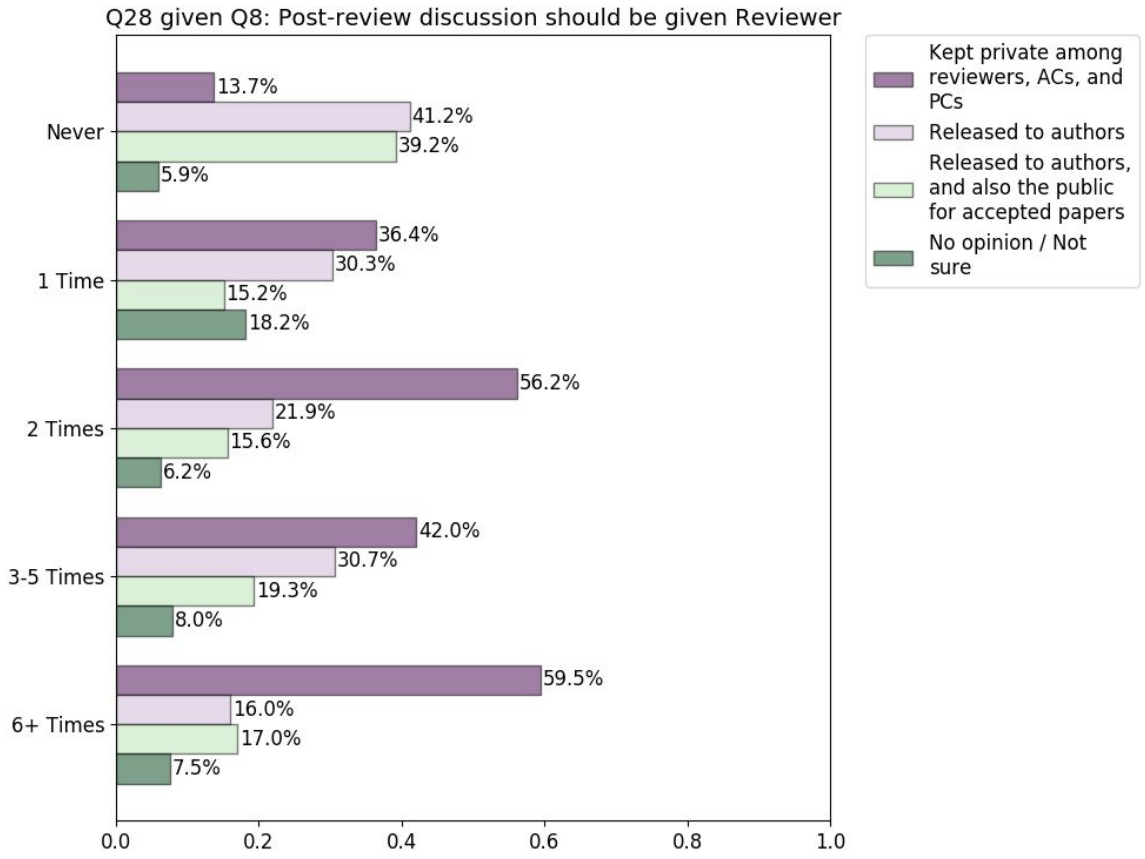


Q28: Post-review discussion should be

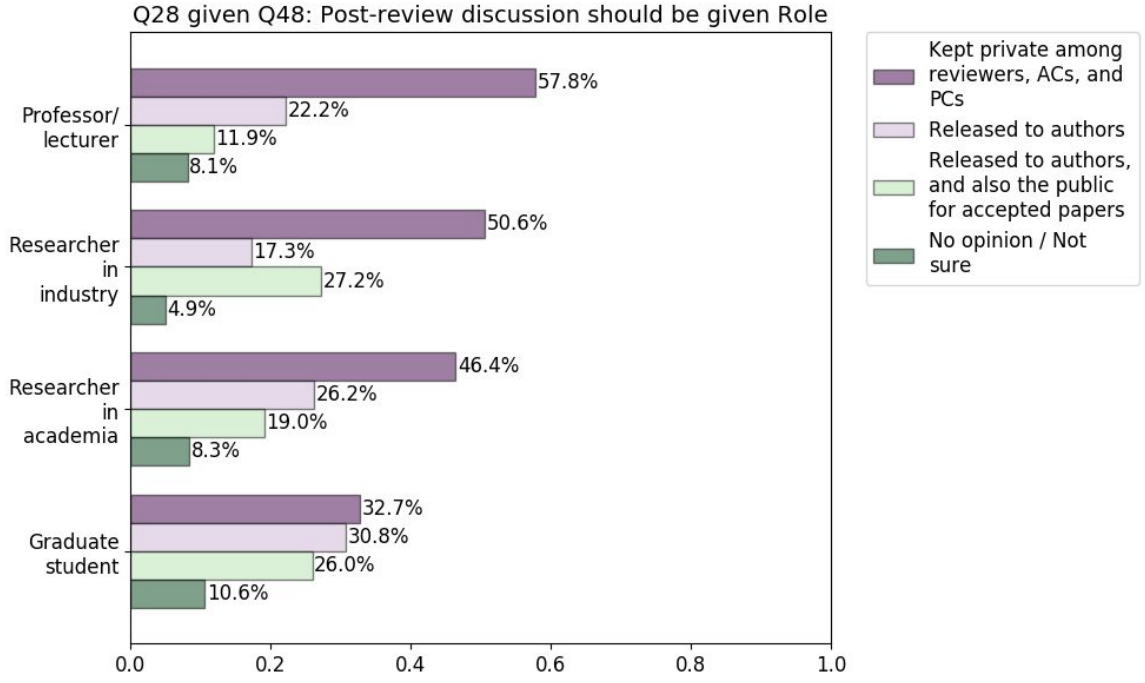
There was significant disagreement in whether post-review discussion should be released to authors, with about half of of those who responded with an opinion preferring that it not, and the remaining half split between release to authors or public release.



Those with greater amounts of review experience preferred that post-review discussion be kept private, while those who had never reviewed strongly preferred that discussion was made available to authors or the public.

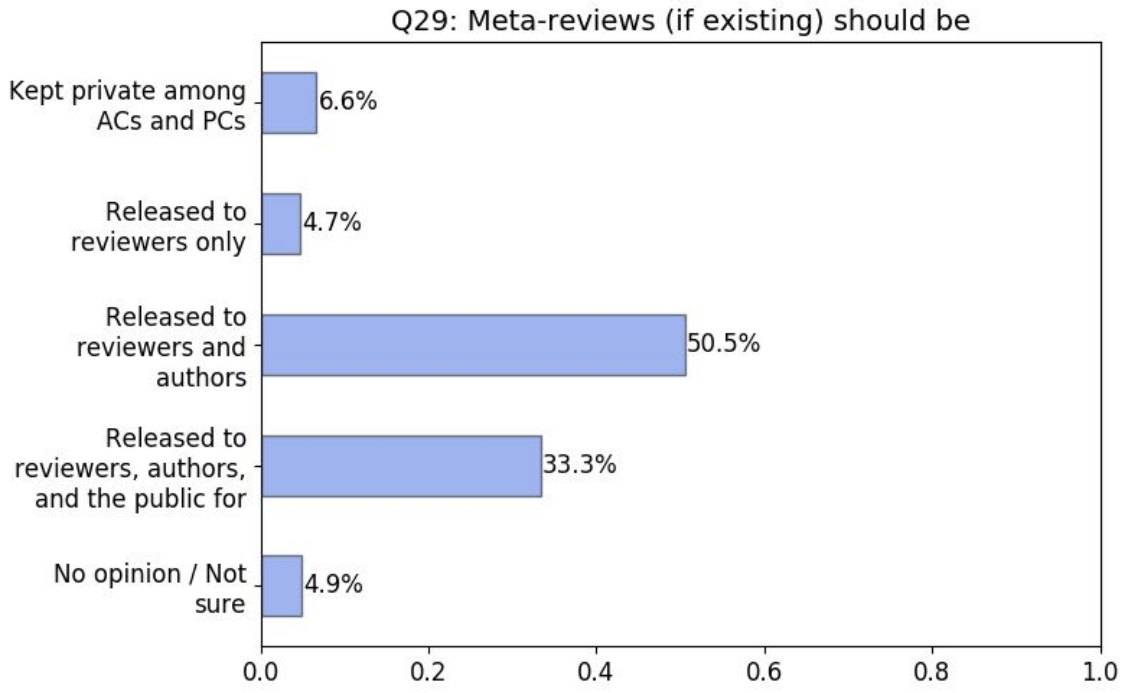


Graduate students were strongly for release of post-review discussion to at least the authors 57%-33%, while professors/lecturers were against by almost exactly the same margin, 34%-57%.

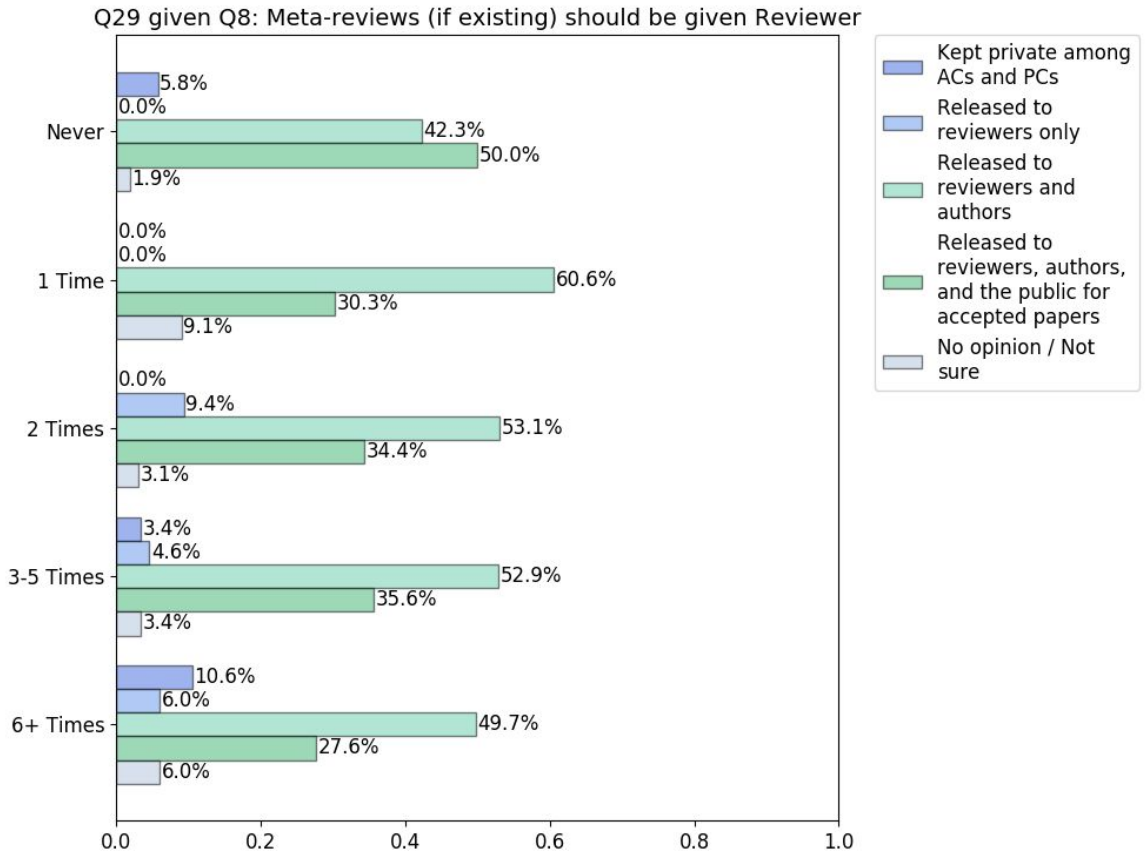


Q29: Meta-reviews (if existing) should be

There was strong support in releasing meta-reviews to at least authors, with 83% agreeing. 33% were also in favor of releasing to the public.



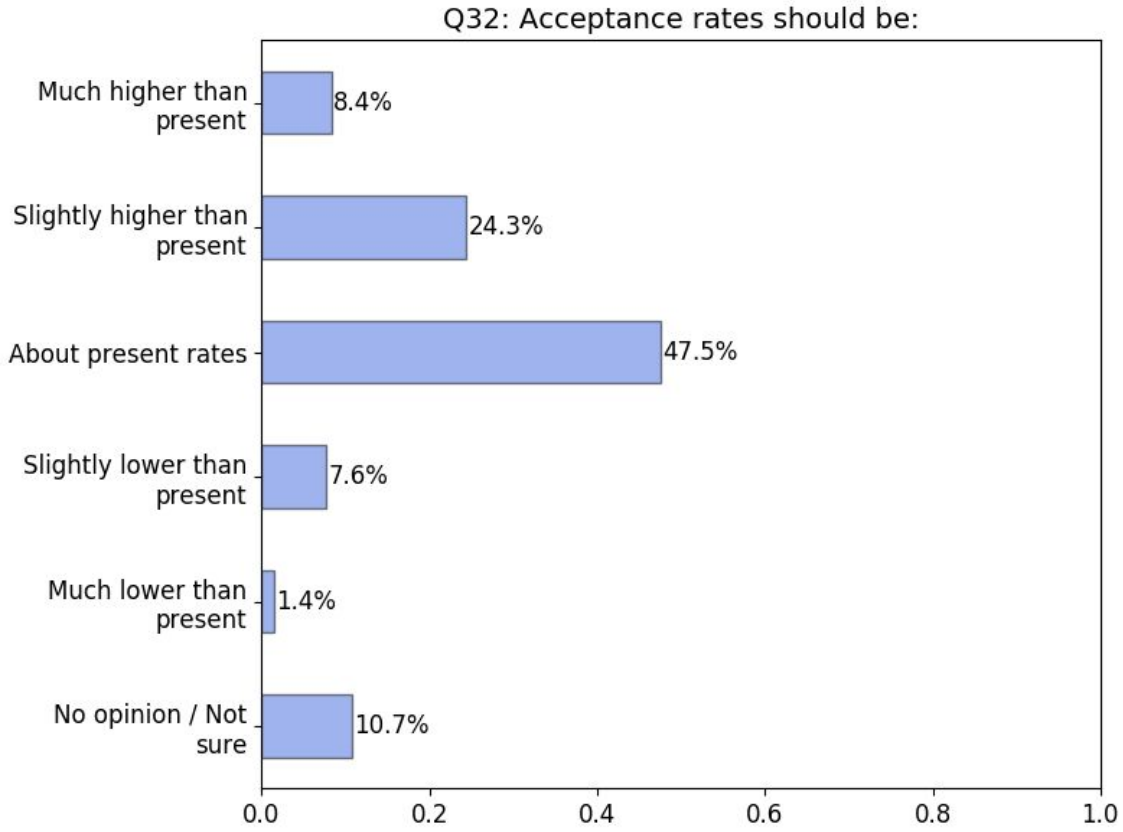
In particular, those who had never reviewed preferred public release of meta-reviews to releasing only to authors, while those who had previously reviewed were the opposite.



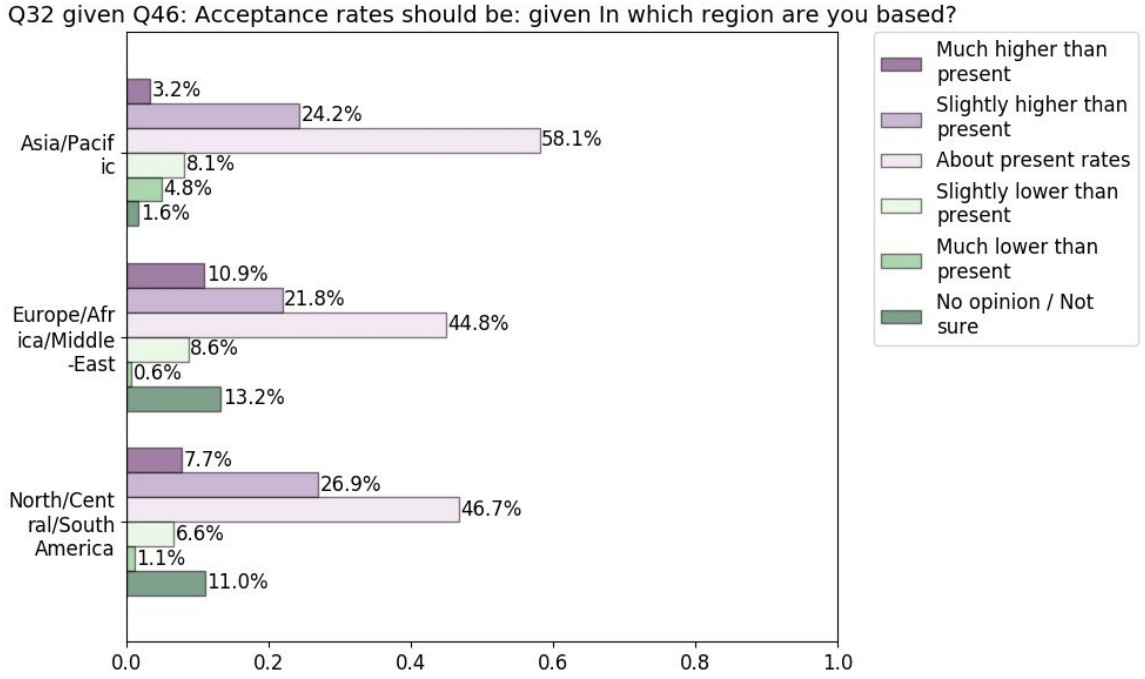
Acceptance Rates

Q32: Acceptance rates should be:

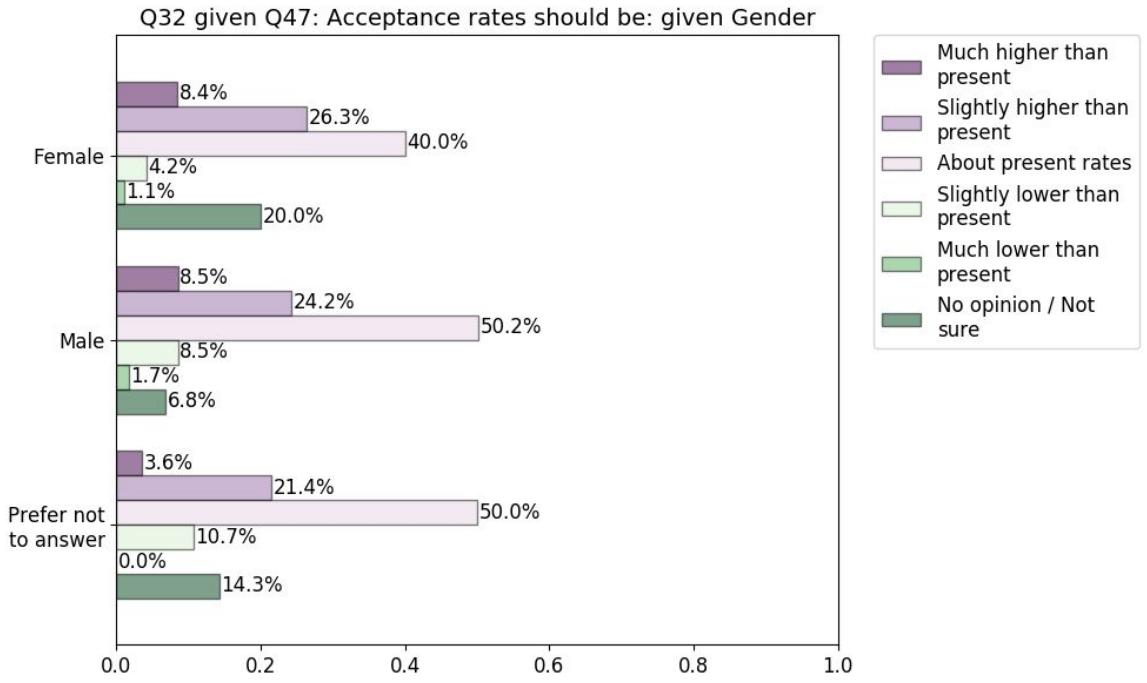
47% were in favor of keeping the status quo with respect to acceptance rates, while 32% were in favor of increasing, 9% were in favor of decreasing.



Respondents from the Asia/Pacific region were much less likely to support large increases in acceptance rates, and were more likely to favor large decreases in acceptance rates.

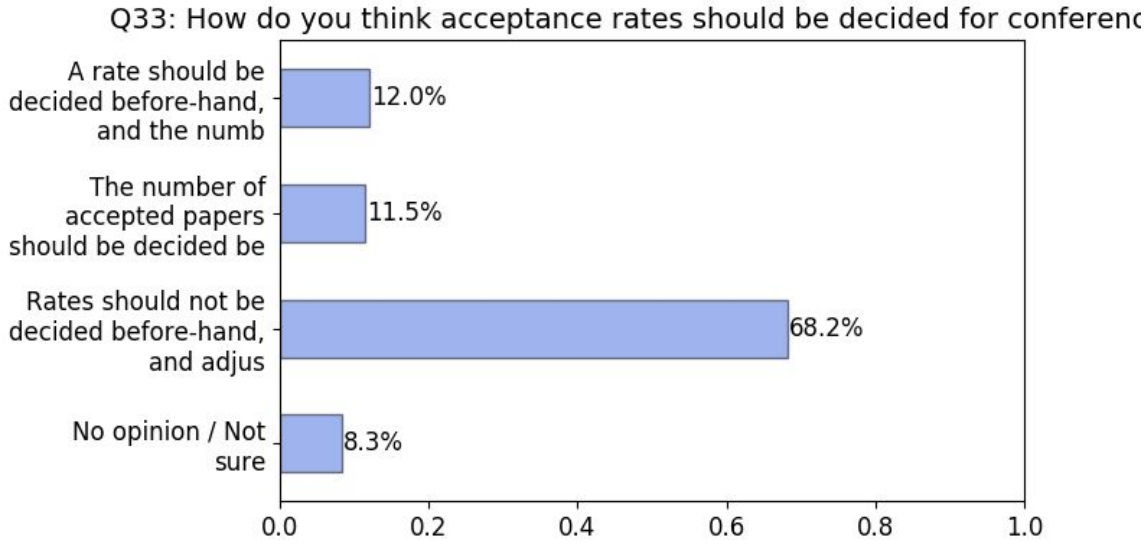


Male respondents were more opinionated about acceptance rates overall, and of those who responded, female respondents were slightly more in favor of increasing acceptance rates.

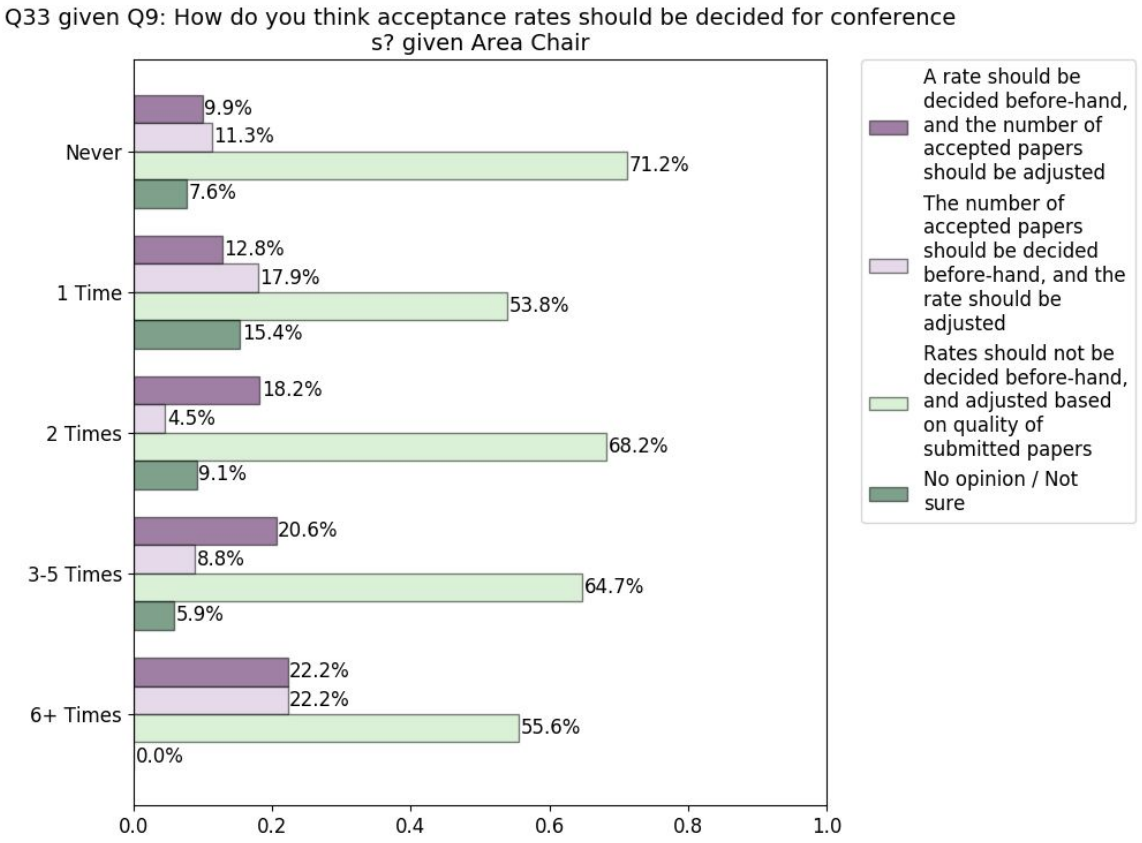


Q33: How do you think acceptance rates should be decided for conferences?

