ACL Reviewing Survey Results

This document contains a series of graphs of responses to part of the ACL reviewing survey. Analysis of the free text responses has been completed separately and can be found at https://www.cis.lmu.de/~hs/acl22/panel/survey.html.

Contents

1 Results for Questions

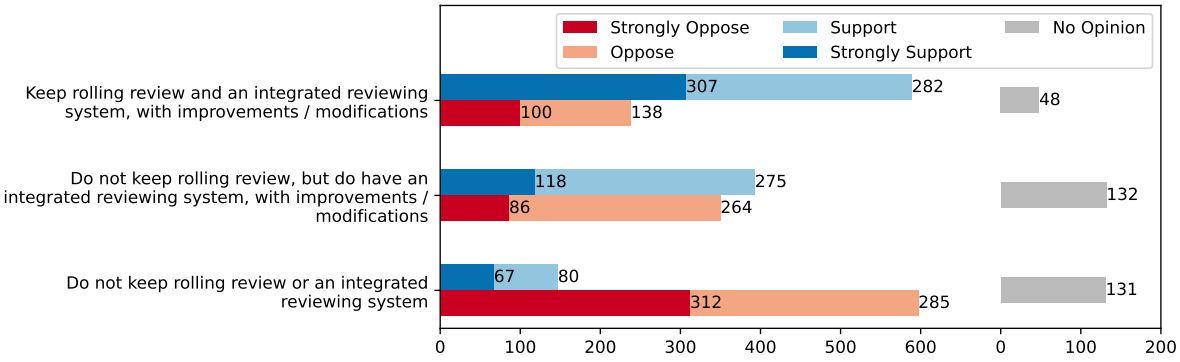
This section presents results of the questions with multiple-choice answers. Every question had to be answered. Five options were given: Strongly Support, Support, Oppose, Strongly Oppose, No Opinion.

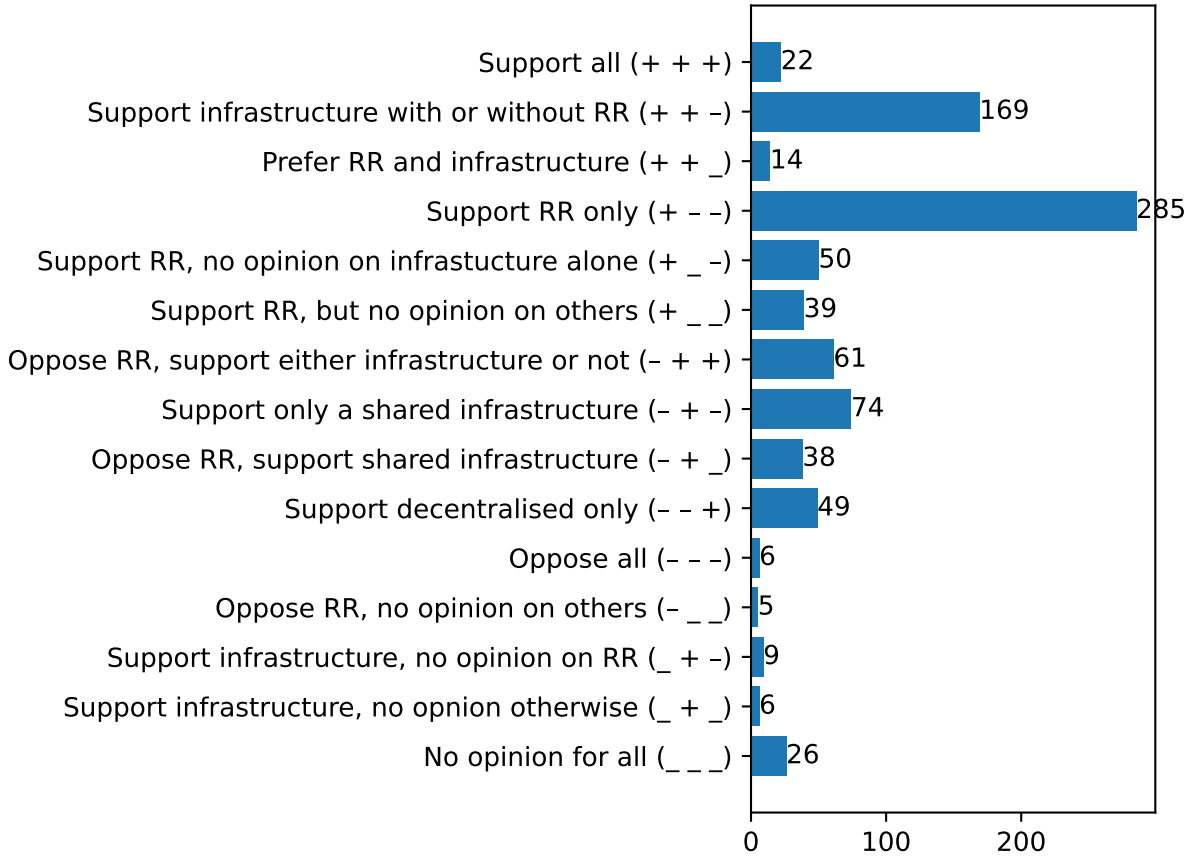
We present the results in two ways. First, the counts of each question, with a horizontal bar chart where blue is support, red is oppose, grey is no opinion, and strongly is indicated by a darker shade.

Second, another visualisation for cases where a set of questions were related (e.g., four options related to tracks). We count how many times each set of responses occurred (e.g., how many people said Support for the first question, and Oppose for the other three). For this visualisation we do not consider the 'Strongly' distinction. In the labels, a plus sign indicates support, a minus sign indicates oppose, and a space or underscore indicates no opinion. The left-to-right pattern of symbols indicates answers to the questions in the first figure top-to-bottom. Patterns given by fewer than five people are not shown.