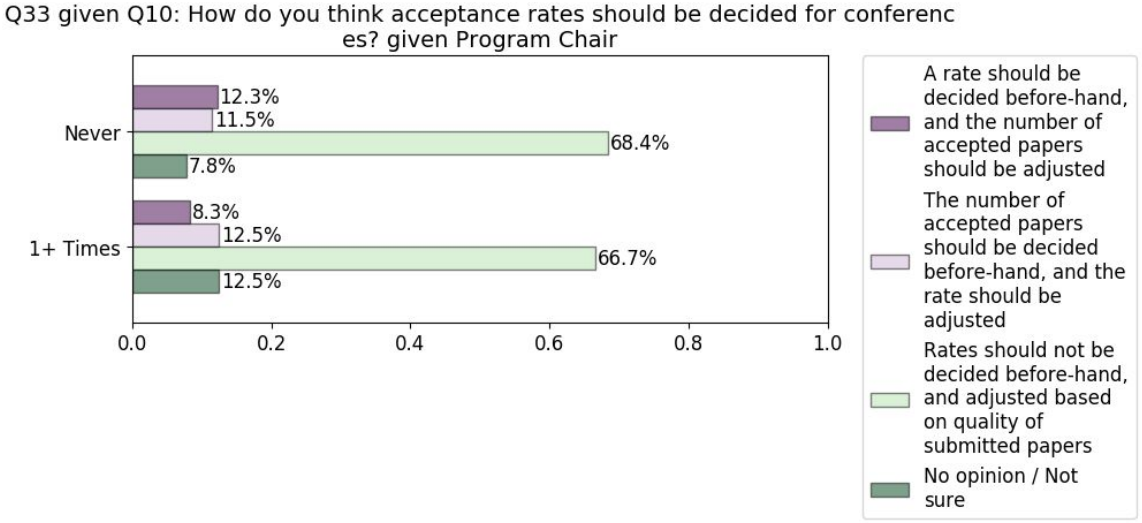
A large majority of respondents preferred that acceptance rates not be decided before-hand, but rather post-hoc based on quality of papers.



This opinion was slightly less popular among those with significant area chairing experience...



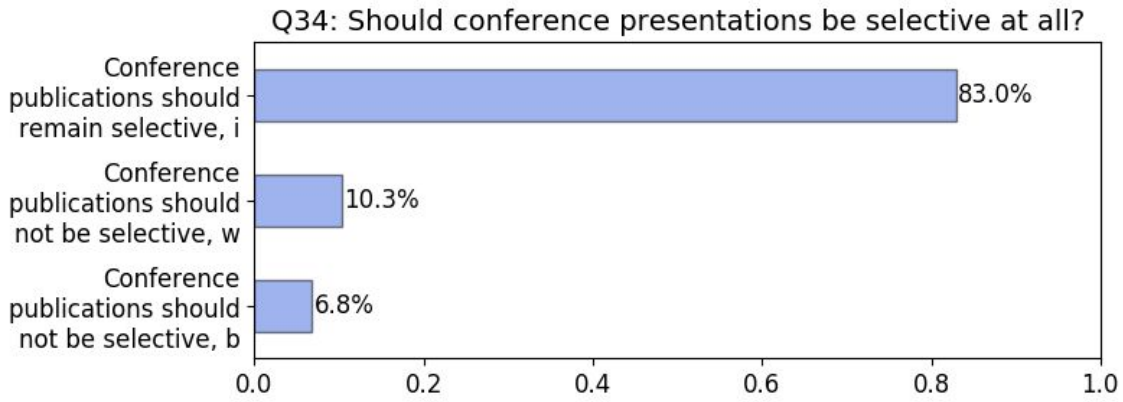
... but there was a remarkable lack of difference in opinions between those with or without experience as a program chair at a major conference.



Selectivity of Conference Publications

Q34: Should conference presentations be selective at all?

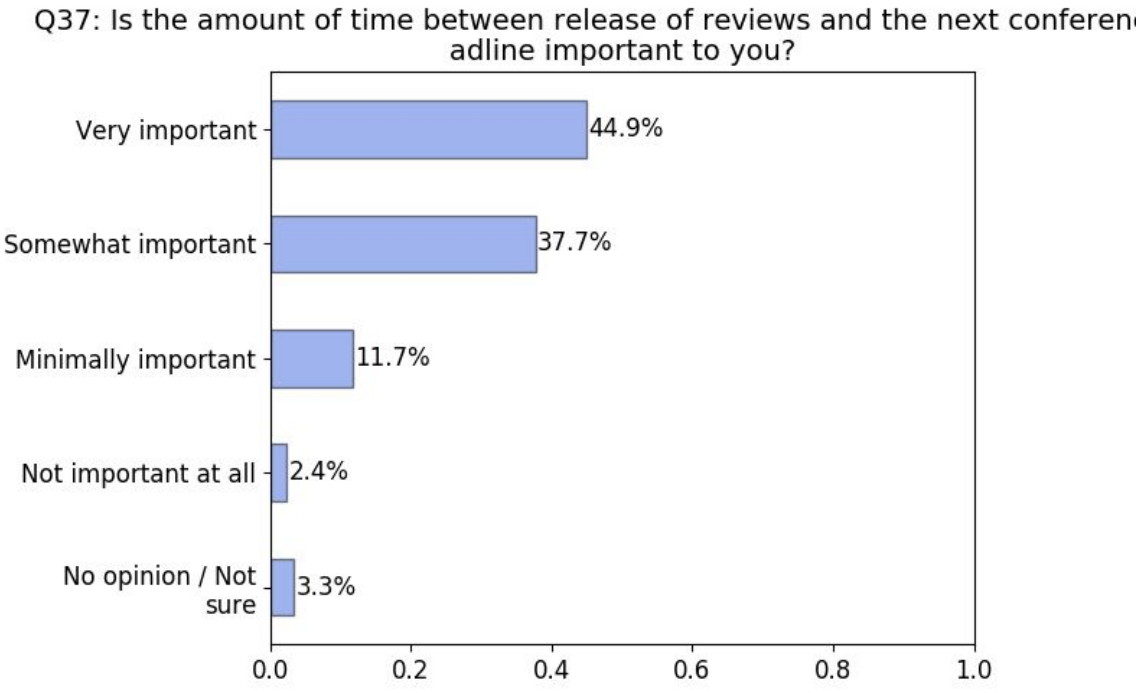
There was strong overall support for conferences remaining selective, and no demographic group had significantly divergent opinions about this.



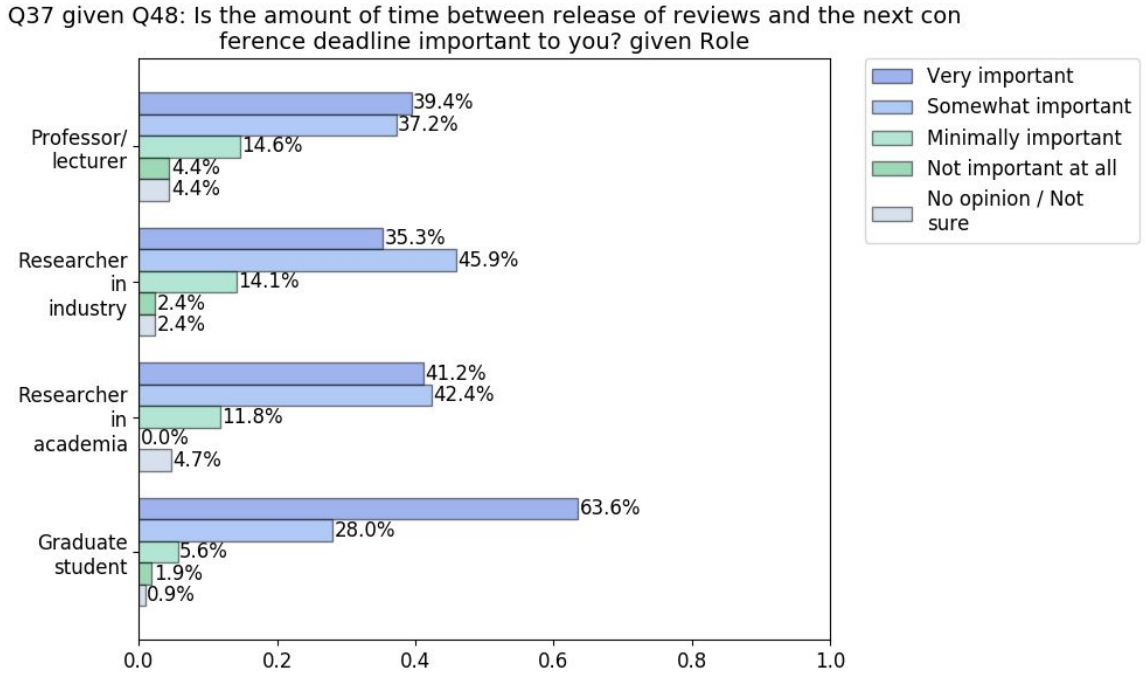
Timing of Review Release

Q37: Is the amount of time between release of reviews and the next conference deadline important to you?

The great majority said that the timing between review release and the next conference deadlines was at least somewhat important, 83%-14%.

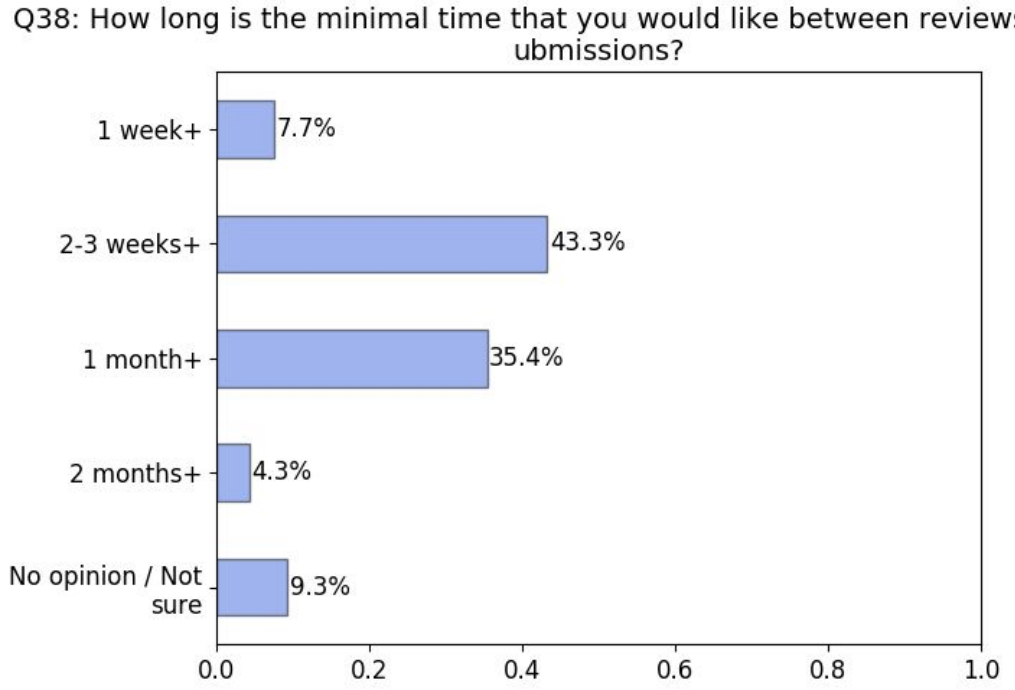


This was particularly the case for graduate students, where 64% said this timing was very important.

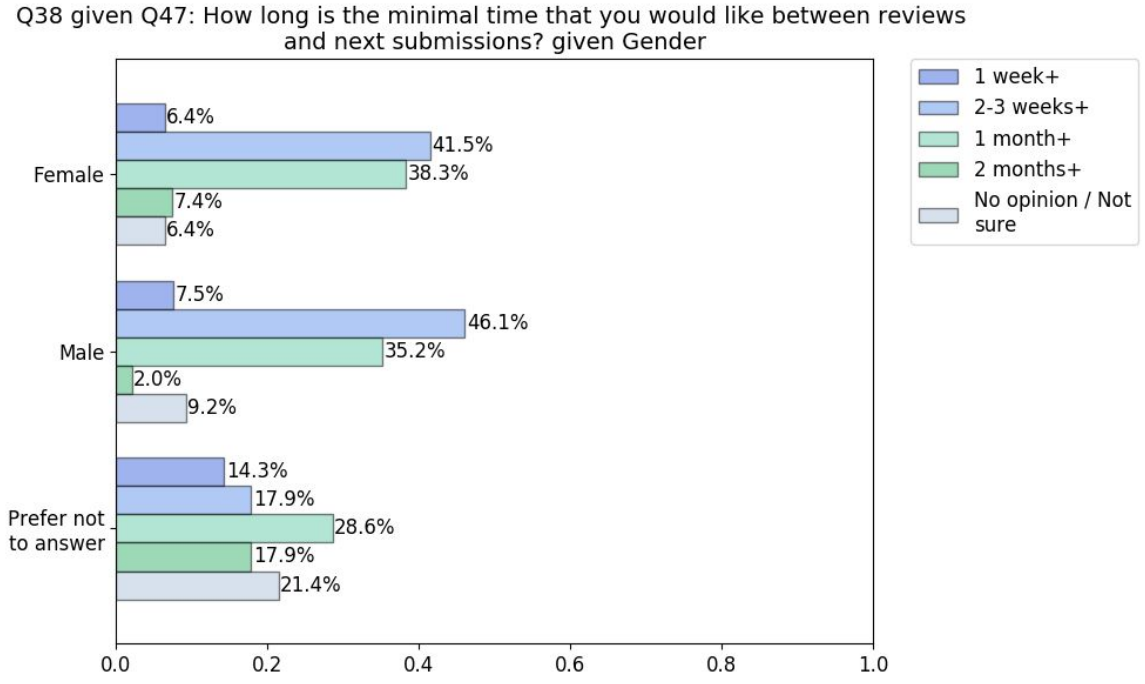


Q38: How long is the minimal time that you would like between reviews and next submissions?

The amount of time that people would prefer was at least 2-3 weeks, and many 39% said at least a month was preferable.



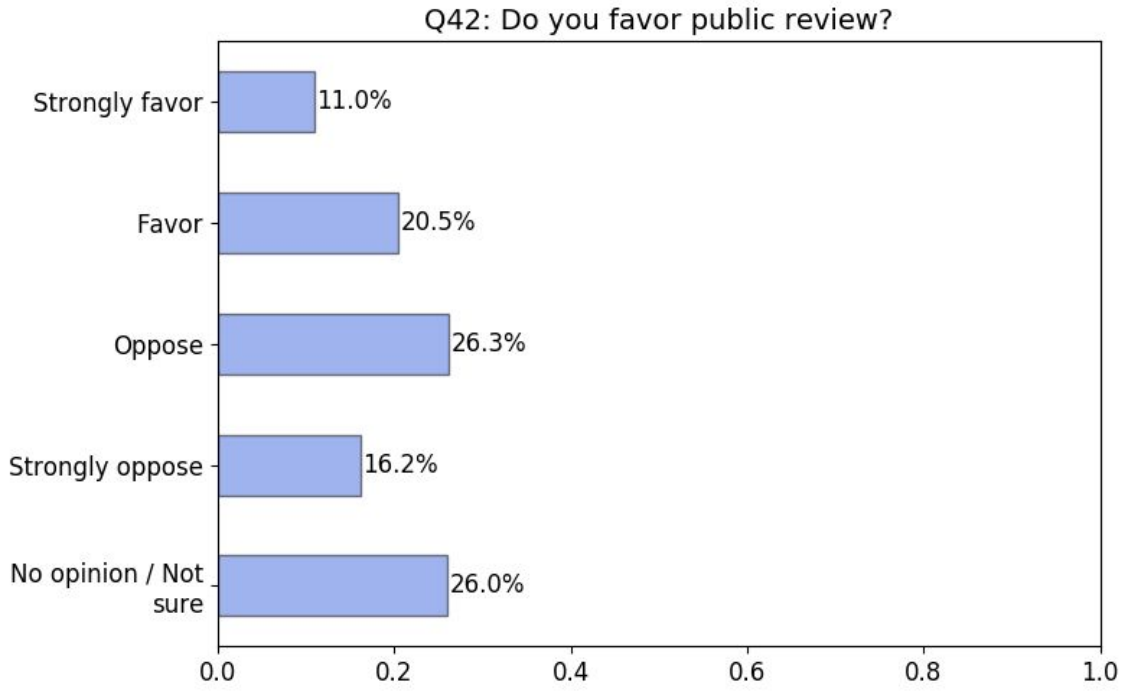
Female respondents slightly preferred having more time between review release and the next deadline, and those who preferred not to state their gender particularly preferred longer reviewing cycles.



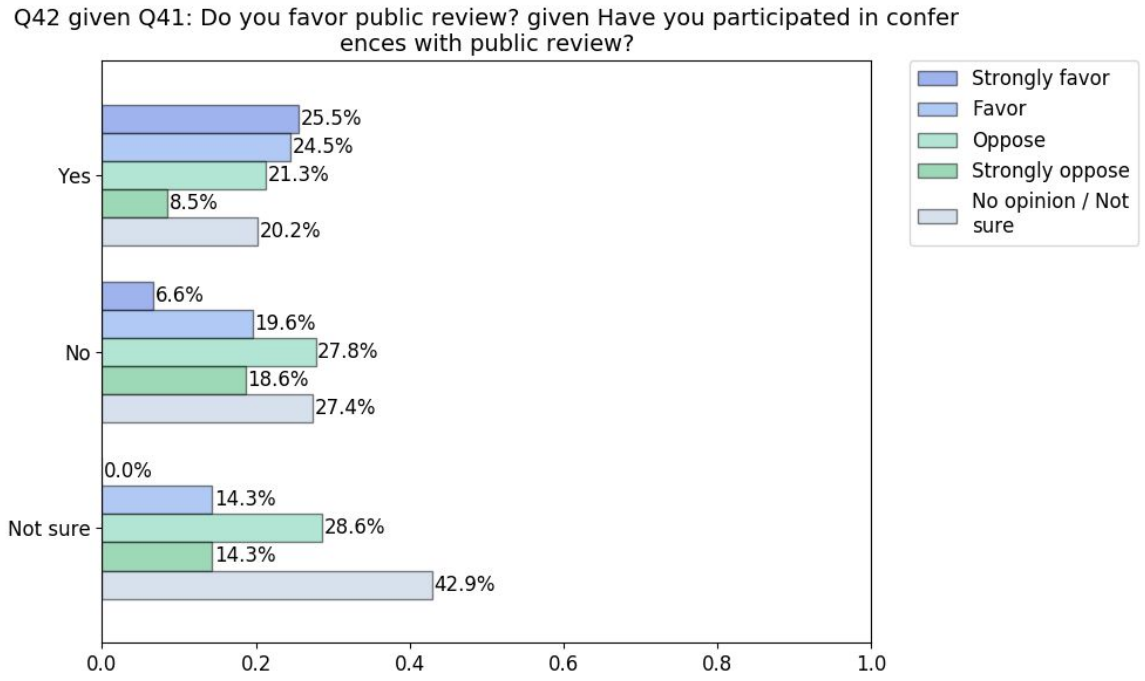
Public Review

Q42: Do you favor public review?

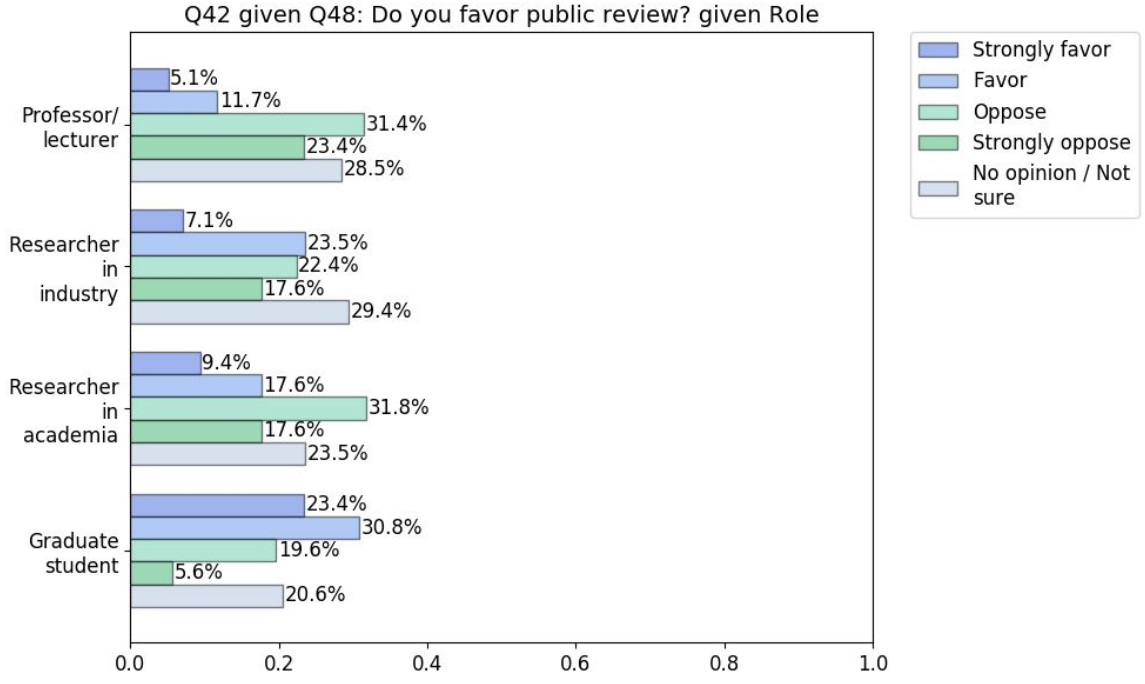
Opinions about public review were mixed and tended to be strong, with 32% in favor and 42% opposed, and 37% holding strong opinions.



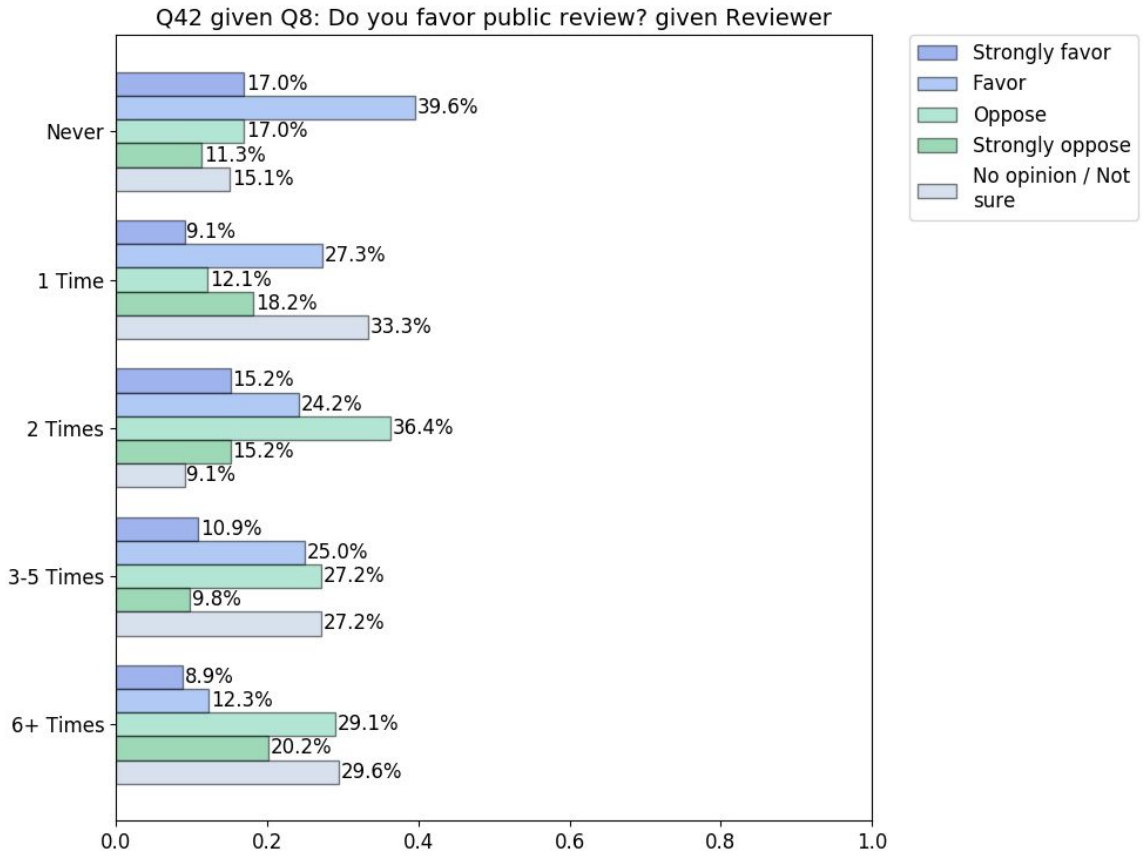
Those who had experience participating in conferences with public review tended to be in favor 50%-30%, while those who did not have such experience were opposed 27%-47%.



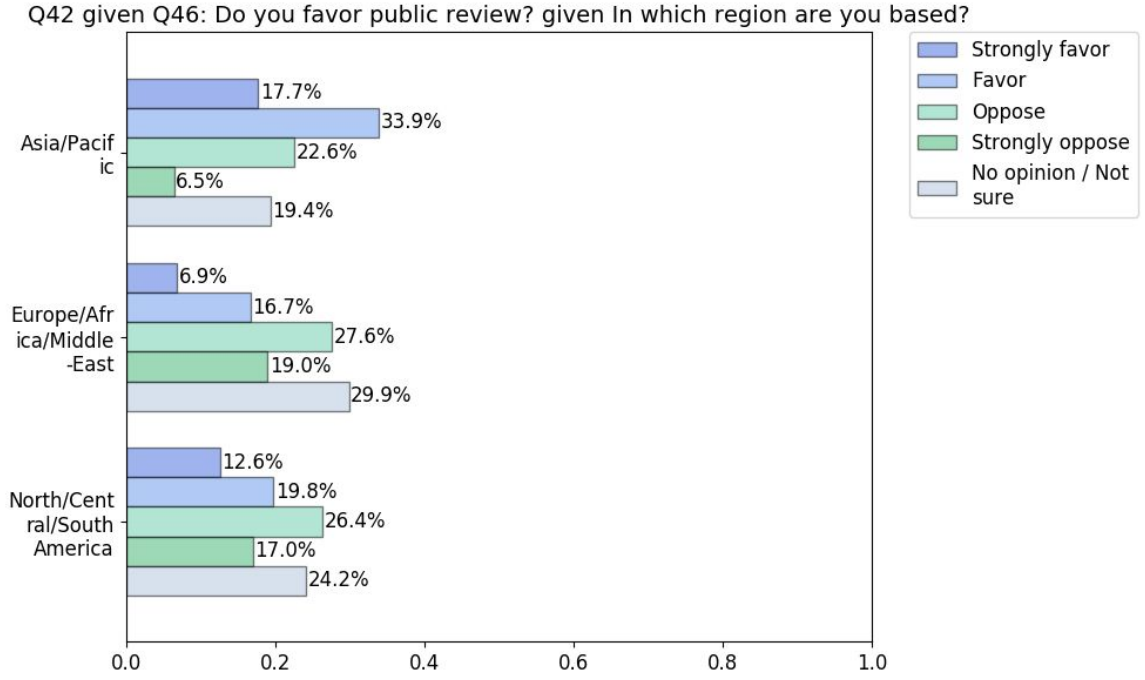
Professors/lecturers were opposed to public review 17%-54%, researchers were in the middle with 31%-40% and 27%-50% for those in industry and academia respectively, and graduate students were in favor 54%-26%.



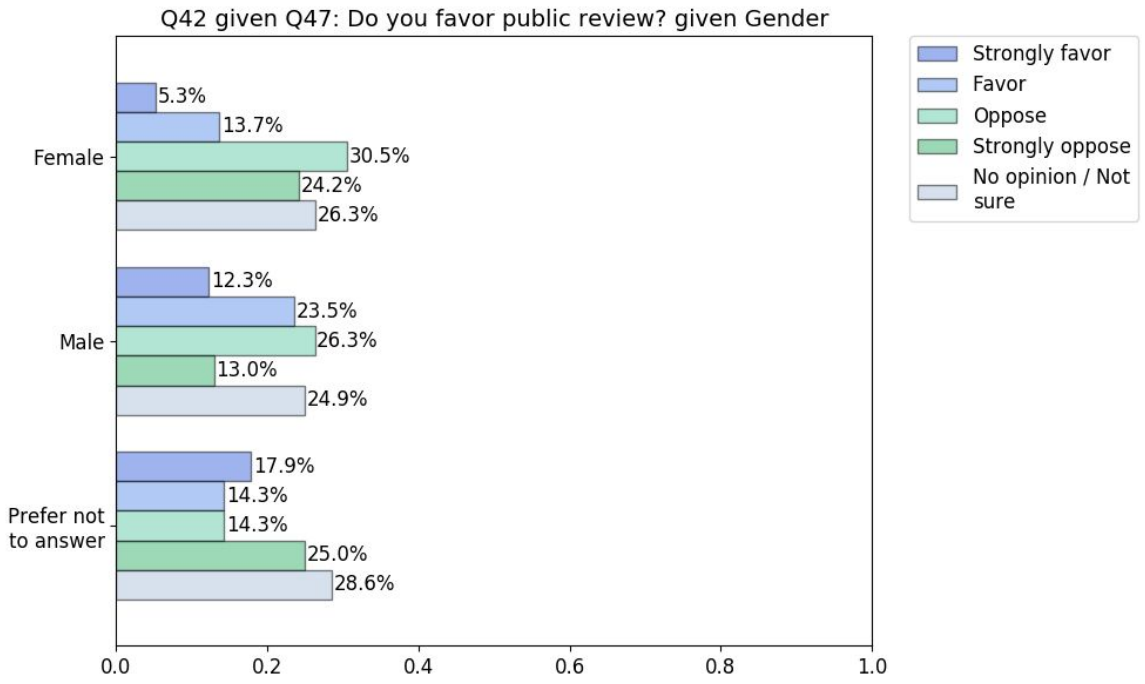
In general, support for public review tended to be inversely correlated with reviewing experience.



Those in the Asia/Pacific region were in favor of public review 52%-29%, while Europe/Africa/Middle-East and North/Central South America were opposed 24%-47% and 33%-43% respectively.



Female respondents were less likely to support public review than male respondents, 19%-55% vs. 36%-39% respectively. As noted in the demographics report, however, female respondents were only half as likely as male respondents to have participated in a conference with public review, and the effects of gender and experience with public review may be conflated in these results.



Detailed Results: Open Comments, Categorized

The open-ended responses were manually categorized for each question and appear below.

Author Response

Positive: Response is useful for ACs/PCs

- The empirical number of papers whose decisions change can be an underestimate of the impact of responses, since the AC can use the response when deciding which reviewers to trust when making recommendations.
- Based on rebuttal comments, at least Area Chairs can see the author's points if not reviewers.
- As an AC, author responses have been critical to my work and enabled significantly higher quality decision compared to what I observed without responses.

- I find that author responses are most effective when addressed to ACs/PCs, rather than reviewers. Setting it up to be like that explicitly also relieves time pressure somewhat.
- It's a useful tool for area chairs to know if a reviewer is completely wrong. With inexperienced reviewers, it's helpful to have a framework for discussion.
- It is probably only valuable to the AC/PCs if at all, or to ask clarification of very specific aspects (unlike a general discussion of why the reviewers are wrong).
- A lot of the criticism of author response relates to impact on scores and impact on changing the opinions of reviewers, but the main place where I personally find it to be influential is as an Area Chair.
- Helpful for ACs, good for authors to critically engage with reviews as well
- I used to be against it, but after serving as a program chair, came to see its value. Even if no reviewer changes their review or score, the AC and PC see the responses and they are sometimes very helpful.
- ACs need to be guided to spot dodgy reviews
- In the review process, I think author response is even more critical to the area chair than the reviewers themselves, it allows the AC to take to know and take into account the authors reactions/response to the reviews when making the final recommendation.
- I was an area chair for AKBC and ACL. AKBC had author responses. Although the reviewers did not care much about the author response, as an area chair, I find the response very useful especially in the borderline cases to decide whether to accept or reject.
- I find it extremely useful as an AC. AR period can run in parallel with discussions between reviewers, and it would be great if we could support interactive discussions rather than one-shot responses (moving to openreview?)
- They are valuable for ACs, not for reviewers.

Positive: Response is useful for Reviewers

- I saw more value in the author response as a reviewer. I was usually more sure about my assessment of the paper after reading the response. However a discussion with other reviewers plays a similar role for me. As a paper author I usually had a feeling that the response does not make much difference.
- As a reviewer, a nontrivial fraction of the author responses have clarified parts of the paper and caused my to change my review. Admittedly, it's definitely less than half, likely even fewer, but it is valuable, especially as the reviewing load grows and I have less time to dedicate to each review I write.
- As a reviewer, it is nice. Sometimes a paper has very confusing descriptions and we have to reject it while considering the worst case due to it. Author response can solve it.
- As a reviewer for recent conferences without author response (NAACL and ACL), I often would have liked to ask clarifying questions to the authors, which would have helped me write a better review. Without an author response period, there is no way to do so.
- As an author, I feel that author responses rarely change the outcome. As a reviewer, I think author response is a good source of additional information but again it does not change my overall opinion in most cases.

Positive: Having review feedback earlier due to author response is an advantage

- As an author, I'm not sure if I've ever swayed reviewers enough to change the final decision, but the early feedback sometimes enables me to withdraw a paper in time to send it to a more appropriate venue.
- I feel the author response is VERY useful to get feedback without the need of waiting for them until the final decision.
- This approach also has the advantage of giving authors some insight into reviewer opinions about their paper early. This year, the acceptance decisions came barely a week before the next major deadline making it difficult for authors to implement requisite changes for resubmission on time.
- It's extra load as a reviewer for sure, but it's useful as an author to get some preliminary feedback and figure out what is unclear to reviewers.
- It also has the advantage that authors get early access to the reviews and can start to rework the paper earlier, based on that feedback.
- The biggest advantage of an author-response period is the fact that authors can decide, if after seeing the scores from the reviewers, they should continue to keep their papers in the conference's review pipeline, or withdraw them.
- There are many cases where small changes suggested by reviewers or additional experiments may get the paper accepted; in this case the authors need not wait for another conference and can publish their research before current results become outdated.
- I've had a paper go from borderline to accept thanks to author response so I can vouch for its occasional effectiveness. Ultimately though the real reason I'd want to keep author response is because it allows authors to receive partial feedback earlier on in the review process.
- Additionally, as an author, I appreciate the chance of seeing the reviews before the final decision and being able to clarify any large misunderstandings.

Positive: Author response is important for ensuring review quality

- I have had very good experiences, both as an author, where I was able to clarify some misunderstandings, and as a reviewer/AC, where the response helped me re-assess or justify a decision.
- Strongly favored both as a reviewer and an author. (...)
- I think author response is essential. Both for clarifying reviewer misconceptions and for confirming/refuting claims.
- I have been in several fields, and NLP has (by a wide margin) the worst quality of reviewing, especially in the last few years. Often the reviewers appear to be graduate students, or have fundamental misunderstandings about basic concepts. Having an author response allows the authors to at least point these out for the program chair (I've never seen a reviewer positively change their review from an error pointed out by the authors before).
- The absence of author response seems to result in common serious errors in the reviews.
- The cons mentioned above are not "serious" cons. They can all be addressed if more time/work is done! However, low review quality is currently a serious issue in our community, and author response is of the most important tools to tackle this issue and take

reviewers responsible. Please do not remove it. Ignoring author response just allows reviewers to judge a paper by their own understanding of it which may not be even close to what the paper actually is.

- It's not just about the misconception. I have observed that reviewers don't even read papers carefully.
- author response is a great tool for highlighting to the reviewers (or ACs) clear factual errors or misunderstandings about the review. I do not believe that removing the author response saves too much time as it adds only around 1 week to the review process.
- As an author, it is important to be able to signal to the reviewers/ACs when the original review reflected a misunderstanding about the paper. As an AC, it helps indicate whether reviewers' critiques are justified.
- The last ACL there was an author response. However, this year I didn't have the option to clear some misunderstandings from the reviewers. I am sure the author response helped to get my work accepted that time. I had a positive experience about author responses in other venues like CVPR, ECCV, and ICCV.
- There are plenty of easily-answerable questions that come up about experimental details or other things that are trivially answered by an author response, and where the answer would meaningfully change my review. In most serious discussions that I've had with reviewers over papers recently, someone has lamented the fact that there was no author response to clarify a point
- I like the ability to respond to review as an author, to get a clarification as a reviewer, and to get additional information as an area chair.
- As an author, I found the response period helpful to (1) answer any questions the reviewers might have, and (2) clarify points that were unclear in the submission.
- I have found author responses useful in cases where reviewers had different opinions based on the extent to which they understood the paper. Sometimes a bad review is due to a misunderstanding which can be corrected during author response.
- I like author response both as a submitter and as a reviewer. I have had scores go up on my papers several times after clearing up reviewer misconceptions.
- Author responses sometimes help me clarify nuances that would either justify the claims made in the paper as a reviewer and an author, or help me further confirm that the paper under review lacks in scientific rigor and thus the results are dubious, which makes a huge difference.
- Authors' ability to respond may improve reviewing.
- For recent conferences without a response period (NAACL 2019) and with a large number of new reviewers, the review quality has decreased overall, leading to an increased prevalence of factually incorrect statements or misunderstandings in reviews.
- The response period can be valuable for the authors because it can help clarify the reviewers' comments, and it puts pressure on the reviewers to be both clear and polite in their comments.
- Yes, processing author responses is a bit more work. But is that really an argument when reviewing quality can be improved? I would bet that having author responses has an impact even beyond having reviewers change decisions, in that reviewers probably feel more accountable to the authors if they are exposed to the response and therefore review more carefully.

- A small misunderstanding can get an otherwise great paper rejected. Without an author response, the reviewers always err on the side of caution.
- I think the author response is helpful, even if it often doesn't work as well as some authors might hope.
- Sometimes reviewers clearly do not understand something, and this gives authors a chance to point that out.
- As a reviewer I have non-trivially updated my review numerous times after reading the author response, and as an author my reviews have changed after the author response period. I think it's a great part of our community's reviewing process that absolutely adds value, and that I brag about to friends in other fields!
- The effort spent on both sides - reviewer and author - definitely increases quality of accepted papers. Even if scores are not raised or lowered by reviewers, the reviewing process comes closer to journal papers. And don't we all claim archival character for our conference papers?
- I think that it is important to have "author response" for authors or conferences to avoid a good paper being accidentally rejected only by a (possibly bad) reviewer that misunderstood the content of the paper.
- People says that authors responses are useless, I beg to differ. … (more comments)
- In NAACL recently, my paper was rejected with two 4s and a 5 and very little meaningful explanation as to why. While I understand the constraints of preparing a good program, it was frustrating not to be able to participate. I also definitely read author responses as a review and have often changed my score (usually about 1/5 of the time, though maybe only by a point) because of points they raised or addressed.
- Our paper initially received scores of 3.5,2,5 and after author response it received 4,3.5,5 and received a best paper award. Yes, it's only a few papers for which the decision is changed but it should be seen whether those papers are just regular papers or have a controversial revelation? If it is the former, sure the argument makes sense but if it's the latter then the author response is serving it's purpose well.
- It is true that one time as an author, the result of the author response was increasing the score, which lead to the accepting the paper. In retrospect, in that particular case this was the correct course of action as the paper is quite often referenced by others.
- They're becoming increasingly necessary as reviewer quality drops and sometimes exhibit complete incompetence (I had reviewers claim there's no appendix, i.e they didn't look in the Supplementary Material). A *very short* response to clear out hideousnesses like this might be a reasonable compromise.
- As an author, I have found that the author response is helpful in clarifying details that the reviewers may have missed in their first pass over the paper.
- The author response is important but way more important is the discussion between reviewers. If a reviewer clearly missed something important, they should be able to be notified by other reviewers.
- Sometimes, it seems a reviewer doesn't understand a concept and the review process allows for clarification.
- If we use the same reviewing process that's in place today (3 reviewers, AC recommendation, PC decision), then I think author response can help make better accept/reject decisions.

- Good science comes from good discussions. Ideally, one paper should be accepted/rejected only after many exchanges between authors, reviewers, and area chair. Removing author response is promoting randomness.
- I have seen changes in all three roles (author, PC member, AC member) which led to improved reviews. One of the later versions did instruct authors to stick to questions asked by reviewers or misconceptions in reviews: I think this option is best (if reviews seem adequate, nothing needed from authors, PC members or AC during this process).
- Author response is the last remaining chance to react to tons of poor, irresponsible reviews.
- In ACL's current format, author response periods are the only way to unsure transparency. Reviewers are overloaded and in my experience have often been innacurate in many of their comments and judgements.

Positive: Author response is important for fairness

- Even a small number of papers changing their outcome can have a profound impact on the people actually involved in those papers. So we must do our best to make the process as fair as possible.
- A chance to rebuttal is only fair!
- Obviously, scores will change in only a relatively small number of cases, but the author response is an important mechanism for improving fairness and mitigating the problem of asshole reviewers.
- Author responses are an absolutely necessary part of the reviewing process, to clarify misunderstandings about the submissions that happen way too often, especially since the reviewer pool has become much more diverse in recent years. Yes it is a bit more work for the reviewers, but it really isn't that much as the reviewers should have read each paper carefully. It is the only way for authors to interact with the committee before a decision, and it is one of the very few ways to make sure the conference is fair and not biased.
- I haven't seen a case of reviewers changing their opinions based on author response. However, I think the authors need to be heard. Considering that our procedures for reviewing are no where as thorough as a journal review process, I think that is the least we can do for authors.
- It is very likely for reviewers to misunderstand the paper, which can be easily fixed with author response. Papers being misunderstood does not necessarily mean they are of poor quality, it may be the result of poor writing, which is common among non-English speakers.
- Reviewer misunderstandings are very common, increasingly so as time pressure and loads increase, and reviewers are increasingly junior. Author responses are crucial for fairness, especially when the consequences of resubmitting an incorrectly rejected paper may mean it will never be published.

Positive: Improves the (perception of) soundness of the process

- Even if the result does not change, the rebuttal can make authors (and reviewers) more confident that the process and result are sound.

Positive: Author response is important when doing research on novel topics

- It is also helpful for papers that work on novel topics that don't clearly fit into an existing track.
- Considering the large volume of reviewers, it is possible that reviewers are not familiar with the sub-area but end up giving low scores without substantial reasons to back it up. It can be discouraging and unfair to the authors.
- A major purpose of scientific reviewing in my opinion is to ensure that the community and society do not let new, potentially revolutionary, ideas get bury under a majority of mundane works. Author response is a MUST in this kind of scenarios, because we cannot trust three largely randomly selected reviewers would easily see the merit of such work out of a high pile of mundane papers. It is of course at the discretion of reviewers or area chairs to read author response, but to remove this chance (however slim it is) simply because a majority of submissions did not benefit from author response is to give up our role as a scientific gatekeeper and facilitator.
- Moreover, more than once reviews contained clearly incorrect statements, specially regarding expertise outside of mainstream "deep learning". The only way to make reviewers responsible for their comments, as well as to allow authors to act and argue against innacuracies in these, is having a rebuttal period or, optionally, having an open review policy.

Positive: Ability to ask/answer questions is useful

- As a reviewer, I also regularly change my reviews based on responses. I appreciate being able to ask questions in my initial review.
- As an AC for ACL 2019, I noticed that on multiple occasions the reviewers had questions that could have been answered in an author response.
- Helpful for ACs, good for authors to critically engage with reviews as well
- It does give you an opportunity to clarify things. Even if it doesn’t sway anyone’s opinion, it gives you something to base later revisions on.

Positive: Increases quality of resulting paper

- Even if very few reviewers change their scores, it is a way for authors to systematically address concerns of the reviewers which in my experience leads to higher quality papers overall. As a community, I believe we should be aiming for *quality* papers (as opposed to quantity) even if the reviewing process takes more effort.
- Even if the outcomes don't change for the particular, I feel that the exercise of author response still results in better research because it requires authors to engage with shortcomings of their work.
- As an author, writing a rebuttal gives me an opportunity to address the reviews, but also helps me write a formal response in detail. It is absolutely extra work, but it's work that I benefit from.

Neutral: Author response needs to be scoped better

- As a reviewer I have had only few experiences where I felt the authors tried to really answer my questions. When they did (even if meant that they had to explain a limitation of their work), it always made me look at their paper more positively. But often I felt the authors were just trying to downplay limitations of their work and weren't really engaging with the reviewers' criticism in their responses.
- It is too tempting to provide new results / more numbers that the reviewers asked for even though this is discouraged.
- There is a very limited role for author response in pointing out factual errors in reviews.
- It would be better if the request for author response could be framed in the form of specific questions from reviewers as opposed to just responding to negative comments.
- As a reviewer, the only author responses I found really useful were those that provided some additional results that were missing in the paper, but easy to obtain within a week. More recently, the inclusion of new results in the author responses was explicitly forbidden. Under these circumstances, I don't think the response period makes a lot of sense.
- I believe that a few day author response can help resolve problems with borderline papers (but I don't think it should be used for papers that are not borderline). It should be used mostly for clarification and/or to identify misunderstandings, but not to actually describe/modify/explain substantial information. While it does increase reviewer load short term, it could reduce the reviewer load on the long run as a lot of borderline papers that get rejected are instantly resubmitted to other venues.

Neutral: May only be necessary in some cases

- I think there should be author response if the paper is borderline accept or less so that misconceptions may possibly be clarified. Papers with high scores need not have author response.
- I think that when paper are borderline or a decision can change based on some additional comments from the authors then author response is extremely useful. On the contrary, when the decision on the paper is clear then it is useless and should be avoided.
- I think it is only ever useful in borderline cases. It could be useful in case of gross misunderstanding of the paper from one or several reviewers but reviewers are unwilling to change their opinion/score.
- Author response could be helpful for a small percentage of submissions where the reviewers are unsure about some points in the paper and the author response could help clarification.

Neutral: More interaction like TACL/ICLR is better

- Maybe the best strategy is to encourage more people to submit to TACL, which has proper author-reviewer interaction
- I think that a more journal-like review process with a proper rebuttal phase like that of ICLR is very beneficial for paper quality.
- Interactive, asynchronous reviewer-author discussions, like ICLR has, are more fruitful and scientifically beneficial than one-time “responses.”

- For author response to work, there needs to be a real 2-way discussion between authors and reviewers, as in journal reviewing or an OpenReview style discussion period. But this cannot really scale to the way that our conferences currently work. As a reviewer, it helps to be able to ask questions, and the answers help to form a clearer opinion of the paper, even though it is unlikely to change the aggregate opinion of the reviewers. It also leads to better science if the authors execute some of the reviewer suggestions, which is then checked before final acceptance (I believe some conferences have a conditional accept) but it won't work unless there are more fundamental changes to our publishing model.
- As an author I feel there needs to be ways to clarify certain aspects of the paper. However, I find that the author response period does not really help. However, something like openreview might be a better way to encourage author responses and reviewer engagement IMO.

Neutral: Format of author responses is important

- If there is an author response, I prefer the model used by IJCAI where a separate reply is given to each individual review.
- The space for author response should be longer. But the experience of several submissions told me that the author response seems to be useless.

Neutral: May be best not to show scores at author response time

- Suggest that the reviews be shown to authors without the scores, so authors focus on the contents of the reviews.

Neutral: It might be better to have the response be only for the ACs

- I particularly liked the setting of COLING-2018, where the author response was to the area chairs
- While author response may not lead to significant changes in reviewer scores, it does provide an important additional piece of insight for area chairs and program chairs. I liked the approach taken by COLING 2018, where author responses were kept but only visible to ACs.
- I liked the following setup from a couple of years back: (describes COLING 2018 process)
- No clear benefit. It would be more effective if authors could signal "wrong"/"bad" reviews in a different way, e.g. directly to the AC.
- WHile I believe author response is useful, in the current format it is broken and ineffective. Reviewers rarely change their minds. I think it is very effective in the COLING 2018 format where author responses go to ACs instead of reviewers.
- I wonder if it wouldn't be best if responses were handled entirely by area chairs, and only for borderline papers. Reviewers would still ask questions, but only ACs would need to review the handful of responses.
- Both as author and as area chair, I found it best when authors replied to area chairs (too).

Neutral: Need better indication of whether the response was actually considered

- Somehow make reviewers to confirm that they have indeed read the response and it was not ignored
- If the reviewers actually incorporate the response into their final review, then it's worth keeping.
- It would also help to know if the response will be read.
- I think author response is useful, but the way it was implemented in *ACL did not seem to be useful, as it was often ignored and empirically it has been proved how little influence it has. Therefore, if author response is not going to be improved (which would be ideal), removing it has not been detrimental.
- I am opposed to the current format of the author response: without a reaction from the reviewers, the response is useless from the point of view of the author. Even a simple "The reviewer read your answers and did not wish to change his/her evaluation" would be a good start. I also think that in order to reduce the amount of misunderstandings and unfair reviews, authors should be allowed to flag one or two critical points of disagreement in order to force a reaction from the reviewer.
- That said, there should be an acknowledgement by reviewers that they have read the author response and they should note whether it did or did not impact their decision.

Neutral: Depends on reviewer load/assignments

- It depends on how many papers a single reviewer/author has to respond to. If I as a reviewer were to respond to 5 papers, I'd oppose, for 2 I'd Favour
- I think a lot of this depends on how reviewers are assigned to papers, and how well-suited each reviewer is to review the paper.

Neutral: If there is no author response, the ACs have a much greater responsibility to vet review quality

- A lot of reviews I get from *ACL reviewers are helpful. however, there's been quite a few times where I've gotten extremely awful reviews -- not even in the sense of being negative about the work, but like, single sentence, unhelpful reviews -- even sometimes on accepted papers. It's an insult to both the authors' and other reviewers' time and effort, and ACs should have more of a responsibility to bar reviewers like this.
- I got a review saying "All in all, it's not a scientific research paper." in EMNLP 2018, and after the author response, the reviewer at least increased the recommendation score from 2 to 3.
- Even there is no author response, I would like to request for ACs to check the quality of the reviews carefully too.

Negative: Has little effect / too much effort for authors

- I believe it's good to communicate and to discuss, but I also know in fact reviewers would barely read my response. This process is not designed in humanity.

- My general view about conference reviewing is that because we have several ACL conferences per year, the effort put into reviewing (via author discussions or meta-reviews) should not increase. If the paper is good enough, it will get through anyway either on one or the next conference. There's really no point spending extra time and effort arguing and discussing the paper as soon there will be a next chance for the authors to submit their improved version. Also, for instance, if the general idea of the paper is good but the format/presentation/writing is bad then I would prefer to reject and tell the authors to improve their writing and resubmit to the next conference where it then would be very likely accepted.
- Apparently the reviewers did not bother reading our answers...
- it sounds like a good idea in the abstract but increases everyone's workload quite a bit plus has very little impact on paper acceptance
- It's a waste of time. Rarely the discussion has an impact.
- Author response is mostly inefficient because authors usually argue with wrong people who review their papers with prejudice. That's why paper assignment is far more important than author response.
- While it can be helpful to clear up misunderstandings in preliminary reviews, my overall impression is that author response has a rather small effect on the final outcome and that the extra time and effort therefore is not clearly motivated.
- Even when taking the time to carefully write a response, reviewers rarely changed their review text (i.e., removed the questions), and only changed numeric scores in a way that would reduce diversity among reviewers rather than address author response
- In principle it could help, but in practice I have not found it useful as either an author or reviewer.
- First, it adds an additional (somewhat significant) layer of work for the authors. Second, most reviewers do not change their reviews based on author responses -- even when their concerns are directly addressed or clarified. I am not even sure that most reviewers read the author responses.
- I found it less helpful to argue against a review and I don't think the response ever swayed the overall decision.
- I have seen author response help but most of the times they don't help. As a reviewer, I have changed my score by 1 point on seeing a clarification but this is not the common scenario.
- It seems like a lot of work for authors with basically no payoff. As an author, the idea that you can correct reviewers' misunderstandings seems appealing, but I've never seen an author's response change a reviewer's mind. If anything I think it makes reviewers dig in their heels even more.
- In my opinion, it doesn’t change the final outcome and just makes you nervous for a month hoping for your response to change the reviewers opinions, which rarely happens
- Despite addressing some point of the reviewers, and pointing out obvious discrepancies in reviewers judgement, the final evaluation remained unchanged.
- For authors it is a chance to say what the reviewers missed about their paper. But I find that rarely changes reviewers minds. It always seems just a lot of extra reviewing time.
- However, in my experience it has never changed the outcome of a decision so amounts to an unnecessary time sink for both authors and reviewers.
- I agree that in some cases the author response can clarify a misunderstanding, but that pro is down-weighted by the extra burden author responses generate.

- In my experience, reviewers rarely change their scores following the author response.
- The author response process seems ineffective and unfeasible, unless we change the current review system in a somewhat drastic manner.
- Not worth the effort - reviewers ignore responses.
- the author response rarely seems to change anything, and it is all extra work both for reviewers and authors
- Most of the time, I feel the reviewers do not really carefully read the responses.
- I am a little disappointed that the scores don't change much though.
- Author may spend long time on how to organize their words to reviewers. But in practice, reviewers seldom change their reviews after author response period. This may due to the lack of discussion raised by AC. Thus, the author response seems no use in most of the time.
- In most cases reviewers do not change the score. I prefer to reduce the workload of authors and reviewers
- However, the ratio of such papers is extremely small, compared to the effort spent by all authors and reviewers. Comparing the cost and the effect of this process, I vote for not having it in *ACL conferences.
- It just doesn't matter: decisions are seldom influenced by the responses.
- Although a good idea in principle, I find that author response is usually ineffective and does not change reviewers mind or reviews , so in practice it is just wasting time.
- It has never materially changed my opinion of a paper. My sample size is small, though.
- In its current form this option is useless because reviewers do not really change their minds based on the response.
- I believe, the major disadvantage of this process is that most reviewers do not take it seriously and there is not mechanism to ensure that the reviewers have seen and examined the authors' answers. This means that authors may have wasted their time for clarifications that were not given the proper attention or were never seen by the reviewers.
- If there were a way to make the whole author response process faster, then it would be worth it, but as it is now, it's just too time consuming for little reward.
- The author response is potentially useful, but in 90% of the cases the reviewers do not have the time or the will to read the response and eventually revise their judgement. As an author, I was often feeling that replying to the reviews was simply useless.
- I saw only one case that author response changed my opinion as a reviewer. I think that the effort is not worth the average benefit.
- There seems to be no impact of author responses on the reviewers, as it is rarely the case that reviewers change their opinion about a paper, especially reviewers who gave generic and non-detailed reviews.
- Author responses have some value, but they add a huge amount of overhead to the review process for both referees and authors. I don’t think they’re worth it on the balance.
- Ultimately, it naturally has pros and cons. But the stress and efforts caused on both sides may not pay off in total.
- I think reviewers have already formed an opinion and in my experience the review response period does nothing to dissuade or correct misconceptions.
- I've heard they don't make much difference, if this is the case it seems pointless.
- Nice idea in principle, but not very effective in practice.
- As a reviewer, I haven't received any author responses that have made me change my mind. Often times the tone of the response of the papers that would need an increase in scores to

get accepted is rather defensive or the author's try to justify their decisions that I have understood correctly but not agreed with the authors. Overall, I have felt that the authors response is a waste of time and when I've felt confident enough in my review (which is most of the time) I do not even read author responses.

- -Lack of evidence that author responses improves review quality.
- While incorrect reviews can be frustrating, I feel the response period mainly just wastes a lot of people's time because usually it doesn't change anything but means more deadlines, writing, reading, etc. I don't recall finding it that useful as an area chair either.
- In principle, I like the idea. But as an author, I have never had my response acknowledged. And as a reviewer, I am often the only person to try and respond to the response. So in practice, it may just be better to shorten the reviewing time.
- Paper reviews in which the author can reply to reviewers' questions are usually more effort than they warrant. Anecdotally, only 1% of the author rebuttals have actually flipped a borderline judgement to 'accept'.
- due to a large number of submissions and the corresponding heavy review workload, not all reviewers will take author responses so serious.

Negative: Reviewer discussion is more important

- Discussion among reviewers is much more important and should be prioritised over author response.
- In my experience reviewer discussions are more useful than author responses.
- Author response is extra-work for everyone, so one should really think hard about it. Note that it goes hand in hand with reviewer discussion, even though the two processes are quite different. I am in favor of the latter, not so sure about the former.
- Discussions among reviewers are much more useful.
- I think a high-quality reviewing and discussion among reviewers are more helpful than author response.
- The job of author response is to allow the authors to point out mistakes that the reviewer made (as opposed to changing the paper, providing additional results, etc.). However, in my opinion, this should be the job of the PC, not the job of the reviewer. I.e., if one reviewer has missed an important detail, or misinterpreted the paper, the other reviewers should point this out during the programme comittee discussion.

Negative: Over-emphasizes the point-based system of reviewing

- Author response seems to feed into the point based view of whether a paper should be accepted. If I can get this 2 turned into a 3 then my average score will be 3.5 which means accept. This isn't a good thing.

Negative: Diversity and inclusion issues with author response

- It also seems to work against groups who may not have as much confidence entering into these kinds of discussions - in particular I'm thinking of women and non-native speakers, who I think in general are more open to reviewing criticisms than other groups. So, I think there is a serious diversity and inclusion issue introduced by author response.

Negative: Not enough time is provided to do an appropriate response

- ... not for *ACL conferences, which have too many submissions and which are too often not accepted after discussion, and for which the timing is too tight.
- For one, as a PhD student I was often pressed to run 8-10 experiments in a 5 day rebuttal period (often orthogonal to the paper just to satisfy the reviewers).
- I'm in favor. But for this to be useful, authors must be given substantially more room for replies, and reviewers more time for assessing them.

Negative: Is irritating/frustrating

- Time sink, dealing with my own and other reviewers could be irritating.
- Also, as an author, it is incredibly frustrating to put a lot of work into an author response that gets completely ignored by the reviewers.
- I have had papers where the author response resulted in my paper being accepted (and it probably wouldn't have been otherwise), but in general, the author response period is kind of frustrating, since the reviewers rarely read and respond to it.

Negative: Not clear that it results in improvements of papers for the current conference

- Whilst I understand the potential benefits of the response process, they put extra strain on reviewers and authors and it's not clear the process helps authors improve their papers for the conference they've submitted. I think it's better use of everyone's time that, in case of a rejection, authors use the feedback to improve their paper and resubmit their work at another venue.
- Author response has a potentially bigger role involving asking/addressing questions, proposing changes to a final version of the paper, etc., but this is only valuable if you can conditionally accept papers.
- Further there is zero guarantee that authors will actually take reviews into account in their camera-ready version, so we don't know whether the possible misunderstanding will be clarified in the final version.
- In my observation, author response is rarely used to resolve misunderstanding of the submitted manuscript. Instead, authors make promises to answer questions in the final paper, leading reviewers to review a hypothetical final paper, rather than the submitted paper.
- -Uncertainty that clarification questions responded to are also addressed in the camera-ready.

Negative: Paper should stand on its own without response

- The paper has to be clear enough not to necessitate an author response phase

Negative: Too much time for reviewers

- A major problem for me as a reviewer is that I hardly remember what the paper is about or what my comments were when reading the author response. As a result, I either have to spend a long time re-reading the paper or have a strong bias to keep my old score. I think this is very common with other reviewers.
- take too much time for reviewers
- This dissuades reviewers.
- As a reviewer I find it a hassle to have to go back and deal with author comments. I do a thorough job of reviewing, and find it frustrating to hear authors comments at that stage. As an author I don't believe it's going to make much difference. Overall I just don't think it's worth the extra effort.

Negative: Not in sync with other disciplines

- It's also not the way other paper submission systems work, so it's out of sync with other disciplines.

Negative: Journal-style multi-round reviewing is a better option

- Better to consider ways of incorporating multiple rounds of review as with journals into *ACL process

Author Discussion

Positive: Useful for ACs

- How open are the reviewers to accept another person's view that disagrees with ours, is doubtful. Author discussions are hopefully useful to the area chairs.

Positive: Useful in clarifying reviewer's opinions

- Never participated in one, but seems like a good idea to clear misunderstandings; sometimes a re-phrasing of a statement is enough to change reviewer's mind
- I think that having the chance to talk with other reviewers, share the point of views and discuss about the paper is extremely useful and help to raise the quality of the reviews.
- More transparency is always better. Having discussion periods for authors and reviewers is very desireable. If we wish to ensure that the reviewing process is meaningful and contains the least possible biases and randomness, this is the way to go.
- My preferred setup is that the authors can reply *immediately* to a review, and both parties can discuss. Then, the paper is fresh in the mind of the reviewer, and I've found the continued interaction to more frequently cause me to change my review. I love seeing, "I'll run that experiment you suggest", then a few days later saying, "Here are the results of that

experiment". Having this interaction happen when the reviewer submits the review saves time, as the reviewer won't have to reread the paper as deeply.

Positive: More effective than single author response

- This must be better than a one-time response. Whether it works in practice to improve the quality of the reviewing process remains to be seen.
- While an author discussion would certainly be better than a response period,
- Is there any strong argument against having an author discussion period? The rebuttal period can still be 1 week long but enabling people to converse back and forth is much better and fulfilling.
- This is preferable to author responses in my opinion.
- If we are going for a lengthy review process anyway, we might as well do it properly.

Positive: Provides a good opportunity to discuss/learn

- What kind of researcher can say "no" to more discussion?
- It is an opportunity to learn!
- Interactive discussion better promotes scientific exploration, mutual understanding, and good evaluation.

Positive: Similarly to author response, it would be effective

- Same as 12.

Positive: Have experience with OpenReview and like it

- The OpenReview style discussion is a good model to follow here.
- I strongly prefer the ICLR style that uses openreviews: anonymity, accountability for both the reviewers and authors since all the reviews and author responses are public.
- the open review schema conducted by iclr may be a good choice for acl conferences
- Make it open like open review, if someone is interested let her have her voice. Community…
- Although I haven't participated in a conference with this setup either as an author or reviewer, I would support this even more than a one-time author response. I would support a system like OpenReview for ACL conferences, although I would prefer if it was open only to the track (authors and reviewers).
- Please use openreview for all *acl conferences
- I would love to have a review process like Open Review. It improves the overall paper and review quality.

Positive: More efficient than a single author response

- Being a dialog, this may actually be just as fast as the one-off response system above, as shorter interactions can bounce more quickly.

Positive: If this could be a way to get closer to journal reviewing it may be useful

- It depends on how it is set up. If it is just an elaborated author response, then I have the same reservation as for author response with the addition that this would take even more time and effort. If it can be used to implement a kind of "accept with minor revisions option", it could be useful but this is probably hard to achieve.
- Instead, it may make sense for *ACL conferences to have a revise and resubmit option where a paper with such a decision can be submitted to the immediate next *ACL conference. In such cases, the new reviewers should have access to old reviews, discussions and author responses.

Positive: Interesting, but not enough experience to say more

- It sounds like a good alternative, but with higher time commitment than the simple author response. It would be worth trying but not sure if it would work.
- This can be good or bad. I haven't had enough experience with it to form a strong opinion.

Neutral: Similar to journal reviewing

- I don't have personal experience with this, but it seems too me like a longer version of the author response and akin to journal reviewing.

Neutral: Would be useful, but only with a mechanism for revision/conditional acceptance

- This just exacerbates the drawbacks of author response without conditional acceptance.
- Author discussion only make sense if there is a process to guarantee that authors incorporate changes agreed on during the discussion.
- Absent a journal-style process where authors can revise submissions, I am skeptical that an extended discussion period would be useful.
- I recall PerCom has the shepherd-acceptance option where the authors get assigned to a committee member to follow up with them before the final version is sent. This allows further discussions with experinced committee members about interesting works that lack some few things.
- This seems like it could be good, but then I'd like us to move further in the direction of rolling reviews. Perhaps TACL should fill half of our conference presentations and be reviewed via something like OpenReview?

Neutral: Would need to reduce number of papers per reviewer

- Author discussion is ok, but reviewers should review less papers in this scenario.
- Same reason as above
- I think it will have advantages and improve reviewing quality, but for it to work there has to be fewer papers per reviewer than currently. And for many submissions, it won't be worth the effort and probably lead to unnecessary extra work for both authors and reviewers (as

authors, understandably, often defend their work stubbornly even after it is clear that it shouldn't be accepted).

Neutral: If reviewer response is just as bad as it is for single author response, then maybe not worth it

- I would favor author discussion over author response. But are the reviewer response rates for author discussion as bad as author response? If so, it might not be worth it.
- I think I may have reviewed for a conference with an author discussion period, but I think no real discussion happened.
- While the idea is great, in practice, there are rarely any discussions at all, possibly due to extra effort required, both by the reviewers and authors.
- It seems like the majority does not participate at all in author discussions sadly
- Experienced this once at ICLR (as author), but there was no discussion after the response to reviews.

Neutral: Maybe useful only in a limited number of cases

- Sounds potentially cool, but may require a time investment on the part of the reviewers and in particular of the area chairs, who have to keep the discussion focused and goal-oriented, that may not be worth it. Maybe do it only for a small number of borderline papers.
- if the point is to have a long debate where the authors will argue "to death" with the reviewers I'm really not sure it's necessary. No one has the time for that. Maybe this could be only allowed for papers with a strong variation in recommendations ?
- Do you mean reviewer discussion? I think this is useful for some but not all papers.

Neutral: Need to increase reviewer discussions, not necessarily with authors

- AC's should push for discussion more often; I've seen more papers with strongly dissenting opinions not get any discussions than do.
- Again, it lengthens the process and this role should be done by the PC.

Negative: Too much work for reviewers/ACs/authors

- Too much workload for reviewers, especially when have also to review papers in other conferences.
- I think it is too much work for the reviewers/area chairs.
- Adds to the time devoted to an article when reviewing. Removes the need for authors to formulate their responses carefully, since a discussion will allow further clarification later.
- with *ACL and EMNLP getting an ever-increasing number of submissions, it may not be feasible.
- I am strongly oppose the model where reviewers have to help authors improve their paper and get no reward for it. Reviewing takes time, and expertise, and it is my view that authors should do their best to submit finalized (rather than in-progress) work, and submit it when it is ready.

- I tend to finish my reviews early and I find that any further discussions with authors or other reviewers often require me to go back and re-read the papers to re-fresh my memory. Since the discussions often happen in a much tighter timeframe that the reviewing process, I often find it difficult to fit these at the last moment without being too superficial.
- Authors seem to rely on detailed feedback, which is costly in time.
- For the same reasons as I object to author response: as a reviewer, it is incredibly hard to keep this paper in my head for such a long time, and in the few times I attended such discussions I didn't feel they were useful.
- This would put too much strain on the already overloaded review process and result in longer review periods and more churn.
- Same reason as above. I think the authors' time is better invested in addressing the comments they receive from the reviewers for the next submission.
- While it's certainly the best way to clarify misunderstandings, that would even create more effort on both sides. And it adds the burden to the reviewer of deciding what and when to believe.
- Sounds like too much work for reviewers, who are already volunteering more of their time than they can afford
- Please see my response to question number 10, above.
- needs too much work.
- Most reviewers aren't interested enough, also very high overhead for authors
- this would even increase the time/effort burden on all participants; this is what happens in journal publications and i would say it is not needed for conferences where reviewers are usually busy with multiple papers and mulltiple conferences
- Reviewer are too busy to go over the same papers after several weeks. Reviewers should actually forget about the paper they reviewed as soon as possible.

Negative: Would slow down the review process/there's not enough time

- Same reasons; this requires a long review period that ACL does not have.
- Author discussion seems more ideal for reviews, if time is allowed. However, if the author response process is ineffective and unfeasible due to the short turnaround and the large number of papers, there is no way to make the author discussion process work under the current scheme, since it requires more intensive work.
- - Would slow the review process down further.
- Will further increase review time.
- I have no experience with this. It seems to be a good idea but may just make the turnaround time to be way too slow.
- This just slows down the whole review process, and most reviewers won't read author responses anyway.
- I think it's a good idea in principle, but there probably is not enough time for it.
- Dragging the review process even longer is unnecessary and the benefits of this is unclear.

Negative: Not significantly better than a single response

- I have seen ICLR with such a step during their reviewing process, but did not see more benefit.

- If a long discussion is required, the paper is on the borderline. I think the situation without discussion (or with only one response) is not so critical for authors compared to the one without any response, which can cause critical mistakes.
- A one-time response should be sufficient. Similar to most journal papers.
- I've never seen that, so I don't know if it would be useful. I think a longer author response period without discussion, but with the opportunity to improve the paper would potentially be better.
- I'm not sure if the benefits would be much greater than having a single author response.
- I'm not particularly for or against this, but my feeling is that the author response (vs discussion) is a better balance of time vs. benefit.
- I don't have strong feelings on discussion, but I have almost always found one round of response to be sufficient or nearly sufficient.
- Not completely opposed in principle, but I feel that having a single author response strikes the right balance between allowing the author to address misunderstandings and taking up too much of everyone's time.
- I oppose the author discussion because the reviewing loads have increased over years which means that the number of sub-par submissions has also heavily increased. We already have to put in much effort reading and reviewing papers that do not conform to the standards of ACL. Adding author discussion would increase the load of this useless work even more.
- Too much work and too easy to dodge.
- While I think having an author response is useful to clarify particular points reviewers may have misunderstood, I do not think it is generally necessary to have more than one turn of this exchange.
- One round of author response should be sufficient. As is, reviewers tend to ignore. Focus on making the single round more responsive.

Negative: Is likely as ineffective as/more ineffective than regular author response

- I think this would suffer from the same problems as the author response scenario.
- I have wasted a lot of time on ultimately fruitless author debates. I've also seen a paper essentially get re-written by the reviewers in the course of a protracted exchange. I don't think open-ended discussion works well in practice.
- Unproductive
- See my comments on author response. Author discussion is even worse. (having to go back and recall paper is annoying)
- Again, there's no point to discuss with prejudiced reviewers since they are just not interested

Negative: May hurt diversity or overly favor those of certain advantaged groups

- Similar reasons as previous question. This often works against underrepresented groups who do not feel the same level of comfort entering into such discussions.
- This framework seems to favor folks who have a flexible time schedule and can afford to keep addressing all comments.

- It would also great an even more uneven playing field as more highly published researchers would best game the system.

Negative: May result in arguments that would be contentious and stressful

- I'm a little worried that this could end up in very long arguments, which a reviewer generally wants to avoid.
- I've never submitted to NeurIPS or ICLR but have witnessed some of the lengthy discussions which not that uncommonly turned unfruitful. Sometimes we saw people report new experiments in the discussion and other things that shall take longer time to improve. I'm worried that this is too much load for the reviewers and may encourage practices similar to what's described above.
- This sounds interesting in principle but my guess is that it would be extremely contentious and stressful in practice.

Negative: Will increase the variance of the reviewing process

- I found that the discussion period increases the variance of review process quality dramatically. If you get lucky with a thorough reviewer, the same reviewer is likely to actively discuss. However, if you get a superficial review, the same reviewer will likely not participate in discussion. Thus, it will make good reviews better, but bad reviews worse.

Negative: The paper should stand on its own without discussion

- I feel like this practice would encourage people to submit unpolished work.
- Same answer (the paper should stand on its own)
- A full paper, even if a short paper, should be clear enough so that reviewers can see what has been done, and how. While a discussion may be useful for the author, it's not so useful for the conference.
- In the end, papers should be clear by themselves. A misconception or reviewers that do not seem to have the right expertise can be fixed by author response alone. If more is needed, revisions are needed or the paper should be resubmitted where a new set of reviewers can have a fresh look.

Meta-review

Positive: Summary is very helpful to authors/PCs when reviews are conflicting (e.g. w/ the decision)

- I think it's very helpful to have the summary when there is likely to be any confusion for the author when looking just at the reviews.
- In addition, they are necessary when reviewers comments/scores do not agree well with AC decision.

- Meta-reviews are especially important when the accept/reject decision conflicts with one or more of the reviewers. It is bewildering and painful when a paper receives good reviews but is rejected, and a meta-review can help communicate the reasoning and make the whole process seem less arbitrary.
- But, a meta-review would be accepted when a paper with a high score (like average greater than 4) gets rejected.
- I find this an useful practice not only for borderline cases, but also for cases where the scores among reviewers vary a lot.
- I'd say the AC workload is really heavy as it is; on the other hand, I've seen some rejects with surprisingly high scores where some explanation would have been appropriate.
- Meta-reviews are especially needed for cases where the ACs disagree with the majority opinions from all reviewers.
- Especially for mixed reviews: sometimes one has bad luck with a reviewer who does not like the paper, sometimes one reviewer was the only one to notice valid points. It is good for authors to know how the AC interpreted reviews. It does not seem necessary in cases where reviewers all point out the same things and reach the same conclusion. This just seems extra work for the ACs.
- I think establishing a meta-review for borderline cases is a great idea as it will help make the reasoning behind contentious decisions more transparent and increase the trust in the process.
- it would help authors better understand the AC reasoning for borderline cases and hence might be a good addition to ACL reviewing procedures

Positive: Helps to clearly and concisely convey reasoning behind the decisions

- Even in clear-cut cases, it's useful to have a senior person give a concise review of what happened during the review. The reviews themselves can focus on minor details that can be confusing for new reviewers.
- As an author, I do not profit from an author response where I might not even see any impact, but I profit from knowing how ACs have factored in the different reviewer comments.
- I think getting the high-level opinion of the experienced AC is very valuable to the authors, who could use it to improve their paper and understand the decision made.
- Very much in favour of authors receiving a clear summary of why their paper was accepted/rejected, rather than having to piece together the reviewers' comments and guess which ones carried more weight with the ACs.
- Having your paper rejected with good reviews and no transparency into the decision is a terrible experience for authors. Meta reviews require the ACs to provide at least some justification for their decision.
- If there is no author response, I would like to get meta-reviews to get more information about the decision.
- For papers that are not borderline, it should be sufficient to write a one- or two-sentence meta-review. This is not a big burden on the ACs and it is helpful to have a concise summary about the main reasons for rejecting / accepting a paper where the reviewer scores are in agreement.

- Rejection/acceptance ranges get now to fussy given the numerical scores (e.g., 5,5,3 gets rejected while 4,3,2 gets accepted). We need to let the authors know the reason, otherwise the reputation of random rejections may get built around *ACL conferences and area chairs being blamed.
- Meta-reviews will improve transparency, and I don't see how requiring them for all papers add substantial work to ACs, as they basically have to be able to justify their decision in any case, a meta-review just force them to put it in writing, and for PCs/senior ACs to justify when their decision differs from the reviewers' recommendation.
- Again, more transparency is better.

Positive: Prefer for all cases

- Generally, as an AC, I try to write meta reviews for all papers. However, I agree that in clear cases they might not be necessary.
- ACs should write meta-reviews for every paper.
- I can see the point of reducing the workload for ACs, but I think it is preferable that all papers get a meta-review.
- I think authors should see all reviews written for their papers, including meta reviews. I think all papers should get meta reviews.
- I don't think it would be fair to only have meta-reviews for borderline cases since these are not a very clearly defined category and a lot depends on lenience/severity of individual reviewers.
- Writing meta-reviews makes the work of the AC more transparent. It is true this is much more valuable in borderline cases, but much like we write reviews for clear accept/rejects as well, we should also have the ACs explicitly write the reasons for acceptance/rejection in these cases as well, rather than simply relying on the scores.
- I think meta-reviews would be great for all papers, but I am not sure that the amount of extra work is sustainable.
- I think all papers should receive a meta-review by the ACs (assuming that the ACs aren't overloaded by this requirement).
- Would prefer that it is provided for all papers, as feedback on the main reasons for acceptance/rejection, but definitely most important for borderline papers.
- Every paper should get a meta-review - in clear cases they can be short, but authors love to get feedback on the final decision.
- Meta-reviews should be written for all papers. It makes it much clearer which of the many comments that reviewers made are most important for the accept/reject decisions that were made.
- I think meta-reviews are always very useful, I see the point of having them only for borderline cases and it's preferable to have that than nothing but I think it is useful in any case
- meta-review has to be required. we talk about transparency, accountability and responsibility over and over in various contexts, but it turned out that we don't take seriously the transparency and accountability of our reviewing/selection process which is arguably the most important function we perform as a scientific community.

- I prefer meta-reviews for all papers.
- I think that meta-reviews are a good process, in general, for the full PC hierarchy, and improve the quality of the reviews. If a conference can support a hierarchical process of meta-reviews, it should do so for all papers, and not just for borderline cases.
- Meta-review should be provided for all papers. However, efforts in borderline cases should be greater.
- For clear accepts/rejects, a short meta-review is sufficient, but it should nonetheless be provided.
- Would prefer meta-reviews for everything, but I understand it's time consuming for the ACs, so if necessary having it only for borderline cases is still useful.

Positive: Doing it for all papers ensures that the ACs will look at/check all the reviews

- I think it is crucial that all reviews are as constructive as possible to the authors. Even in the cases where a paper is a clear accept/reject, I think it's useful for the ACs to at least take a look and make sure that reviews are not just short opinions.
- The meta-review imposes more discipline on the reviewing process which we badly need these days. But papers that are clearly accept or reject don't benefit as much from meta-reviews, so could be removed from the process to save reviewer and Area Chair time.
- For the ACs, writing meta-reviews is a lot of work and sometimes a bit annoying. However, it ensures that the ACs think thoroughly about each paper and can justify their acceptance recommendations, thus again this can have an impact on the quality of the acceptance decisions. I would probably argue for ""only borderline cases"", but with a rather broad ""border"". A 5/5/5 paper needs no meta-review; papers in the range of an average of 3.0 to 4.0 might.
- In my experience, writing meta-reviews made me understand pros/cons of papers better, and this was an indispensable step to make decisions.
- However, writing a meta-review would hopefully ensure that the AC has read all the reviews carefully. In case a review is generally 'shallow' or 'unreasonable', the AC would be able to point that out. Therefore, a meta-review would hopefully make the process more fair to the authors.
- I see meta-review as a way to prove the area chair is aware of the paper and of the discussion reviewers have done during the review period. Hence, if a paper is a clear accept/reject then there is no real reason why the area chair should focus on that paper. On the other hand, when a paper is borderline / one of the reviewer gave a score that is clearly out of the trend, then I'd like to know (as an author) that the area chair gave an extra eye to the review process and, in this case, the meta-review would be valuable.
- Area chairs have a crucial responsibility in the decision process, since PC chairs largely follow the area chair view (I guess). Compulsory meta-reviews put strong impact on this role: an area chair should make it very clear why she/he wants a paper to be selected (given the reviews). It is important because there is a temptation of relying only on the reviewers' views but not all reviews are always of good quality.
- Having meta-reviews written only for borderline cases will help AC to notice, and circumvent, authors' unjustified bias.

- Meta-reviews are good for multiple reasons and should be made for all papers. They encourage area chairs to be well-considered in their judgements. If the meta-reviews are made public, they are a useful contribution to scientific dialog.

Positive: Not a significant amount of effort for ACs

- I wrote meta-reviews in this decision cycle and found it not a huge burden -- meta reviews don't take much more time if you've written an informed email opening discussion on a paper.
- I felt the SAC/AC structure worked well at ACL 2019 - as an AC I only had to write meta-reviews for 14 papers, if I had a lot more I would be less enthusiastic.
- This should not take too much extra time, as some reasoning must be performed on border-line cases anyway, and should be written down.

Positive: Important for borderline and low-scoring papers

- Meta-reviews make certain that reviewers talk to each other (because the area chair requests/semi-demands it), the the meta-review if done properly is NOT just a summary of the reviewers thoughts, but a clear statement by the area chair about the paper. There are generally so many borderline papers as well as papers with 2 good reviews and 1 bad review, that I think it could be dispensed with for only those few papers where authors agree on high scores. For very low scoring papers, I think it is still important to make certain that the authors understnad the problems with their paper.
- Border-line papers may be just rejected in case of ACL. I prefer meta-reviews save good but will-be rejected papers.

Positive: Important for borderline and high-scoring papers

- I think that for clear rejects meta-reviews are not needed, but for clear accepts they are useful for later-on deciding which papers would be eventually accepted across areas or sub-areas.
- Maybe borderline and positive. Not sure there is a point in writing one for a paper which will never get in anyway.

Positive: Meta-reviews should be shared with authors

- Should be returned to authors
- Should be shared with author
- Meta-reviews for borderline cases = ABSOLUTELY. However, they should be made available to the authors. As area chair, I personally experienced cases where authors didn't see my notes, and have not included the remarks in the final versions of their papers.

Positive: Makes it clear that the decision lies in the ACs

- Meta review puts a lot of work on area chairs, but I think makes it clear that decisions lie with area chairs. I think author response and discussion presume that area chairs are inert and don't take an active role in the decision process (and are able to override bad reviews or break a tie, the objective of achieving reviewing consensus seems deeply flawed).

Positive: Important for catching reviewer mistakes

- There are cases where the reviewers may miss important points in the paper.
- Sometimes (often) reviewers make mistakes, are not competent or are lazy/sloppy. The aC job is to correct. “Borderline” based on scores only abdicates responsibility from AC to do their job for all papers.
- Some rejected/accepted papers that have particularly low quality reviews *need* the meta-review in order for the AC to identify the not-so-good reviewers and force them to justify their decisions. Of course this is conditioned on the AC caring enough, which in my experience has not always been the case.
- Meta-review once saved my paper which was eventually received very well by the general audience.
- I saw lately a 1.5 2.5 2 paper that was accepted for a very good reason (in short the reviewers rejection was not because of paper quality but too linguistic findings), this can happen.
- Given the large variance in reviewer quality the role of the AC is important as it provides a more reasonable counter balance to mediocre and sometimes incorrect reviews.

Positive: Necessary for the PCs

- When the conference grows to the size of ACL 2019, there is no way the PC chairs can read all the reviews and come to a qualified opinion. In this case, meta-reviews are essentially necessary.

Positive: Would prefer for message from PCs also

- It might be nice for the program chairs to occasionally write comments to authors also (either in the metareview or in a metametareview), to explain their final decision.

Positive: It is nice to have the (more) expert opinion of the ACs

- Some expert opinion is icing on the cake.

Neutral: Only necessary in borderline cases

- There is no point in meta-reviewing most papers. The AC only has something to add in cases where they had to correlate points the different reviewers had in common, weigh the pros and cons of the paper carefully, and perhaps consult the paper itself (ie, borderline cases).

- If by borderline you mean the range [2.5-3.5] inclusive! This reduces the load for meta-reviewers and at the same time increases review quality for the gray area.
- Maybe review borderline papers + a random sample of other papers?
- A paper that gets all 1's and 2's doesn't need to have the AC point out all the flaws in the paper all over again. Similarly, a paper that gets all 5's doesn't need additional words of praise from the AC. I think the AC's should definitely have to verify that a paper with all 1's and 2's really is terrible and that a paper with all 5's really is good, but an actual thoughtful written recommendation for anything but a borderline paper seems unnecessary.
- I strongly favour the second -- meta-review only for borderline or edge cases; not only to lessen the burden of the chairs but also to let the majority votes and scores of the 3-4 reviewers come into play!
- Otherwise it increases the reviewing load substantially
- I think that using meta-reviews for papers with a clear accept/reject judgement would make useless the work of the reviewers.
- ACs do have overloaded work already, and I do accept having meta-reviews to borderline papers
- I really do think it's a lot to ask to area chairs. but if it's only for borderline cases, why not? Whatever to make the selection process let's vulnerable to randomness.

Neutral: Borderline is unclear

- what does borderline mean?
- Seems like a reasonable compromise. How do you define borderline? Who makes that decision? Is it imposed on the AC? Or coudl the AC decide for themselves what constitutes borderline in their area?
- Although "borderline" seems to be harder to define these days. It seems many people are not using 6 to highlight exceptional work, but rather as the new 5.
- hard to decide what is borderline
- For clear cases the meta-reviews can be short. I favor having them rather than not, because it may not always be clear what *is* a clear case and what is not. The meta-review can clarify this, with a short but clear statement. It may help to clearly specify this so that they don't take too much time to produce and to consume.
- If only some papers get meta-reviews then it would be good to have a simple rule determining it (e.g. rejected papers with a mean score above 3).
- The definition of borderline is subjective. I understand the limitations of an AC given the volume of the papers.
- Depends what "borderline" means. It would be nice to have decisions explained, especially given how not deterministic the scores can be.
- Meta reviews can be very short for clear accepts, so do not see why we need to limit them to borderline cases. Moreover, how would you define a borderline case? It should not be based on scores alone.

Neutral: Seems like a lot of work, but no personal experience

- No experience of my own, but it sounds difficult and time consuming.

- Although I've never been an area chair, meta-reviewing sounds like a lot of work if your area has lots of submissions. Meta-reviews for borderline cases only definitely seems reasonable, but is probably still a lot of work depending how many borderline cases there are.
- Yes, maybe writing meta-reviews for borderline cases would be a good idea. It depends how much work load it adds to the ACs and whether people are willing to make that additional effort.

Neutral: Have particular comments on recommended format

- Just as the review form has a space for comments to editors and a space for comments to authors, maybe the metareview should have a space for comments to editors and a space for comments to authors; both should be optional. It's much more work to write a metareview for the authors.
- The meta-review can be short for accepted papers (underlining modifications that have to be taken into account in the final version of the paper). For rejected ones, it may be a summary of the main weak points that explain the negative decision

Neutral: ACs need to try harder to write good meta-reviews

- Meta-review is good only if AC takes time to write it. If not, it is worse than useless.
- Meta-reviewers need to step up and act more decisively on borderline papers. Too often when I've been reviewing, the meta-reviewers only act when absolutely necessary.
- The meta-reviews I have received for rejected papers have been useless. They have only repeated phrases from the reviews that I can already see, instead of helping me to understand why the paper ultimately did not reach the publication threshold.
- Often, meta reviews are not very helpful (e.g., one line summary of the what the work does or summarizing what the other reviewers already said).

Neutral: Can re-use some discussion points

- Meta-reviews are a lot of work for the ACs, though I found myself able to reuse what I had sent reviewers to start discussions of unclear cases (a short summary of what I got from their reviews and where the disagreements lie). Perhaps meta-reviews are a way to promote discussions between reviewers?

Neutral: Need to make clear the target of the meta-reviews

- Some clarity on who the audience is for the meta review would also be nice - as an AC, I wrote meta reviews targeted at the PC, where I was very terse in clear-cut cases; I would have written very differently if it had been for authors (in a way that would have taken a lot more work from me). But authors rightly want to see the meta review. It seems hard to find the right balance here.

Neutral: Should not be public

- Good for internal discussion, shouldn’t be public

Neutral: Not feasible when handling a large number of papers

- However, all of this gets much more tenuous when area chairs have more that 10 papers to oversee.

Neutral: Prefer removing the role of ACs

- I think we should remove the role of ACs to make the reviewing process faster.

Negative: More work for ACs and little benefit

- more work for unclear benefit
- I understand it is too much work for ACs to write meta reviews, but at the same time a meta review is only helpful if the ACs review the paper themselves and not just summarize other reviews.
- Being an area chair was so much work I don't think I can do it again, adding more work to it seems insane.
- As an area chair, I was asked to write meta-reviews without necessarily reading the submissions themselves. This was a weird cognitive challenge for me; and I ended up reading many of the borderline papers. This means a lot of work for the AC; and again, I'm not sure it's worth it, as I don't know that it improves the quality of the decisions.
- For authors, they are useful. I wouldn't have been able to do the job as area chair if I had had to write them though.
- Doing meta-reviews properly takes too much time, if it is done for all submissions
- It's a huge burden on the ACs, which is already a demanding job. And the meta-reviews themselves tend not to be too informative. At the borderline, it's difficult to make an informed decision. And it's even harder when that decision has to be justified to the authors. I'd rather just acknowledge that it's an imperfect process.
- As an author the idea of a good meta-review (primarily in cases where reviewer scores are widely different, not even necessarily for all borderlines) seems appealing. However, I think this may be a case of trying to impose regulations to improve behavior, which rarely works. If area chairs are not conscientious, requiring meta-review is not going to fix it. If they are, requiring meta-review only makes their job even more onerous. (And I will say as an area chair that yes, occasionally decisions come down to slightly arbitrary choices. Making me come up with some post-hoc explanation for that isn't going to do anyone any good.)
- Authors read the reviews they are given. Having an AC summarize them is not useful, except to reveal how the AC weights the arguments presented by the reviewers. I don't feel like they add much (either as a reviewer or an author).
- In principle I trust the reviewers and they should be able to find the agreement. Adding a meta reviewer is more or less like flipping a coin, with the only difference that it costs much more time.

Negative: Maybe not necessary as reviewer reviews should be enough

- ACs should trust their reviewers. If the paper got excellent or terrible scores, and the reviewers cited more or less the same reasons, there's no value added by having the AC

summarize. Let's keep in mind that ACs are not necessarily smarter or more knowledgeable than the reviewers.

Negative: More work for reviewers

- They seemed to want to "justify their existence" bugging reviewers for no reason. Not useful. Not helpful. Reviewing is a lot of work. Dealing with this people is a nuisance and they don't add.

Negative: Presubmission mentoring is a better option

- - Pre-submission mentoring, as available for the Student Research Workshop, is a better alternative. While probably infeasible for the bulk of submissions, it could be a one-time service offered to anyone who submit to an ACL* main conference for the first time. (It should continue being available for anyone submitting to the Student Research Workshop.)

Structured Review Forms

Positive: Structured forms improve quality/clarity

- While occasionally annoying, I think that structure helps with the quality of reviews and can improve their clarity.
- At least some structure is important to get reviewers actually providing detailed comments. A more structured form can also help PCs nudge reviewers to place value on the kinds of contributions the conference is looking for (e.g. not just novelty in algorithms, for example).
- I do like the lightly structured forms (strengths, weaknesses, etc.) as those items are almost always required, and easy to distinguish, and it can enforce complete reviews and discouraging lazy 2- or 3- line reviews.
- There was an adjustment period, but now I appreciate structured review forms... they help me organize my thoughts while reading and reviewing.
- Writing more structured reviews (eg for NAACL 2019) was definitely really challenging, but I think it's important because it helps calibrate reviewers on what aspects to factor in when making a final recommendation for a paper.
- The main load in reviewing is reading and `processing' a paper. Once this is done, the reviewer's opinion should be communicated clearly, to improve the reviewing process and benefit the authors. Structured forms facilitate this.
- Continuity makes sure that all reviewers have at least considered all options of the review process. However, I wouldn't make it formulaic, as the weighting may need to differ per paper (in the view oif the reviewer).
- I find it useful (as a reviewer and author) to have more than a single numeric score. Some structure for text comments (strengths/weaknesses etc.) is useful as well.
- More likely to make reviewers consider the paper more objectively.

Positive: Helps ensure that people comment on all relevant aspects

- I do see, though, that a free text review field makes it more likely that reviewers ""forget"" to comment on important aspects.
- It definitely helps me (as a reviewer) remember to touch on the many aspects of research that I should be reviewing.
- and it can force reviewers to be a bit more focused in their thinking.
- This is valuable at least as a reminder about the points to keep in mind, e.g. originality, attention to related prior work, clarity, ...

Positive: Structure is useful as an AC

- As a reviewer, I prefer no structure, but as an area chair, I sometimes appreciate a form with minimal structure. This is because I tend to write fairly lengthy reviews that usually hit on all the points raised by structured forms anyway, but in a format which I think is more natural to me (and, I hope, helpful to both the authors and area chairs). But as an area chair, I know that not everyone writes such comprehensive reviews, and it is maddening and unhelpful to receive very short reviews that do not treat important factors such as the paper's strengths, weaknesses, novelty, etc. Imposing some structure on the reviews can mitigate this problem for some "low-quality" reviewers.
- Some structure makes it easy for the AC and the author to pick out the salient points of the reviewers' writeup,
- I have never chaired a conference with structured forms, so I don't know if structure leads to more fine-grained advice. It might make it easier for ACs to get the important points out of the review more easily, which would be good. Perhaps I would try a single text box + score along with an "executive summary" box in which the reviewer has to give a 1-2 sentence justification for their score.
- I found a minimal structure useful as an AC.
- The main purpose of structure in a review form is to make the job of the ACs/PCs easier. I don't think the purpose is to teach/force the reviewer to write a good review. Admittedly, some reviewers need some help in this area, but we need to recognize that there's more than one way to write a good review.
- I take pride in my reviews and have been recognized as a top reviewer before multiple times. The highly structured review form is quite limiting to me. Minimal structure is probably a good middle ground; though I would prefer none, I can see that it would be helpful, cynically, for folks who don't put a lot of effort into reviews.

Positive: Liked structure from ACL 2019

- I liked what ACL did, where you remove all of the scores for various parts and just have a single number, but have a few different boxes encouraging a more structured review text. The subscores basically correlated with the overall number anyway, and wasn't very informative.
- ACL was good.
- I liked this year's ACL structure. The novelty/soundness stuff from a few years back was too much.

- This year review form was perfect.
- I liked the 'reasons to accept' and 'reasons to reject' from ACL this year. I also like to be forced to write a certain amount.
- This year was good. Some previous years went overboard.
- This years ACL review form was nice!, except for the one redundancy (strengths/reasons to be accepted)
- I love this ACL form. I hated previous years.

Positive: Having at least some positive and some negative points provides a more balanced feel and is important

- I like the (fairly recent I think) trend to ask for 3 positive and 3 negative points. This forces reviewers to find good aspects in even weak papers and provides for a more balanced feel.
- But I think enough structure to guide reviewers to highlight both strengths and weaknesses is important.
- As an author, I appreciate having the strengths/weaknesses structure a lot, but did not get additional value from extra structure.
- I think that separated boxes for strengths / weakness / questions can be useful for authors to have a clear picture of the reviews and for reviewer to organise the ideas and give a clear feedback to the authors.

Positive: Can help reduce bias and increase fairness in judgements

- There are psychology studies that suggest that having more structured reviews are likely to reduce some types of biases in judgements.
- A more structured form could make the entire review process fairer and help mitigate some of the differences between what reviewers will write.
- I prefer structure and review criteria that substantiates the reviewer's judgements. The structure identifies the criteria by which papers are accepted to the conference, and reminds reviewers that they must be consistent with such criteria. The free-form reviews allow even the most expert reviewers the wiggle room to be sloppy, partial, and unfair to the authors.
- I like structured reviews as an author because we can understand much more information about why the paper was rejected than no-structured reviews, and sometimes we can find "bad" reviewers that wrongly evaluated the paper.

Positive: Structured forms are helpful for new reviewers

- I think the contributions, strengths, weaknesses set-up is good since there seem to be a ton of reviewers these days that have never reviewed before
- It is also helpful for new reviewers.
- Structure is great for beginners, but when you already know what you're doing, it sometimes gets in the way of the flow of your arguments and artificially separates various aspects of the review

Positive: Would like a few more numerical scores

- In addition to what I marked above, I still favor originality/novelty and soundness as additional aspects for rating.
- Writing free-form text for each aspect (eg novelty, soundness) is challenging though, so one suggestion would be to just ask reviewers to provide scores on each aspect but leave comments optional. That way, all reviewers consistently grade a paper on the given aspects, but don't have to spend too much time explaining their scores.

Positive: Liked structure from EMNLP 2018

- The recent review form used EMNLP 2018 was a big improvement -- I would favour this.

Positive: Like at least separating strengths and weaknesses

- As a reviewer I often find that too much structure is not helpful. However, it is important for authors to get specific feedback regarding their papers' strengths and weaknesses, at least.
- I thought the old classic forms were okay, although probably had a few too many scores required and I do think it's better to ask for both strengths and weaknesses explicitly. A few extra questions/scores such as novelty could be useful.
- I prefer a simple structure. I find stating the positive and negative points is an important point.
- At the same time, encouraging reviewers to explicitly state the pros and cons of the paper is of value.

Positive: Positive comments about specific aspects of review forms

- For score boxes, I really liked the graded scores of ACL 2019. E.g., scores 1, 2, 3, 4, and 5 were explained in words, but there were also 1.5, 2.5, 3.5, and 4.5 (without description) so you could settle in-between. It made deciding on a score *much* easier in some cases.
- I do like the reason for accept and reason for reject box.
- I have found it helpful when there is separate structure in the form for typos, editorial comments, etc. so that those don't get mixed in with the scientific comments.
- I missed questions about reproducibility and availability of code in ACL this year. I'd like to see this question again without the remark that not having it applies to most papers.

Positive: Does not add significant work

- This really doesn't add work to a review.
- I think more structured forms do not increase the load on reviewers, but rather lessen the burden when they are well structured with multiple-choice forms.

Neutral: Constantly changing formats are hard for reviewers to deal with

- It would be nice if the format would not change so often - so that we might get use to it and get prepared to provide the answers as needed in the review form.

- Please stop changing grading systems! It is not clear that any scale is better than another, but it is pretty obvious that changing it all the time is harmful.
- Don't keep changing it!
- As a reviewer, as long as we converge to some standard I will be happy to work with it (I don't like having to rediscover how my review needs to be broken up for each conference).
- If using a structured form, keep it the same across all conferences and years and provide clear instructions what to review in each field. The current system is very time-consuming for the first paper in a batch and especially for the first time reviewing.

Neutral: Structure may be good, but only if it's optional

- It's fine to have some options where you just rate 1-5 or choose a radio button, but I otherwse prefer a single text box for comments. This box can perhaps encourage sections on strengths and weaknesses, but should otherwise be free.
- Could make structure optional as default text that can be erased
- When designing a structured review form, it might be wise to think about which sections are required and can be optional. Optional sections allow reviewers to leave blank, making reviews efficient.
- I don't like being forced into answering certain questions. Different questions are appropriate for different papers. And some papers don't deserve much feedback anyhow.
- I appreciate having a somewhat structured review form, where you can break out individual points as you see fit.
- I think it is fine to have a default structure template in a text box, but reviewers should not be constrained if they prefer a different organization of comments.

Neutral: Should vary structure based on paper type

- Ideally, there is structure related to paper type. Some questions do not make sense for papers that are not experimental, for instance.

Neutral: Minimal structure + free text space is preferable

- So it's good to have a minimal number of boxes for summary/strengths/weaknesses, but these should be small; and it's good to have a big box for an unstructured, detailed review.
- Further comment section (free-form) should be included btw.

Neutral: Prefer more numerical scores but single text box

- Structure is fine/good for the quantitative scoring portions of the review, but unhelpful for the written portions of the review.
- I would stick to simple structure and use marks for the various criteria.

Neutral: The purpose of the structure should be explained to reviewers

- Reviewers need to be comfortable with the structure and understand its purpose. Otherwise, they may refuse to review in the future. (I've been told so by reviewers.)

Neutral: Structure should be adjusted to fix systematic biases in *ACL papers:

- I think that the three main problems in *ACL conferences these days are the following. (detailed description on: 1. TECHNOLOGY RATHER THAN SCIENCE, 2. LACK OF DIVERSITY, 3. INSUFFICIENT MULTILINGUALISM)

Negative: Too much structure increases reviewer workload

- Some structure is very useful because it improves comparability, but too much structure is counter-productive as it limits reviewers and increases their workload at the same time.
- We have to be realistic that as a community we are too large to allow reviewers to simply respond as they wish. This doesn't scale up.
- I'm strongly opposed to such structured reviewing forms as that for NAACL this year. It makes the life of the reviewer hard at no apparent gain.
- I think having too much structure puts an undue burden on the reviewer, especially if minimum length limits are imposed. I feel that having structured response be optional may be a solution.
- NAACL 2018 had a very structured form, which felt repetitive and resulted in a lot of overhead in reviewing.
- When these are overly complicated, they impose a burden on reviewers.
- I have a certain amount of time as a reviewer to spend on each paper. In the end, it's the paragraph engaging with the work that's the most valuable for other reviewers and for the authors. Every bit of structure you ask me to fill in takes time away from writing that paragraph. Eventually, it doesn't happen at all, and the the reviews become completely useless to authors.
- Too much structure makes the reviewing difficult in my opinion.
- I can write a good quality, fair, concise review in less space than the damn boxes require. Waste of time.
- Much structure means much loads for reviewers, and sometimes the structure is not comprehensive.
- I think this is highly correlated with whether there's an author-response or author-discussion, as the combination of either of these with a heavily structured form seems too much.
- The very detailed structure used by NAACL a while ago was terrible, I'm totally convinced that it made me spend more time and write lower-quality reviews than a more relaxed format.
- Some structure is helpful, too much structure increases workload for all parties without much benefit.
- Detailed structure is a huge burden to reviewers and effectively doubles the reviewer load.
- The hyper-structured NAACL 2018 review form was so far from what reviewers can tolerate that it damaged the entire reviewing process. As a reviewer for that conference, I was unable to deliver a coherent review and my reviews took 2-3x longer than usual, meaning I had to reduce the time I could spend writing meaningful comments instead of answering micro questions about each submission.
- Overly complex review forms like that at NAACL 2018 lengthen the review time and I don't think aid the process and may even detract from it. As an AC I noticed many reviewers

ignored many questions and put everything in general comments, which I myself have done when the review form is overly complex since it is frequently hard and time consuming to make your comments fit an overly complex form.

Negative: General negative comments about complexity

- We have WAY too many variables
- In m'y opinion, there should also only be one score
- I really did not like the hugely structured forms introduced a year or two ago.

Negative: Structured review forms make reviews disconnected and fragmentary

- I am glad that ACL is trying to push for non-trivial reviews. But the form isn't necessarily the best way to do it. Making the reviews more disconnected and fragmentary makes it more difficult to express/perceive what issues the reviewers find genuinely important and which ones are just there to fill in a box.
- Strengths/weaknesses can be good, other structures I've seen lead to shorter reviews because it's a small amount per text box and harder to understand what the core reviewers' comments are.
- Some recent form incarnations have been much too complex, particularly when some results, questions, or concerns overlap across multiple form fields.
- It's silly to ask for comments on strengths and weaknesses in separate boxes, as they are often two sides of the same coin, and need to be discussed together. Sometimes an important property of a paper is both a strength and a weakness at the same time. More generally, what I want to write as a reviewer, and what I want to receive as an author, is a critical essay about the paper, with an organization and emphasis that will depend on the properties of the specific paper.
- To phrase a coherent and well-structured review, free text is the only real choice in my view. I struggled with some of the recent partly highly structured review forms.
- I think ratings w.r.t. soundness, novelty, etc, make sense, but writing a good review is already a challenge without having to split it into a bunch of separate sections.
- Too much structure is counter-productive, hampering reviewers' focus.

Negative: Structured review forms reduce freedom of expression

- The highly structured form at NAACL 2018 made reviewing much more difficult for me, and I don't think straightjacketing the process actually helped the quality of my reviews, but rather the opposite - more freedom allows better expression, I really struggled to organize the review as required and to decide where to put each of the things I wanted to say, and the result were worse reviews.
- Having more structure, instead, could lead to too much restriction, leaving not enough space for comments that lay out of the form questions.
- I have my own review structure that I favor, so it can be very annoying to have to try to cram it into odd boxes.

- Not everything needs to be commented on. What matters is that reviewers justify clearly their scores. I dislike review formats in which we need to give at least n positive and negative points for each criterion.
- I don't think structure is helpful with my review writing, it largely limits what I want to say.
- I feel strongly about this. Every paper is different and as a reviewer, I feel I want to have the freedom to emphasise certain aspects of the work more than others. Having detailed textboxes also increases my reviewing time by a huge amount.
- I think each paper is unique/review and prefer a single text box to structure the review as appropriate.

Negative: Having to write strengths/weaknesses for every paper may make some strengths/weaknesses that are relatively insignificant seem more significant than they are

- I like a separate box for typos and grammar, but feel that enforcing strengths/weaknesses pushes to write strengths or weaknesses that may marginal with respect to the overall feeling of the paper and may be misread as significant.
- Forcing people to write strengths/weaknesses makes it very difficult to get an idea of what their overall take on the paper is. Both from a reviewer and AC perspective, I find an unstructured review easier to write/interpret and much more useful than a structured one.
- I find that structured reviews become schematic and uninformative very easily. Reviews should be able to express nuances, and this gets suppressed when the form asks for strengths and weaknesses.

Negative: Some structured elements are irrelevant for some papers or redundant

- Some elements of a detailed structure are irrelevant for certain papers
- When I'm reviewing, I don't always have opinions for every subtopic in a very structured review, and it makes the reviewing process take a lot longer and feel more fabricated when I have to search for something to say.
- However, I've seen some highly structured review forms that force the reviewer to fit all papers into a fixed template that doesn't always suit.
- As a reviewer, I don't have the time for really complex forms that often don't align well with what needs to be said about the paper.
- Structure inevitably doesn't fit some papers, and is generally frustrating and onerous for both reviewers and ACs.
- Currently (ACL2019), it has some forms which require almost the same content for review. That's too bad.
- Different papers require different review structures. Also, different reviewers like to communicate differently. If you don't trust someone to write a proper review in a free textbox, don't let them review.
- If there are different categories of papers with different review forms the problem can arise that the paper does not fit any of the categories. Having an open category might help here.

Negative: Minimal word/character counts are not good

- but for the love of god, do not have a minimum word count! It totally backfired. People got around the minimum word count for NAACL by ranting about the minimum word count.
- I'm especially opposed to boxes with a minimum character count. This is both burdensome and leads to poor reviews.
- In some cases there are minimum character counts; the effect I observed was that (other!) reviewers paraphrased the same points over and over in order to make up the space.
- An important point: minimum word limits on weakness/strength boxes is BAD.
- If you have a structured form, then remove the minimum length for the free text form.
- I think that requiring a minimum length for each part of a structured review form is not a good idea.
- During recent years I have seen several structured forms and sometimes it is really difficult to conjure up the required amount of words in every required box.
- I just pad the extra fields with low content comments. By the time I write the review I've read the whole paper and formed my opinion. I'm not going to go back and re-read just to find more favorable things to say about a crap paper just to meet the minimum word count.
- I was very unhappy with the NAACL 2019 review form, and in fact have decided that I will no longer review for NAACL. A complex review form with min word counts make reviewing even more of a hassle, and leads to questions of whether reviewing is the best way to spend my time
- During recent years I have seen several structured forms and sometimes it is really difficult to conjure up the required amount of words in every required box. This is especially difficult for very weak papers that will be clearly rejected. If in that case I have to generate at least 100 words about the strengths of the paper then all the text will be artificial and the strengths are sort of pseudo-strengths.
- As someone who is an experienced review (and won a Best Reviewer award recently), I have never been anything but frustrated with the structured review forms. At most, minimal structure (strengths/weaknesses/overall) is best. The minimum word count for some boxes is just frustrating when papers that are clearly incomplete/very poor quality come in, and one has to spend a great deal of time working through the review process rather than making clear rejects and spending that saved time building detailed reviews for the accept/borderline papers.

Negative: Structured forms are ineffective, people will “game” them

- I've seen author gaming the structured system to get away with whatever they want to write, so I would say: if an author doesn't want to make a good job, they'll find a way. More structure won't solve the problem of unmotivated reviewers.
- Too demanding structures irritates reviewers. Sometimes they just copy and paste the "contribution" part of the paper
- I used to like the structured one, but it looks like a lot of reviewers just phone in most scores and AC's don't pay much attention to them.
- If a reviewer needs some forced structure it just means that he is a bad reviewer.

- Structured review forms encourage lengthier reviews, but tend to leave reviews that copy statements from the paper (as strong points) and are not that useful for area chairs to highlight what aspects are really important in the review.

Negative: Even with structure it's not clear how to combine the various aspects into a final recommendation

- It would make more sense if there were a principled way to combine the sub-category scores, but different reviewers seem to weigh different aspects of the paper differently (or sometimes to something like a min of all categories), so it ends up being more frustrating to get something like: 5/5 soundness 4/5 novelty 5/5 appropriateness 2/5 strength of contributions 2/5 overall. It is more helpful for me to just hear the pros and cons and know why the specific overall score was given since that's what ultimately makes or breaks acceptance.

Negative: Makes reviewing difficult for clear accepts/rejects

- For short papers or papers that are easily rejected, more structure or word limits on certain sections seem more a hindrance.
- Makes reviewing harder for very good/very bad papers

Review Transparency

Positive: Transparency is important, e.g. for fairness

- Why hide anything? This is significant work that people have put on the paper, and it could value the community. The only problem with it is some papers may get "poisoned" by unfair / low quality reviews. I Revealing the discussion, and allowing authors to post their responses to the reviews should mitigate this.
- Including discussions improves transparency; except for some confidential topics (which could still be hidden), reviewers should not be writing comments that they would not be willing to stand up for, and they would still be anonymous.
- Transparency is key!
- scores could be kept private from the public, but i believe any outcome of reviewing process (other than discussion which is a process toward which the outcome was created) should be made public (anonymized) should the submission have been accepted.
- Meta-reviews are the lens to transparency. They should be released!
- This question is absolutely strange. Why should it not be transparent? I can only think of bad reasons…
- If we want to make sure our reviewing process is meaningful we must be transparent.
- I have seen several papers that were accepted although they shouldn't. Sometimes this happens because it is controversial and will initiate a discussion. However, many authors and reviewers consider everything published in *ACL as a "gold" truth and that everyone should take it for granted. A recent example is the paper [Title omitted by editors], accepted

in NAACL 2019 which has clearly many flaws but initiates a discussion, concerning attention mechanism in neural nets, useful for the community. The least ACL can do is publish the reviews and discussion in order the community to have a complete view of why this paper was accepted. Of course the reviewers, meta reviewers and ACs should remain anonymous.
- Everything should be public.
- Part of me thinks that the reviews themselves (for all submissions) should be made available behind some kind of license agreement pertaining to maintaining anonymity/professionalism/etc, so that we can better instrument the review process to understand how it can be improved.
- Transparency should increase quality of reviews and discussion.

Positive: Public reviews are informative to readers

- I find the ICLR approach very useful. Sometimes I find the ICLR reviews and meta-reviews more useful than the paper itself. Often the reviews and meta-review clearly summarise the strengths and weaknesses of the paper better than the paper itself. If reviewers are going to provide excellent reviews, it seems a shame not to make them available to a wider audience.
- Publishing the reviews would have the benefit that summaries from different view points, identified weaknesses and suggestions for future work would become accessible.
- - A well-done review is a huge job, offered for free to the community, and can greatly contribute to the final paper. Making such a review public is rewarding for the reviewer.
- - Well-done reviews are simply interesting. Readers of a paper may benefit also of the reviews, and this can trigger interesting reader/author/reviewer discussions and collective efforts.
- It is also scientifically useful for future readers.

Positive: Like OpenReview

- You don't even have the OpenReview option, that is sad.
- Recommend the open-review.
- It would be great if ACL adopted openreview. In addition to transparency, it fosters good discussions, which I think serves both authors and reviewers alike.

Positive: Reviews should be released, even for rejected papers

- I think releasing to the public for rejected papers would be even more useful
- They could even be released for not-accepted papers as far as I'm concerned

Positive: Reviewers should be de-blinded for accepted papers

- Reviewers of accepted papers should be made public along with the review: they get to know who the authors were they reviewed, but the other way around it still stays hidden. Publicizing the reviewers of accepted papers allows full transparency, and might further prevent mean behavior (not that it's that prevalent), and 1-sentence reviews.
- Re. transparency, I would like to see the conference release the list of reviewers for each paper (not associated to each review, just the list of the 3-4 reviewers who worked on a

paper). This would increase the sens of accountability from reviewer, while keeping the actual comment attribution somewhat hidden.
- In post-review discussions the reviewers' identities could become relevant (e.g. to establish particularized expertise). But metareviews should ideally contain a synopsis of post-review discussions. I would want to hear more about other conferences' experiences with making reviews public for accepted papers.

Positive: Public reviews may help prevent authors from making low-quality submissions and wasting reviewer time

- I'm very much in favor of public discussion. I like how openreview is set up in this regard. If you're worried about getting very negative reviews that will publicly stick to your paper, maybe that will also help people self-monitor low quality submissions, helping with the exponential growth in submissions :).

Positive: Public reviews will encourage reviewers to not do poor/uncivil reviews

- Having the review process publicly available (after the notification) could be a good incentive for reviewers to give good and constructive feedback and to avoid 2-line reviews where one mainly asks to cite his/her own works.
- I think reviewers should be held accountable for what they write and some incentive for that is good.
- Transparency is very important in all communities and I think it will encourage more civility.
- IMO review transparency provides the best incentive for high quality reviews.
- - If a review is made public, it motivates both the reviewers' and the authors' for a high quality work.

Positive: Releasing meta-review/discussion to authors would be helpful

- It seems to me that releasing the discussions (question 24) to the authors is a good thing if reviewers make an effort to be constructive. There's in principle no reason why you can't be candid or dislike a paper without being disrespectful.
- I experienced my paper with good ratings was rejected without reasonable excuse (maybe by area chair's decision). That time, unfortunately, the chair published a very similar paper on the next *ACL conference. We cannot prove anything, but it was the worst in my academic life. Solving such cases at all is difficult, but I would like to see enough meta-reviews for uncertain rejection.
- Releasing meta reviews to author is likely important. Plenty of papers with good reviews get rejected and it is unclear why.

Positive: Releasing meta-reviews publicly is a good balance

- I think meta-reviews should be used to summarize these discussions.

- I think some version of the meta-review should be released to the public, along the lines of "here's what makes this a good paper for the ACL." Pointing out each paper's flaws could be seen as the Association undercutting the research—if there are flaws, then why accept in the first place?

Positive: Releasing reviews to authors is important to make papers better

- Part of the purpose of peer-review is to allow self-reflection by the authors. Without an understanding of why their article was rejected, they will not benefit from the peer-review process.
- Not all discussions are better when done in public, but the (rejected) authors have a right to know what needs to be fixed…

Positive: Public reviews provide a learning experience for inexperienced reviewers/authors

- - If reviews are public, it is possible to learn from experienced/outstanding reviewers how to write a good review.
- I think open review such as ICLR helps new authors like me a lot. I can read from reviewers' opinion and learn where is good in this paper and where is bad.

Positive: Public discussion, with an option for private would be good

- Of course, there should be an option for private discussion, but the default strategy, I believe, should be to keep discussion public.

Neutral: Public reviews are only relevant to the submitted version, not the final version

- Don't like the idea of published reviews, since they are only relevant to the submotted version, not the final version.
- I like the idea of reviews becoming public, though there is not much sense in comments from reviewers that were fixed afterwards. One of the problems with conferences is that you can never be sure whether comments are addressed. It can really help the audience if they see potential pitfalls of the paper or concerns from authors.

Neutral: Reviews should remain visible after the conference

- The reviews should be visible after the conference. I don't like how START deletes them after the conference.

Neutral: It should be possible to opt-in or opt-out

- and would want akk authors to be able to opt in or out,

Neutral: Even if reviews are released many people won't care

- Ultimately, only authors care whether their paper was accepted or not, so reviewers don't really need to see a meta-review, unless they want to improve their reviewing ability.

Neutral: The amount of transparency needs to be made clear up front

- Every form that an author, reviewer, etc., fills out needs to state very clearly who will see the response, and when. In a recent conference, reviewers were blind to one another (which is fine) but there was a rumor that the reviewers would be revealed to one another after the discussion period. These things should be made absolutely clear up front.

Neutral: Public release should be on a case-by-case basis

- The committee should have the option of releasing or not releasing certain information for a particular paper.

Neutral: Discussion among reviewers is important

- Having the reviewers not able to see each others reviews until they submit their own review, but then able to see them seems good to me. Certainly, being able to discuss with the other reviewers is vital -- no one wants to read a paper in isolation, unable to clarify simple points with the only other people that have read the paper.

Negative: Public release would make it harder to have frank discussion

- I think that there needs to be frank discussion of papers and of reviews/reviewers, and we won't get this if we make everything publicly available.
- When the reviewers talk about a paper, they are pretty candid. I know I would not talk about a paper if all my comments were to be released. Meta-reviews at least as I have done them should be available to authors.
- I have accidentally and intentionally deanonymized myself in reviewer discussion and would be uncomfortable with those discussions being public.
- The PC discussion should be kept private + ideally anonymous to allow people to say what they think. I would have concerns about giving an honest opinion if I knew it would then be released and the author woudl see it.

Negative: Public release would be hard on reviewers/make it harder to recruit reviewers

- Releasing reviews or discussion to the public will only make it harder to recruit reviewers. We aren't in a position to do that given the numbers of submissions we now face.
- The volunteer review and AC team must be protected
- - Additionally, it will likely be more difficult to identify people willing to review if using public reviews (concerns about increased time in the review preparation process).

Negative: Releasing meta-reviews/discussion may not be good because of lack of anonymity

- The reason why I would be reluctant to release meta-reviews to the public is that the ACs are known, and such there is no anonymity.
- I love transparency but releasing post-review discussions to public can lead to some research that “hacks” the way to acceptance (trying to mold future submissions to styles favored by the community in the past). It may also prevent ACs and PCs to express their opinions freely since they are technically non-anonymous.

Negative: Reviewers may become unintentionally de-blinded through public release

- - Releasing reviews to the public puts the double-blind peer review at risk. Considering that writers can convey their 'voice', reviewers may be recognized, e.g., by prior co-authors who recognize writing style or by prior area chairs or program chairs who have seen prior reviews prepared by the same individual.

Higher Acceptance Rates

Positive: Some separate method of quality control (e.g. Journals) and higher conference acceptance rates would be better

- I think the field would benefit from moving towards a more journal-based publication culture. However, this probably has to be done gradually since at present we do not have enough high-impact journals to accommodate this change.
- Conferences as the main diffusion of science is an expensive, exclusive model and is an exception compared to most scientific fields.
- That being said, I'd prefer to move to a model in which presentation and archiving are separated, like medicine.
- I have mixed feelings about this - I'm strongly oppose to gatekeepers in general, and I like Omer Levy's proposal for reviewing. I'd be happy with a world where papers are posted without gatekeepers at all, anyone that wanted could have space to present their work as a poster at a conference, and we deal with judging quality of papers in some other way. I'm not sure how to make this happen, though.
- We can decouple two questions: Reviewing and Attending Conference. I don't see why these should be tied together? I propose the following system: ...
- I find that there are too many papers with technical problems accepted at the conferences. So, in my opinion, the question should not be how selective the conferences are but what how large the variance on the reviews is. I think the underlying problem is that too many reviewers do not take the time necessary or are not familiar enough with the matter (this is obviously because the number of submissions increased so much). One way to address this problem would be to let journals take the role to mark high-quality work (and I am in favor

of that). Another possibility would be to lower the acceptance rate and (very important) punish people for submitting low-quality work, e.g., but forcing authors to disclose previous submissions and adding a statement how they improve over the previous version.

- Conference publications should remain selective but with higher acceptance rate and with other venues such as journals taking a more important role.
- Considering the increasing number of paper submissions at *ACL, conferences need to be selective somehow. However, I think journal publications should be more promoted in the sense that they can play a more significant role in determining the relatively quality of papers.
- Selective conferences are a mistake, because the conference review process is not robust. It is also unfair to people (such as myself) who sometimes cannot travel because of personal issues. We should learn from more established fields and stress journal publications
- The quality of reviewers is now so poor that good work is being rejected by reviewers who don't know the field or who are very junior (such as grad students reviewing without supervision). Better to be less selective at conferences, and push the `quality` to journals, where editors and reviewers are hopefully better engaged.
- We may need more journals if conferences no longer have a quality status.
- Since conferences as they are today are (i) environmentally criminal (ii) based on the author-pays system (since it is mandatory to both attend and present the paper) (ii) exclusive by design, I would favor completely replacing (or transforming) them with TACL-like journals with acceptance rates similar to ACL.
- I think I would basically be in favour of all papers at top ACL conferences having to go through TACL, so a combination of journal submission and conference presentation, with papers submitted independently of particular conference deadlines.
- I think a combination of higher acceptance rates at conferences but also more emphasis put on journal publications might help. It seems clear that the large conferences have become too large, so increasing the number of accepted papers may not be a good idea in that respect. On the other hand, if conferences were there as a networking tool, with journals taking on the role of quality indicator, less people would presumably submit to conferences. This goes with a perhaps unrelated topic, which is that for some countries / some universities, conferences have become unpayable. Publishing in open access journals, however, is free and thus more inclusive.
- In the long term, I think publishing in larger increments (e.g. journal papers) will encourage better and more discoverable work, but I am not sure of the best transition path.
- I strongly support a more journal-based model of publication in our field, which would lift the travel constraints, as well as acceptance rate constraints, while at the same time moderating the publication pace in favor of thorough review and validation. Basically enlarge TACL (substantially), have it indexed in WoS or whatever (for inclusion across many countries where such things matter), give it a well-curated large pool of reviewers, and then have TACL papers presented at conferences, but keep the workshop format as-is.
- As mentioned above, I'd be in favor of decreasing the role of per-conference reviewing and increasing the role of TACL and TACL-like journals for our community. If this happened, then conferences could have a mix of TACL-paper presentations and on-going work talks/posters. Perhaps this would require reducing the talk-pubs to non-archival abstracts, however.

Positive: Having different levels of selectivity, even within the same conference, may be useful

- I think a hybrid model might work where borderline papers can get accepted to a special poster session, either non-arxival or separated in the anthology. This would improve the inclusivity of the conference and give novice authors a chance to present their work. It would also relieve some pressure off of PhD students who have to compete with industry researchers for the ~20-25% of the accepted papers.
- However, there could be a poster track with a different, publicly available higher acceptance rate to encourage more participation (not that we need any more attendees!). The workshops already serve that function pretty well, but some universities only provide financial support for meetings where you have an accepted paper, which may just be the workshop and not the conference as well.
- Should consider higher rates for posters to encourage participation
- The reviewing is the key component here, as many accepted papers are really debatable. Reviewing practices should be better. Also, it seems to me that the ACL conference have a bit of an identity crisis and all are more or less the same in terms of the areas included. Maybe there's a way to redirect the submissions, for instance accepting only long papers in one venue and only short papers in another.
- Maybe more strict separation of the major conferences into the sections is necessary.
- We need more conferences, probably with focuses on a topic or region. Perhaps such lower tier conferences can have higher prestige spotlight presentations. *ACL could also experiment with this.
- In the short term, having selective and non-selective tracks (e.g. reviewing only abstracts) for conferences might be the best thing to try. In the long term, most quality papers should be published in journals.
- Perhaps keeping long conference papers selective but accepting more incremental work as short papers (and giving reviewers an option to accept a long paper as short) would allow people to get their ideas out in the short term while incentivizing more substantial work as more prestigious. Another possibility would be to shift some of the burden to workshops/more focused conferences like CoNLL and *SEM, perhaps with an option to accept a main conference submission to a colocated workshop/conference instead without an extra round of review.
- The reviews appear to be getting more random (especially with the latest round of the Toronto Paper Matching system use, rather than bids). If the science is good and sound, I think papers should generally be accepted. Perhaps creating a tier system (i.e. the "super amazing" papers one day, and the rest for the rest of the conference, and this is determined by the reviewers/PCs), but it doesn't seem like it benefits either the field/science as a whole, or our careers, to have work stuck in the publishing pipeline for no reason other than competitiveness. That diminishes the science of the field as a whole, and sends folks publishing (not peer reviewed) to arxiv, which really hurts us.
- I think that conferences should make public and indicate the clear accepted papers, to give them an advantage to stand out among the many other papers in the conference (also in proceedings).

Positive: Higher acceptance rates are better for the majority of people, other than those going for tenure/positions at competitive universities

- Very low acceptance rates tend to be favored by a small subset of the community (who are struggling with tenure and career issues at very competitive universities). This is nice for those in that group, but not good for the rest of the community.
- I'm in favor of keeping the same acceptance rate (I wouldn't want to go back to my PhD days if acceptance rate for major conference would have been 10%). Since the number of submissions more than triple itself in the last couple of years, the consequence is that there are many more papers presented in each conference (mainly many more parallel poster sessions).

Positive: Physical limitations don't seem to be a strong argument

- With poster sessions I do not see physical limitations as a reason to keep out high quality papers.
- The number of posters can be increased to support more papers. Most *ACL talks are low quality compared to other fields.

Positive: Low acceptance rates lead to inefficiency

- I think we should increase our acceptance rates, otherwise we smother ourselves with resubmissions over and over again.
- Many high-quality papers do get rejected and I believe that is also causing the review overload of the papers.
- Currently low acceptance rates and increasing number of researchers in the field leads to the infinite re-submitting process.
- We really need to acknowledge that our current models for reviewing aren't scaling at all, and we need to be willing to change. We need conferences to be somewhat selective, but I think our obsession with 20% or whatever is extremly harmful right now. Bigger venues, more acceptances will help lessen pressure on our system due to resubmissions and all the overhead associated with response, discussion etc. (which we ought to stop trying to do). Get the work out there and let the community decide what is important. Many of our conference papers already go on to have little impact, so let's admit that and maybe focus far less on a highly precise reviewing process (let's go for recall).
- Rejecting papers which have unanimous accepts to keep the balance makes no sense at all to me.
- While I think that acceptance rate should remain selective, selective rates combined with the anonymity policy and multiple submissions being disallowed this year at both EMNLP and CoNLL make the situation very risky for researchers. In case of rejection, a researcher will have lost a few months.
- in favor of doubling (approx.) current acceptance rates; thereby having more posters and fewer oral presentations. I have a better experience with speech conferences, which have acceptance rates >40%. I don't believe that quality of these conferences is lower and both writing and reviewing papers is a better experience. At the end, with ACL papers a lot of the work goes on the quality of the writing, not on the experimental results. You end up

carefully looking at each single word trying to anticipate any possible criticism. Is this exercise really useful?

Positive: Low acceptance rates lead to variance

- Also, it's fair to say our reviewing process is very noisy, and we probably reject a number of pretty good papers (and accept some that aren't so good).
- The ultimate decision of the interpretation of the papers should be left to potential readers and NOT on a three randomly chose reviewers.
- I don't think lower acceptance rates will lead to higher quality programs due to the variance in review scores. The conferences should move to the same model as every other scientific field: conferences for feedback/networking/sharing work, journals for significant research contributions.
- The review process is quite good at identifying papers for a clear accept or clear reject. it is somewhat arbitrary for the vast of majority of borderline submissions.

Positive: Low acceptance rates lead to rejecting good papers

- Conferences should be selective and high-quality, but at a 25% acceptance rate, we are rejecting papers whose quality is quite high enough to merit inclusion.
- If we go on with the current model, my feeling is that we should increase acceptance rates, because many good papers are left out.
- In an ideal world I would support increasing acceptance rates, because in the last few years we have seen a substantial amount of really good papers (even with three positive reviews) being rejected. This is not only a problem for the authors, but also adds burden to the review process when these papers are inevitably resubmitted somewhere.

Positive: Should have similar acceptance rates for short/long papers

- - Increase the acceptance rate for short papers to roughly the same as for long papers, but do not exceed an acceptance rate of 1/3 of submissions.

Positive: Low acceptance rates are harmful for novel papers/fields

- But very high selectivity weeds out new ideas in my view.
- Using acceptance rates as a proxy for paper quality is an unfortunate situation that should be counteracted as much as possible. It leads to the incentive to do research that fits in with dominant paradigms etc., which seems to me to be unhealthy for the field.

Positive: Important to prevent others from flag-planting

- it is really hard indeed for reviewers to not appear as gatekeepers. In this time where many people submit their work to arXiv, double blind peer reviewing is dying and is sometimes coming from well established groups in other fields (thinking about the DL people flag planting their work in NLP and then submitting to conf), we need to think about how to be inclusive, 1st in terms of diversity of content and then in terms of diversity of authors.

Having higher acceptance rates (maybe something between ACL and COLING) would certainly help.

Neutral: Two clear accepts should be enough to get in

- Papers with two clear accept reviews should get in (this is AAAI's model). That seemed to be about the case for ACL in recent years too.

Neutral: Current system seems to be working

- I don't have any experience in this area so I have less confidence in my opinions. While the growth of the community raises challenges about how to scale, I still feel that the system we have in place right now works well enough.
- On the point of papers being rejected multiple times, in my experience while I may have felt frustrated at the time, in hindsight rejected papers were not ready in one way or another and the subsequent versions are substantially better. That strikes me as the system working.
- I really like the current *ACL setup; I think it strikes a good balance of selectiveness and inclusion. I also see the benefit of an unselective conference. It is not clear to me what this would buy us, though; wouldn't just as many papers need to be reviewed at TACL as there are now for ACL?
- In my experience with other fields, Comp. Ling. has it the easiest: there are not many areas in which publishing a paper at a conference is considered enough. Recent conferences have also struggled to fit as many papers as possible, with the result that talks are now (in my opinion) slightly too short. I think the ACL has so far managed to strike a good balance between quality and easy access, so I don't see a need to make things even easier.
- I'm fine with conferences being selective or even increasing selectiveness, given that we continue to have co-located workshops, which I believe maintains inclusivity.
- We have workshops and arxiv for work in progress or work that is not substantial enough for a conference/journal.
- There already are other big conferences that are less selective, and journal papers take a long time to be published. *ACL conferences fulfil a particular role, favouring the frequent publication of high quality material.
- It is fine to have an acceptance rate, but there has to be some flexibility based on the average quality of submitted papers, and I disagree with having quotas for different areas, which may lead to varying quality across areas.

Neutral: Current system is better than other models

- I primarily work in medicine. I wouldn't suggest moving to this model.
- Acceptance rate should remain about the same. Some conferences in other fields such as WSDM, KDD, Web Conference enforce a very low acceptance rate (10-18%) and at that rate, the reviewing process becomes too random. The decision to accept papers is better to be dependent on other factors not just getting SOTA results on a task.
- I like that the NLP community has a strong conference-first culture and I would be sad to see this shift to the journal culture of other fields. I often explain this to people as our field is

very quick moving and so by following the top conferences you get to see the top quality work at the forefront of the field.

Neutral: Clarity about why papers got rejected with high scores or accepted with low scores is important

- Recommend setting overall score that determines acceptance / rejection and using it to determine the acceptance rate. Papers that fall outside this range (accepted with lower scores or rejected with higher scores) should require a meta review with reasons for the decision.

Neutral: It may not be possible to simply threshold on quality due to logistics

- Re. no. 29: While I think that ideally the acceptance rate would be adjusted based on the quality of submitted papers, I think in reality there are logistic issues that constrain the number of papers accepted.

Neutral: Acceptance rates should be adjustable

- If it was my decision, I would probably estimate an acceptance rate, and then adjust it based on the quality of the papers that year.
- Obviously, sub-par papers should not be accepted just because there are still posterboards left. It is probably a good idea to decide on a rough number of papers that can be accepted (based on the venue), and then the number can be nudged up or down a bit depending on quality. It seems challenging to scale this to 2000 submissions, though.
- I don't think that deciding the acceptance rate in advance is in anyway useful, one should make a decision for each paper independently from the other submitted manuscripts.
- I'm in favour of not setting the acceptance rates beforehand, but accept all papers above certain limit.
- - Increase the acceptance rate slightly as justified by quality, but do not exceed an acceptance rate of 1/3 of submissions.
- A paper could and should be accepted whenever: (1) the results or claims has been convincingly verified, and (2) the results or claims are worthwhile for publication. I cannot see any reason for a pre-defined number for either acceptance rate or counts.
- Acceptance rate has nothing to do with how good the papers are. We should ignore it.
- I strongly feel that acceptance rates should depend on quality. If, for instance, the current increased popularity of our venues has led to more papers far below standard being submitted, the rate should clearly go down. If authors start self-filtering and only send top-notch work, it should go up.
- The acceptance rate should be about the quality of the work. The reviewers must not look at the rate as this might lead to unduly harsh decisions or feedback
- The acceptance should be based on the quality of the paper itself.

Neutral: Should target number of papers

- I think that there should be some combination of number-of-papers target and max accept rate

Neutral: Changing acceptance rates in itself may hurt people in the future

- The drastic change of acceptance rates (especially to be higher) might harm the evaluation of papers presented at past conferences. If the acceptance rates of future conferences are changed to be higher, the past papers will be possibly underestimated. If future conferences are not selective, the past conference papers will be ignored. Because making *ACL conferences more selective may destroy the future of many young researchers, I would like to suggest to keep the borderline quality (rather than just considering acceptance rates).
- If we did this, we might want to consider changing the names of the conferences, since otherwise the existing notions of conference-paper quality might be applied to such non-archival / lightweight productions.

Neutral: We should focus higher acceptance on specific topics

- If we promote research papers and limit purely technological papers, the pressure on conferences will be weaker, and the selectivity rates can get higher, without diminishing the quality.
- The duration of a conference is decided before the call for papers is sent out. A successful conference is one that forms a community. There will be the popular topics, which will have a lower acceptance rate, and the less popular but still important topics, which may have a higher acceptance rate. One does not create or support a scientific community by focusing on acceptance rates. The community is created by getting people talking about the broad range of topics that make up that scientific study. Consequently, the conference should be organized to maximize the amount of community forming for the time that it has, and the way to maximize is to know the maximum number of papers that can be presented and published.

Neutral: Transitioning to a journal system is impractical

- As a practical matter, there is no way our journal system can handle the reviewing needs we have.
- Not too sure about these factors. Switching over to journals is a much bigger project that will take a decade or more to implement.
- Unless there is a new, faster moving journal model proposed, we cannot diminish the value of *ACL conference papers. I believe it is a strength of NLP that we can submit and publish relatively quickly compared to other fields because the nature of the research is extremely fast paced. I'm certainly open to alternatives that keep or even increase the current pace of things (not just arXiv either... peer-review is important), but I definitely prefer *ACL conferences to the existing alternatives.

- The reason for keeping conference papers selective is that, in computational linguistics, the research is often outdated by the time a typical journal review cycle elapses. Therefore, conferences are serving the same goal as a journal review.

Negative: Selective conferences as a measure of quality control is important

- However, they shouldn't be *much* higher because the current model archives publications and we want to have a "seal of quality" on those.
- With the high volume of papers nowadays, I think it's important that only high-quality papers are accepted in major conferences like ACL in order to make it easier for researchers to find the good ones.
- no selectivity was done for many years at LSA (but I don't know if that still holds). I favor selectivity because otherwise the conference can end up with really poor papers. But very high selectivity weeds out new ideas in my view.
- As long as we don't take over the publication mode of e.g. medicine where only journal papers count, we need to ensure quality of conference papers.
- We can't have the cake and eat it: as we (the field) practically stopped publishing in journals, our only chance to maintain scientific rigor lies in much more thorough conference reviewing. This implies, inter alia, more selective conferences.
- I'm strongly against the idea of making the conference publication non-selective. I'll shift field if we make this happen. CS is a fast-turning field and journal reviewing cycles are too long to capture w/ this pace.
- I am strongly in favor of having selective conferences. We already have workshops and arxiv for fast dissemination of result. The relative stamp of quality and fast turnover that conferences provide is invaluable.
- As is, it is difficult to keep up with the pace of work appearing at conferences, which would only get worse if they become less selective.
- This makes it harder for everybody to find the better works to visit their talk/poster. One can only see a handful of papers in each session.

Negative: Higher acceptance rates are impractical logistically

- With the number of submissions increasing year-by-year, one of the biggest bottlenecks is providing space to allow for good quality papers. Over the last few years, AAAI has done its best to maintain an acceptance rate of about 25%. In 2017, it accepted about 638 out of 2600 submissions. In 2018, it was about 933 out of nearly 3800. But in 2019, when submissions went through the roof (at around 7095), AAAI accepted only 1150, for an acceptance rate of 16.2%.

Negative: Low acceptance rates are important for jobs

- As people in NLP move more into other academic fields, acceptance rates become increasingly important for tenure decisions. The current model is already hard for these decisions and for grants, where journals often trump everything.
- However, my understanding (from a past ACL meeting) is that keeping acceptance rates <=25% is a prerequisite for conferences being highly ranked in rankings such as CORE. This

is a big deal in my country (Spain). Conference papers without a CORE A/A* ranking do not count at all for hirings, promotion, etc. Since NLP journals are sadly few and not especially well-ranked in the JCR index (also important here), if conferences were no longer highly ranked in CORE, this would be a huge blow for Spanish NLP researchers and probably put the whole field at risk in the country (of course the best solution would be to change silly local bureaucracy, but that's not going to happen). I think it's the same in some other countries.

- ACL is famous for its stringent reviewing, and it has allowed us to count *ACL/EMNLP papers as journal-equivalent in promotion and tenure cases. THIS IS VERY IMPORTANT and dropping it would wreak havoc.
- This is a complex issue. In my country, PhD students progress is evaluated based on the number of papers conforming to certain criteria they manage to publish and the student needs at least 3 papers to defend. For instance, ACL conference papers count but workshop papers do not. This all depends on where the proceedings are indexed. If the conferences would be more inclusive but as a result the proceedings would not be indexed in scopus then the ACL conferences would not count anymore as valid venues. Thus, the student would have to publish 3 journal articles (TACL?, CL?) which would make the process more difficult. Compared to medicine, there are much less journals in our field and that's why I think similar process simply would not work immediately. If there would be a larger number of generally good quality journals on different levels then maybe making conferences more inclusive would work.

Negative: Selective conferences are important to get lab to support travel

- If the conference was not selective I don't think my lab would routinely support registration and travel costs for presenters. Currently, we rarely support non-presenting attendees. Since we would then need a bigger open access publishing budget for journal articles we would probably cut the travel budget to keep overall budget the same.

Timing of Reviews / Resubmission of Papers

Positive: Having more time is important to ensure that resubmitted papers increase in quality

- Having several deadlines a year means papers can (and should) be adjusted based on feedback and resubmitted, assuming that reviews suggest a way forward (and the paper is not a complete loss). Giving people more time would allow for that. And with lower acceptance rates, this would become a better model
- I think a short turnaround incentivizes people to make minimal revisions and essentially submit a reworded paper that was just rejected. I've personally had success doing that, but I think that it was because of the extremely high variance in reviews which I don't think it is ideal. Ideally, authors will have adequate time to thoughtfully respond to the reviewers' comments in the next version. That would show that the reviewers are actually valued and not treated as a random draw from a distribution of paper scores.

- If we do not give authors ample time to improve their work, many will re-submit to the next conference (which is likely to be less selective) and hope for the best. This is a disincentive to do good science.
- I think it is important for the community. We should give a chance to the authors to address to comments / improve the paper. We want to reduce the proportion of papers resubmitted as is.
- Setting up and running another experiment etc. usually takes a few weeks.
- However I realize 1 month+ is often not feasible! 1 week is too short; just encourages no/cosmetic changes and re-entering to hope for randomly better reviews. (Could we pass the reviews from the previous conference on to the next one and require a statement of what's been changed?? More like journal re-review I suppose)
- I would like it to be possible to actually improve a paper, eg: run some additional experiments based on a review.
- If there is enough time between reviews and the next deadline, authors may have more time to improve papers and submit something more valuable to the next conference. It's in the interest of nobody to see the same papers getting rejected multiple times in *ACL conferences.
- since short time encourages resubmission with minimal / no changes.
- As mentioned earlier, it will help a lot if reviews are seen by the author before they submit their work for the next conference. Their next submission would be better since they would have a month fixing the drawbacks of their earlier submission, rather than just a week. As a suggestion, if the area chairs and program committee do not wish to have an author response period, at the very least, it would help, if the authors see the reviews a few weeks before the final accept/reject notification decision is made. This is, of course, assuming that there is little to no reviewing done on the last fortnight before the notification deadline (around the same deadline as the author response period).
- The current time between the ACL notification and the EMNLP abstract submission is only two days. It is difficult to imagine any useful changes that can be made within such a short period of time. I worry that this will result in ACL rejects being resubmitted to EMNLP and just contributing to the unmanageable number of submissions.
- Reviews released 1 week before submission encourages low-effort "flips." I think having more time would be beneficial, even if it means a NAACL paper can't be flipped for ACL.
- The schedule is far too back-to-back this year, and folks are just serial resubmitting (partly due to the randomness of the reviews this year with the Toronto Paper Matching System), which only increases everyone's load.
- Back-to-back submissions are a poor way of handling the fact that conferences are the only way to validate your work quickly. Would be solved by the field putting more emphasis on (shorter and quicker) journal publications

Positive: Having more time between resubmission is important for giving reviewers time to rest

- We have a crazy schedule that exhausts reviewers, area chairs, etc. Focus more on the wear and tear on reviewers and this will lead to a better situation for authors. Reviewers get minimal benefit from their labors, and yet are asked to review over and over again with few breaks. this is not sustainable.

- If we want people to improve rejected papers before resubmitting, we should have them time.
- Most of the time the reviews are asking for more than polishing a rejected paper. Most of the time the reviews are more about experiments and analysis. Which needs time to be addressed properly.

Positive: Would like conferences to be more evenly spaced throughout the year

- Given the limited number of *CLs and associated conferences there should be enough room to space them out with more than a month's gap between review notifications of one and deadlines of another; an alternative (and proposed, I think) solution is more conferences? e.g., the EACL
- Would be good to have better spacing thorough year
- It's a bit ridiculous that all three major conference deadlines in 2019-2020 occurred in less then a 6 month time span. What about work that is completed the other 6 months? Spacing the deadlines roughly evenly throughout the year allows ample time for review and buffer before the next deadline.
- I *STRONGLY* feel that EMNLP and ACL should be 6 months a part so that the quality of papers is spread between both of them.
- I also think it would be better to distribute submission dates evenly over the year.
- NAACL/EACL, ACL and EMNLP should be evenly distributed across a year.
- It would be good to have deadlines spread more throughout the year. This encourages less deadline-driven science where you submit when your paper is ready rather than to EMNLP because the next deadline is too far.
- Main *ACL conference more evenly spread out throughout the year, so that there's always a deadline nearby.
- - Submission deadline for the first ACL* conference should occur earlier in the fall to enable more time.
- Fixed, equally spaced deadlines would be great.
- I do hope that the *ACL conference deadlines can be more evenly distributed throughout the year.
- Why are ACL, NAACL, and EMNLP submission deadlines all in Dec-May, with none of them in the other half of the year? Even if these conferences remain in the summer/fall, perhaps there could be an early round of review in the fall, so that strong papers get accepted early and others get a revise-and-resubmit for the normal round of review. Authors could be incentivized to submit early by stipulating that (a) the acceptance rate for the early round of review will be higher, and (b) authors of papers accepted early could choose which conference they want to present it at. This would hopefully spread out reviewing responsibilities over the year and give more flexibility to authors who cannot afford to travel to certain conferences. Of course, it would require management of an extra review process and, ideally, infrastructure to ensure reviewers of revised-and-resubmitted papers can see the original reviews.
- We should strive for having the deadlines of conferences uniformly distributed in the year. Some years like 2019 have all the conference deadlines in the first half of the year. A

conference schedule that does not compresses conferences in one semester makes the waiting time for publishing a paper shorter.
- Especially since most deadlines are Dec-May, it can be tough to properly revise and resubmit to the next conference when the turnaround is sometimes 1-2 weeks.
- There are 3 or 4 major NLP conferences each year. How difficult is it to have the deadlines distributed evenly throughout the year with enough space after one notification and the next deadline? e.g., currently naacl,acl,emnlp deadlines all happened in 5 months and there is likely no deadline for next 6 months.

Positive: Workshop deadlines should also be considered

- Also take into account workshop deadline timings if possible.

Neutral: An alternative is to release the reviews earlier, through author response or not

- One way of providing feedback earlier than notification time is to pre-release reviews to authors while discussions among reviewers and ACs/PC are still going on. Note that in case author response happens (which I do not favor), authors are also informed about reviewer views early on. This can inform authors well ahead about possible criticism of their work, and help them to improve their work while awaiting decisions. This can be profitable for the time to camera ready version on acceptance, or submission to another conference in the case of rejection. Note that authors do not need to retract papers as long as final decision is not taken.
- If there is author response, we can use more time to revise the paper since we usually get the initial reviews 2 weeks+ before the notification.
- Perhaps we could try releasing finalized review comments at some point before the final decision?
- If there is no author response period, there should be a way for authors to get preliminary reviews around the same time the response period would start.
- It could help to send the content of reviews out early (maybe without scores), i.e. as soon as the PC discussion period ends and the reviews themselves will no longer be changed. This way authors can already start working on improvements whether this is for their camera-ready version of the current conference or an improved version to be submitted elsewhere.

Neutral: The issue is complicated by the arXiv blackout window

- This is also complicated by the arxiv blackout window, obviously.
- If there won't be author response and the ACL policy of anonymity windows remains in effect, there must be more time between decisions and the next *ACL conference, ideally though faster turnaround (e.g. NAACL took 3 months!). Right now, this is putting papers under continuous embargo until they are accepted.

Neutral: Should have controls on people immediately submitting rejected papers

- I think we might want some controls on people immediately submitting rejected papers to the next conference, in order to reduce the JND publishing. But I'm not sure about this.
- I would argue against having people resubmit the same exact piece of work before resubmission.
- We should have a rule so that a rejected paper cannot be submitted to a different conference 1 week later (i.e. like in TACL). We will save reviewing time and improve submitted papers quality.
- it would be nice if we could prevent resubmitting papers which got a clear rejection in one conference to the next conference, unless indications provided that a major improvement was made.
- I would favour measures that discourage people from 'scrambling to submit'. One option is to release reviews+decisions for the previous submission deadline after the deadline for the next conference (i.e., so that authors have to wait for the next next conference to resubmit).

Neutral: Reviews should be sticky for the next submission

- And explicitly ask (and check on the system) if a paper is a *ACL resubmission, and may be also circulate the review. Having a bad paper reviewed by three person is a waste of time; having it reviewed by six or nine experts is not acceptable and burns a lot of energy that we do not have (given the current submission numbers).
- A great step forward could be to keep track of resubmitted papers. Having a paper submitted along with its previous reviews (and a potential response which indicates how the authors addressed the concerns) could greatly alleviate the review load. Though this would require serious book-keeping which might be hard to implement in practice.
- Authors must declare whether a paper is a re-submission or not. If it is a re-submission, reviewers are automatically provided (1) the previous reviews and (2) a pdf diff of the re-submitted paper and the originally submitted paper. This allows them to very quickly determine whether the paper has had any major changes to address the previous reasons for rejection. Of course if the previous reviews are rubbish the reviewer can disregard them and write their own review.
- Again, let's leverage TACL and push hard for rolling reviews.

Neutral: Review quality is a more important concern

- Encouraging authors to "flip" minimally revised papers to the next conference is useful only when we think the reviewers are likely to be biased/mistaken. I'd rather have a higher-quality review process than simply assume reviews will be noisy and encourage researchers to roll the dice repeatedly.
- Low quality reviews discourage me from playing the game over and over again with a submission. High quality reviews but a rejection encourages me to resubmit but trying to within a week or 2 seems rather rushed...

Neutral: Reviews and decisions being released quickly is important for travel arrangements

- Funding to attend hinges on presenting a paper. Cancelling travel plans can be extremely costly, and so is only having 4-6 weeks' lead time to make the plans. Not all attendees of ACL have their employer handling travel for them, many have to pay up front and be reimbursed (especially students, who make a pittance of a salary to begin with). I know I've personally had to pay money out of my own funds because the late notices of some *ACL venues (conferences, workshops, etc all are guilty of this) because the travel costs at that point in time exceeded what the employer/funding organization budgets for conference travel, due to the closeness of notification versus the actual date of the conference. To make the conference more accessible, we should have reviews and decisions be released as quickly as possible, to allow for people to be able to make afforable and reasonable travel arrangements.

Neutral: Bizarre that papers get released before next conference

- In recent years, accepted papers from the next conference (e.g. ACL 2019) are released even before the previous conference starts (e.g. NAACL 2019). This is at least strange, as at NAACL there will be presentations about papers that may not be state of the art anymore...

Negative: People shouldn't feel that they need to submit to every conference

- I think it is a problem that people feel like they need to submit to every single conference in the field. As we have more conferences (AACL, EACL going annual) it will become impossible to keep the deadlines non-overlapping and this is fine. The purpose of having more conferences is to accommodate the larger number of people in the field, not more submissions per person!
- My group targets one or two meetings a year. That's already too much carbon. People that need to submit to everything are doing it wrong.
- With so many NLP conferences a year, we don't have to wait that long for the next possible submission ;-)
- I submit my work only when it is ready. I can wait for next conference to make my paper better. Why do we need to rush?

Negative: Short turnaround is good

- This year has been great in terms of reviews and next submissions time (NAACL->ACL-> EMNLP). I especially like the fact that at EMNLP, the deadline for workshop submissions was set just after the author notification of the main event.
- ACL has enough conferences to stagger things to always have another deadline soon after the rejection comes through. There's too much of an emphasis on summer conferences and not enough taking advantage of the 4-6 major conferences out there. Sure the more prestigious conferences can be in the summer, but it's great to always have another deadline around the corner.

Negative: People should not have to worry about reviews to improve their research

- The research should continue regardless of whether or not there is a review on a paper. The next conference submission should represent an improvement to the results that were previously reported. A good researcher knows the weaknesses of their research without the reviews, and would know how to improve their results independently of the reviews by the next paper submission due date.

Public Review

Positive: Yes, but it should be voluntary

- I think it should be voluntary choice for authors and also with agreement from reviewers
- Have reviewed for *journals* with open review, and support it in that context, but don't necessarily believe that it should be all or nothing.

Positive: Would like to try once

- I would like to try.
- But I also think "If it ain't broke, don't fix it!". So rather than moving all our conferences to the ICLR model, perhaps we could try it with just one of our conferences, such as CoNLL or EMNLP. Then if it works well and the community likes it, we could use it for more conferences.

Positive: General positive comments

- Why should it be private?

Positive: Is good for being able to release research results more quickly

- The ACL anonymity rules largely prohibit sharing of work from one month before submission until three months after submission (or until the end of the review period). The typical timespan for an ACL paper today is 3-6 months from start of project to submitted paper, so this leaves a significant risk of duplicate work. For example, there were very likely some papers submitted to ACL 2019 that are very similar to papers submitted to NAACL 2019 where the authors of the second paper were unaware of the prior work during the 3-4 months when the work was performed. For the authors of the later ACL paper, the feeling of finishing and submitting a manuscript to ACL a week after the NAACL accepted paper is posted on ArXiv would be demoralizing (as that is when the NAACL paper would finally be able to shared publicly). This is especially problematic when the NAACL paper was finished several months prior! Being able to do a proper literature review on papers already written

while in the early stages of a project is essential to good scholarship and avoiding duplicate work.

Positive: With papers already being public on arXiv, it'd be nice to have systematic review

- It seems like people are already doing something like this by posting on arXiv. It would be helpful if the reviews for pre-prints were more systematic.

Positive: It might make review quality better

- Reviews should be criticized too. The quality of reviews are becoming worse and I believe keeping reviews private is one factor.
- In principle, open feedback on papers is a good thing, though I agree with the last two cons. I do not agree with the first two cons, on the contrary, it is an illusion that peer-review prevents inaccurate statements: in the public setup, it is at least clear that the work has not been reviewed yet and since more people may look at it, chances of problems being seen by at least one person are bigger.
- Another advantage not listed above is that it may encourage reviewers to give higher quality reviews.
- I would love to have a review process like Open Review. It improves the overall paper and review quality.
- I like the idea of Open Review, more to hold reviewers accountable. Currently, there is no incentive to actually take the time to try to understand a paper, thus mainly work that presents 'obvious' outcomes such as pushing state-of-the-art on a given task rises to the top. As an author, receiving negative feedback is part of the research process, so it shouldn't matter to the author if it were made public if they truly believe in the work. However, as a reviewer you must hold yourself to a higher standard, which I often fail to see.
- Public review can help rule out rubbish submissions and un-qualified reviewers.

Positive: Public reviews can be useful as a reader of a paper

- I like the ICLR approach. Sometimes I find the ICLR reviews and meta-reviews more useful than the paper itself. Often the reviews and meta-review clearly summarise the strengths and weaknesses of the paper better than the paper itself. If reviewers are going to provide excellent reviews, it seems a shame not to make them available to a wider audience.

Positive: It could be a good way to recruit new reviewers

- Public reviews could be the best pool to recruit new reviewers from!
- It may increase the time span for the reviewing & decision process, but it is a good way to reduce incorrect decisions. Now that we are currently in shortage of good reviewers, why not just open up to the entire community as well as related fields?

Positive: Public discussion after acceptance sounds good

- I support having public discussions of papers after they are published (which is common practice in medicine, for example). So if someone sees a major problem with a paper, he/she can express these concerns and the authors can respond if they wish

Positive: Could help continuity

- It seems like this would be a helpful way to ensure continuity of reviewing if a previously-rejected paper is resubmitted. As it stands, the previous set of reviews are forgotten. This would incentivize more careful reviewing (since it is public) while also disincentivize immediately resubmitting a rejected paper without making meaningful changes, which would also cut down on reviewer load.

Neutral: Semi-public for only ACL members w/o COIs might be good

- What about semi-public review? Any ACL member can log in and comment on any paper w/o a COI. This would diminish the cons of negative comments and inaccurate statements reaching a wider audience.

Neutral: It's complicated

- You've bundled a bunch of different issues into this. I said previously that I'm very much in favor of public discussion in the openreview style. The one thing I'm not sure about is the _participants_ in that discussion. I'm not sure about letting random people on the internet give comments that attach themselves to my papers. Random people on the internet probably won't care enough to do this, but it seems odd to allow anonymous random people to attack papers. The benefit of the review system we have is that we think the anonymous reviewers at least know something of what they're talking about when they post reviews. On the other hand, if the random anonymous person has good points, that adds value to the discussion, and I don't want to just exclude them if they _do_ know what they are talking about. It's complicated.
- I can see arguments on both sides for OpenReview. Maybe it is worth trying.

Neutral: If it happens, double-blind is essential

- I think public reviewing is hard to reconcile with double-blind and would introduce biases. I think ensuring that the process is as double blind as possible, and ensuring fairness, are crucial.
- I very strongly believe that double-blind reviewing must be maintained. I know that arxiv makes it difficult already, but we do observe in neighboring disciplines that famous authors can get papers of moderate quality accepted more easily, and we should fight this as well as we can.
- The main problem with public review is that it is not possible to preserve anonymity across multiple conferences or reviewing cycles in this model.

- This is a hard question. I favour public discussion. However, you cannot keep the paper anonymous for too long after making it public. So, in the next (resubmission) round the paper will not be anonymous. The authors should have a chance to preserve their anonymity if they choose to (important for smaller labs, less known researchers). I am afraid I am not willing to give up on anonymity in order to enable public reviews.
- I believe the benefits of double blind review outweigh the costs.
- If the authors are not public, then i am ok about public review. Otherwise, public review will much help the “famous” authors and will not be fair.
- We can preserve double-blind with Public Review, and do so better than ICLR. At a time when reviewing quality is suffering, more openness, really does allow work to be better vetted—-(not just “theoretically” as you say above).

Neutral: It might not be used widely

- This is a good idea in theory, but I think (not sure) most papers are not discussed, so the value of it is limited.
- Seems like it would be very unevenly used
- For all the discussions about it, I'd like to see how many papers actually give rise to a discussion, let alone an interesting one.

Negative: Bad reviews will follow the paper around forever

- negative reviews follow rejected papers around fireverr, making it hard to get iid reviews
- If all of the reviews authors recieved were of high quality I'd be in favor, but given the handful of times I've gotten absolutely terrible, unhelpful reviews back (and not knowing how ACs allowed such a thing to be released), I'm not sure if we'd want that kind of public impression of researchers in our field…

Negative: This may hurt diversity / unfairly favor particular groups

- This is problematic for increasing diversity.
- Public review favors the extroverted and self-assured, on the side of both authors and reviewers

Negative: This may be stressful / discourage submissions

- This sounds stressful.
- Public reviews put pressure on authors to respond immediately to public comments, urgently run multiple experiments required by commenters to address their critique, even if the authors don't agree with the critique. Also, public comments may bias the reviewers .
- - Many would not submit if public review is adopted. It is entirely counter-productive.

Negative: Could hurt early-career researchers

- Public review highly discourages submissions from students and junior researchers
- - Public review can hurt early-career researchers and students' careers. Please do not do this!

- I also worry about new entrants to the field: it can be daunting enough to get reviews anonymously from experts on your work, but I can imagine many people (and perhaps more likely underrepresented people) feeling very uncomfortable about posting something for the world to see, especially since I suspect you're more likely to get unconsidered negative comments in this situation.

Negative: Reviewing will become a rich-gets-richer popularity contest

- With this mode, reviewing becomes a gimmicky popularity contest that will be dominated by big players with strong PR machines and a wide reach, rather than judging objective quality. This is already a problem in citing work (the same work done by a bigger lab will get more citations than that from a smaller one), and it would exacerbate this rich-get-richer situation exponentially. All cons are magnitudes more serious than the tentative benefits of this proposal, and would seriously damage the reputation of NLP as a mature academic discipline.
- I am generally suspicious of the court of public opinion, so this process scares me a bit, to be honest.
- This seems to lead to popularity contests in other conferences I have seen. While I like the incentive to have work ready to be seen publicly when you submit, it doesn't seem to actually improve the review quality at all.
- and also that it makes the field even more of a popularity contest. (Possibly less so if author names are not revealed during this period, but I'd also feel weird about having my work open to the public yet anonymous.)
- with public reviews we need to think of a way to allow the same chance for every paper. My concern is that similar to publishing in the arxiv, papers from well-known, strong groups have a major advantage over papers from unknown authors - they can easily reach more people and obtain more support.
- Public review does not encourage diversity and is biased towards authors from 'famous' institutions and specific social groups. This model also has huge potential for abuse.

Negative: There is no filter on the quality of the comments

- I don't favor public review before papers are accepted because one does not know the quality of the public reviewer. That person might have an interesting point, but that's a lot of extra effort for some likely small gain.
- It's a nice idea in theory, but it only takes a few assholes to spoil it for everyone — and they are sure to do so.
- Reviews should be partially supervised to prevent inaccuracies or trolling. This may be done through some automatic (NLP) method, quick human confirmation of each review, rating the usefulness of reviews by review readers (providing reputation to the reviewers), or other reviewer metadata.
- Papers should be judged by experts, not by the public. A public opinion, if given, is irrelevant in this case.
- Reviewing should remain a controlled process: reviewers are supposed to have some expertise in the field, which commenters will not necessarily have. I don't think multiplying the number of lines written about a paper helps to indicate that a paper is sound.

- I have seen examples of public review mostly from ICML, and I believe that there is less to gain from the court of public opinion than from established reviewers.
- More data does not necessarily correspond to more information. At this time, giving my opinion about the average quality of reviews I would add more noise to scientific communication.

Negative: It is better for papers to be reviewed before they are made public

- I have seen a lot of papers on Arxiv that would have really benefited from being reviewed first. I think reviewing serves a valuable purpose, and it is much more important to get the paper right before disseminating it than to get it out fast.
- It encourages public flogging. People sometimes make mistakes and should be allowed (and even encouraged) to fix them at their own discretion.
- I can see potential upsides, but I'm concerned that this encourages people to make work public before it's really ready to plant their flag,
- Sometimes the best scientific ideas do not review well. Protect the authors by allowing them the opportunity to mature their ideas in a confidential review process. After a paper is published, it has an eternity for the public to review and respond to it.

Negative: It is too easy to game / will cause turf wars

- Too easily gamed / faked
- Public review is often misused. Behind the anonymous comments, there are often other authors that try to bash the work from others, instead of providing objective and constructive comments. Public review would be useful if the commentators are required to register under their true name.
- in theory, I'll be pro open reviewing but one could imagine a certain crowd targeting a paper because a) they have issue with the author(s) b) they got incentive from his competitors c) they're working on the same thing so they don't want to be scooped. If we know how to circumvent this, why not ?
- Public comments should be disallowed before official reviews to prevent astroturfing and trolling.
- It seem impossible to avoid conflicts of interest with public review. What would stop anyone from reviewing their own or their friends' papers?
- Great idea, but I think tends to amplify turf wars and takes up time we don't have.

Negative: General negative comments

- Are there really conferences that do this? Sounds like a pretty terrible idea.
- I don't think this is needed now. It would raise more issues than it would solve problems.
- Doesn’t add much value.
- This sounds awful
- i can only see disadvantages for ACL

Negative: It is redundant given that people can post papers on arXiv

- The two pros points can be obtained by publishing a paper on arXiv, so I don't think that the public review is required.

Negative: It may take too much time

- Another issue with public reviews, e.g., in ICLR, is that the review process is very long.

General Comments

Neutral: The review process should be blinded as much as possible

- Please make reviewers blind to each other and to AC as well.
- The discussion should not be influenced by the ACs or the ACs should be blind to the author's identity.
- Meta-review should also be blind. That is, meta-reviewer should not know the identity of the reviewers. Even for program chair/meta-reviewer there would be a tendency to give more weight to reviews from more senior people, which would not always be the appropriate weighting approach (if they used a secondary reviewer not mentioned, or if they didn't have time to read as carefully as a less senior reviewer who had more relaxed time constraints). Meta-review should be based on the quality of the review and the level of details in the review.
- Area chairs and program chairs should be *blind* to the authors of the papers they are leading the discussion for. Conflicts of interests can be handled through other means (e.g. listing at submission time all reviewers and self-selecting the authors they have conflicts with; if this is not done faithfully, reject without review - similar to WWW or ACM MM)
- I think that the review process should be as anonymized as possible. That is, also the identities between reviewers should be unknown, as only the content of the argument made should influence the decision whether a paper is accepted or not, and not social factors related to hierarchy and power.

Neutral: The review process should not necessarily be blinded between reviewers

- I believe the reviewers should be able to know who each other is while writing their reviews. I've differed to folks more experienced than I am in subareas, and I've also fought more diligently when I've seen students be overly harsh on papers.

Neutral: Some measure of how experienced or negative/positive reviewers are would be useful

- Scores are always subjective, so as few scores as possible. Can we analyse reviewer bias post-review so if you get 3 harsh reviewers (below the median) you get some addition to

your score and if you get 3 lenient reviewers (above the median) you get a substraction from your score?
- More generally, I think the qualifications of each reviewer should be made available to the review committee/authors/PCs, so that the very high and very low scores given out by folks very new to the field who may not even have a PhD yet can be tempered by the more experienced reviewers.
- In order to include junior researchers early-on in the reviewing phase while still maintaining high standards, I would favor a reviewing system for junior reviewers that starts with a freshman status before admitting them as fully responsible junior reviewers:

Neutral: There are bigger issues than the reviewing process

- Reproducibility is the primary issue!!!

Neutral: We should be more open to diverse work

- ACL should increase the diversity of the nature of the research it accepts. Rather than just accepting research related to the primary technological approaches of the day, it should allow a more diverse range of research that is relevant to the topic, e.g. application of the technology, participatory and action research. This will expand the community's awareness of issues related to the research and technology as well as allow for greater multi-disciplinary research to be shared within the community.
- - Ensure that there is more balance in the accepted program for attendees who come to computational linguistics from linguistics/behavioral and cognitive sciences.
- - In 2019, it seemed as if multimodal papers, with a prominent language component but which also reported results separately on other modalities, were penalized at NAACL. This was disappointing considering that there is increased interaction among researchers across fields. Maybe the multimodal topic should be made clearer or review criteria for multimodal papers should be commented on in the CFP?
- We need to move away from a one-size-fits-all publication and reviewing model.

Neutral: It may be impossible to sustain double-blind review

- As much as I regret it, I am afraid that we will eventually have to give up double-blind review. The anonymity period was a reasonable compromise to protect double-blind reviewing, but there are clearly many links and parts of the community seems dissatisfied with this policy. If we can move towards more journal publications, this will be less of a problem.

Neutral: Reviewers are assigned too many papers per conference

- I have recently declined to be a reviewer in ACL conferences because of the large number of papers allocated for each reviewer. I would be more willing to participate if I had the option to choose the number of papers to review for each conference.
- The review load for ACL 2019 was extreme. The meta-reviewers pushed it over the limit, unnecessarily. I truly hope next year their involvement will be pushed down to something

useful. Papers that are a clear reject do not need meta-reviewing. An agreement on a reject is an agreement, settling on whether is a 2 or a 3 is inconsequential.
- Review load is getting too high, anything that reduces that load should be experimented with.

Neutral: We should have more venues for publication

- We need more conferences, more paper slots and/or more TACL to accomodate the increasing amount of good papers incoming. Papers shouldn't be rejected with three positive scores as is happening later (it's even dubious to reject with two positives and a negative in most cases). However, keeping acceptance rates as they are is important for countries where conference rankings are essential for hiring, promotion and grants (as is the case in mine). The better solution therefore is more venues for publication, be it in conference, journal, or hybrid (TACL) form.

Neutral: UX improvements are necessary

- Submission / reviewing UX should be much cleaner. All profiles should require setting up TPM before anything else happens; should be clear who will be able to see identity when e-mailing, and all e-mails should be default include appropriate link for next action.

Neutral: It would be nice to more systematically log reviewer disagreements

- It could be useful to have a system whereby the reviewers can highlight comments in the other reviewers' review points they either agree with or strongly disagree with. It could help the AC to understand where the agreements and disagreements lie. Also, if a reviewer has grossly misunderstood a work or made a remark that is erroneous or unfair (this does happen), this would enable the other reviewers and the AC to easily catch out these cases so that the paper does not get unfairly penalised.

Neutral: The overall score range is important

- One area not covered here is the overall score range (1-5 or 1-6). I am strongly in favour of 1-5. If a change does need to be made to reduce reviewers giving 3s then having half-points would be better than shifting the scale.

Neutral: Would be better to have some quality control before full review

- I fear having public reviews would in the end substantially increase workload, we should try to find an option to decrease reviwers work maybe by doing minimal quality control before full reviews.
- It may be worthwhile to consider different levels or aspects of review that are dependent on an overall strength of the paper. For example, a very weak or deeply flawed paper may not require in-depth technical feedback. A more dynamically-structured review may help save reviewer time, and also encourage more meaningful in-depth reviews for papers that are strong and/or borderline.

- I've encountered quite a few papers with unacceptably low quality of writing. I believe that allowing for some kind of basic check for this, instead of full reviews of all papers, would possibly lower the overally reviewing workload.
- Area chairs should be allowed to desk-reject a proportion of submissions with clear weaknesses that will make them being rejected for sure, such as (really) bad readability, trivial results, inappropriate evaluation metrics, etc.

Neutral: Concerns about desk-rejection of papers

- Papers should not get rejected without review if they fail to meet submission criteria, such as including an appendix as part of the main submission and not as separate file. Authors should be given up to 1 week to resubmit in the correct format before being auto-rejected.
- Please explicit. allow "Supplementary" section inside the paper! We received a desk rejection. Having the supplementary work separated from the main paper makes absolutely no sense, because in the final paper version the supplementary section will be included. This policy is very ridiculous and should be reconsidered.

Neutral: Hope for the new senior AC setup

- Hope new senior AC setup helps to improve review quality

Neutral: There is too much power in the hands of ACs/PCs

- I am worried that the independent numeric score assigned by reviewers is something that seems to be being downplayed currently, which concentrates more power in the hands of ACs (or ACs and senior ACs). This goes too far in my opinion. I think this score should be considered as the strongest signal, and occasional deviations from it should be carefully justified.
- I think we should remove the roles of AC and PC, and place the acceptance decision back in the authors' hands.

Neutral: Reviewer quality is a major concern / training for new reviewers might be useful

- One of the critical issues is a lack of expertise. Most of the reviews are very shallow and formal. People are just take a brief look at the table with empirical results without getting the idea of the paper. This "super-empirical" approach is not very effective is such a field like NLP. There is also a huge problem when professors redirecting reviews to the PhD students who are actually quite new in a field. This happens very often since the number of submissions grows quite fast last years.
- There should be a good way to select qualified reviewers and henceforth ensuring review quality. For example, area chairs could prepare for some quiz questions that whoever review for that area should be able to get a score at least 3 out of 5, or just let every reviewer candidate to write review comments for one well-known published paper in that area.
- Maybe we can provide examples of good reviews as a training resource for junior reviewers? I like the structure more given how many new people we have entering the field,

but maybe we should handle this by having some 'reviewer training module' for first time reviewers rather than over-structuring our review forms.

- Everything during the reviewing process should be made available to the authors, including names of the reviewers, and the authors should decide to make it public or not. It is also important for further discussions. I cannot think how this can be a problem. It is time for reviewers to be more responsible. I see too many of my colleagues reviewers that do not care at all and write reviews of 6 papers in less than one hour.
- The fact that it's well-known that reviewers must be spoon-fed doesn't speak well to the current state of affairs, but this is largely, I believe, due to the saturation in the field, which would be difficult to solve. Also, the line between engineering and research is almost non-existant, so it's hard to see where progress is actually being made.
- A problem is that there lack enough qualified ACs and reviewers.
- Reviewers are overworked, obviously, but as a result often outsource reviews to students. That should not be allowed. Secondly, only people with >= k first author pubs in *ACL conference should be allowed to serve. I am very disappointed by some of the lazy and incompetent reviews (about 1/3) this last ACL, both as a reviewer and author.
- I feel like questions regarding poor quality reviews from certain reviewers and how these can be mitigated are missing from the survey. These reviewers, in my opinion, are unlikely to be influenced by author responses or ICLR style public reviews/discourse with authors and need to be moderated by the ACs more carefully. I saw an example of this with Sasha rush and iclr reviews this past season. But in one of my experiences the AC disappointed and wrote a poor quality meta review based off poor quality reviews. Are there ideas for promoting high quality review practices, dealing with folks when they clearly just don't care?

Neutral: We should be more scientific about how we examine the review process, not do surveys

- Why is that a popularity contest? Check results you are scientists. Did it save time, did it affect the quality of accepted papers or just randomly rejected between borderlines etc.

Neutral: It should be easier to participate without a paper

- It should be easier (i.e., cheaper) to attend *ACL conferences for early career researchers who did not get their paper accepted

Neutral: We should write fewer, but higher-quality papers

- What I'd really like is for the community to be writing fewer papers, of a higher quality. The publish-or-perish culture is extremely damaging, and means we all spend time reviewing poor work. (I just think of the number of papers I've rejected from ACL* conferences in the last 10 years.)

Neutral: Like Omer Levy's proposal of "reviewers review authors accept"

- I like Omer Levi's proposal. Not sure it'll scale for 3k papers but.. see here https://twitter.com/omerlevy_/status/874141786003750917 TL;DR: 2 weeks review cycle ; If one disagrees with the reviews he can publish his paper anyway as long as the the reviews are included, if not he can resubmit with fix at a latter point.

Neutral: Bidding is important/not important

- Please restore the bidding process for ACL! It doesn't take much time, it ensures the best fit with reviewers' interests, and it is a small perk for reviewers, giving them a survey of what people are working on.
- I prefer the no-bidding review assignment procedure of ACL2019 to the with-bidding procedure of NAACL2019, but the quality of paper-reviewer matching could still be improved.

Neutral: Considering review for tutorials is important

- - Add review criteria for tutorial submissions. For example, should tutorials be able to be offered to a large group to have a chance to be accepted?

Neutral: Better instructions about expected format is useful

- It would be good to have instructions for specific paper types.

Neutral: Reviews should be less based on opinion and more based on previously cited scientific fact

- Reviewers should be held to the same standard as the authors regarding their statements - if a reviewer claims something different than the author, he should find some evidence for that claim in the scientific literature. Otherwise how can we assure the quality of the papers if the reviews are substandard?

Neutral: Would like ACL to be hosted in Africa

- When will ACL be hosted in South Africa or Nigeria?

Table of Contents

- Neutral: Similar to journal reviewing
- Neutral: Would be useful, but only with a mechanism for revision/conditional acceptance
- Neutral: Would need to reduce number of papers per reviewer
- Neutral: If reviewer response is just as bad as it is for single author response, then maybe not worth it
- Neutral: Maybe useful only in a limited number of cases
- Neutral: Need to increase reviewer discussions, not necessarily with authors
- Negative: Too much work for reviewers/ACs/authors
- Negative: Would slow down the review process/there's not enough time
- Negative: Not significantly better than a single response
- Negative: Is likely as ineffective as/more ineffective than regular author response
- Negative: May hurt diversity or overly favor those of certain advantaged groups
- Negative: May result in arguments that would be contentious and stressful
- Negative: Will increase the variance of the reviewing process
- Negative: The paper should stand on its own without discussion

Meta-review

- Positive: Summary is very helpful to authors/PCs when reviews are conflicting (e.g. w/ the decision)
- Positive: Helps to clearly and concisely convey reasoning behind the decisions

Positive: Prefer for all cases

- Positive: Doing it for all papers ensures that the ACs will look at/check all the reviews

Positive: Not a significant amount of effort for ACs

Positive: Important for borderline and low-scoring papers

- Positive: Important for borderline and high-scoring papers
- Positive: Meta-reviews should be shared with authors
- Positive: Makes it clear that the decision lies in the ACs
- Positive: Important for catching reviewer mistakes
- Positive: Necessary for the PCs
- Positive: Would prefer for message from PCs also
- Positive: It is nice to have the (more) expert opinion of the ACs
- Neutral: Only necessary in borderline cases
- Neutral: Borderline is unclear
- Neutral: Seems like a lot of work, but no personal experience
- Neutral: Have particular comments on recommended format
- Neutral: ACs need to try harder to write good meta-reviews
- Neutral: Can re-use some discussion points
- Neutral: Need to make clear the target of the meta-reviews
- Neutral: Should not be public
- Neutral: Not feasible when handling a large number of papers
- Neutral: Prefer removing the role of ACs
- Negative: More work for ACs and little benefit

- Positive: Public reviews may help prevent authors from making low-quality submissions and wasting reviewer time
- Positive: Public reviews will encourage reviewers to not do poor/uncivil reviews
- Positive: Releasing meta-review/discussion to authors would be helpful
- Positive: Releasing meta-reviews publicly is a good balance
- Positive: Releasing reviews to authors is important to make papers better
- Positive: Public reviews provide a learning experience for inexperienced reviewers/authors
- Positive: Public discussion, with an option for private would be good
- Neutral: Public reviews are only relevant to the submitted version, not the final version
- Neutral: Reviews should remain visible after the conference
- Neutral: It should be possible to opt-in or opt-out
- Neutral: Even if reviews are released many people won't care
- Neutral: The amount of transparency needs to be made clear up front
- Neutral: Public release should be on a case-by-case basis
- Neutral: Discussion among reviewers is important
- Negative: Public release would make it harder to have frank discussion
- Negative: Public release would be hard on reviewers/make it harder to recruit reviewers
- Negative: Releasing meta-reviews/discussion may not be good because of lack of anonymity
- Negative: Reviewers may become unintentionally de-blinded through public release

Higher Acceptance Rates

- Positive: Some separate method of quality control (e.g. Journals) and higher conference acceptance rates would be better
- Positive: Having different levels of selectivity, even within the same conference, may be useful
- Positive: Higher acceptance rates are better for the majority of people, other than those going for tenure/positions at competitive universities
- Positive: Physical limitations don't seem to be a strong argument
- Positive: Low acceptance rates lead to inefficiency
- Positive: Low acceptance rates lead to variance
- Positive: Low acceptance rates lead to rejecting good papers
- Positive: Should have similar acceptance rates for short/long papers
- Positive: Low acceptance rates are harmful for novel papers/fields
- Positive: Important to prevent others from flag-planting
- Neutral: Two clear accepts should be enough to get in
- Neutral: Current system seems to be working
- Neutral: Current system is better than other models
- Neutral: Clarity about why papers got rejected with high scores or accepted with low scores is important
- Neutral: It may not be possible to simply threshold on quality due to logistics