1.1 Rolling Review and Infrastructure

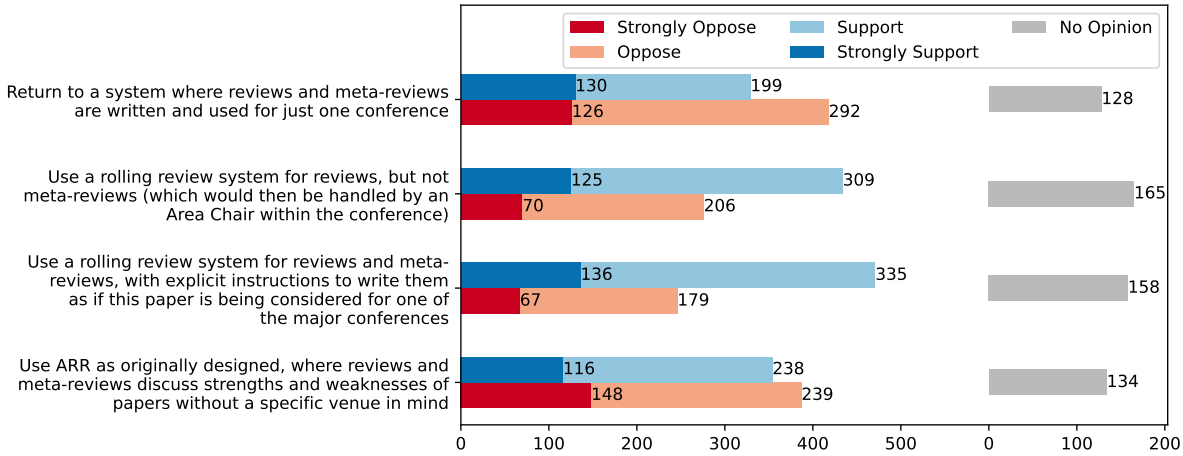
This question asked for respondents to provide their opinion on overall approaches to reviewing. Each respondent provided three answers, one for each statement. Note that there were no constraints on how people responded, for example, they could respond "Strongly Support" to all three. The labels on the left of the second figure (e.g., "Support all") are our interpretation of each answer.

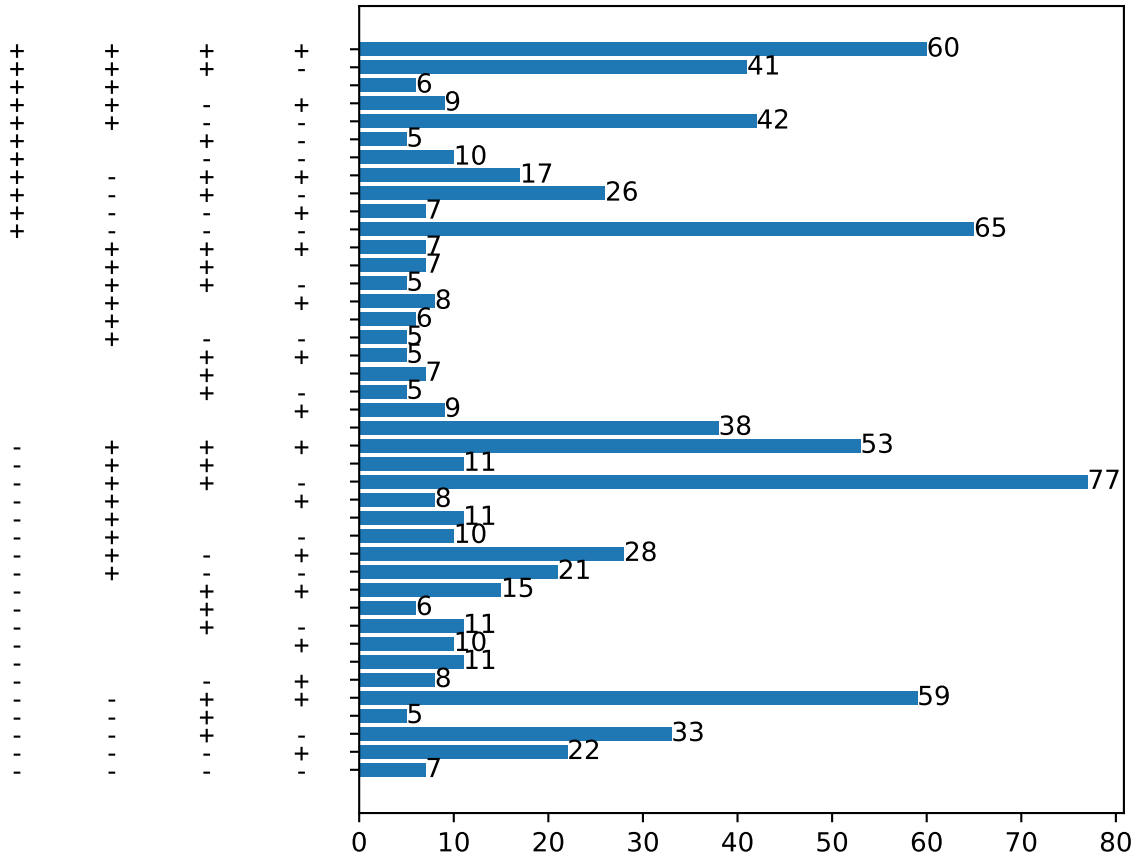




1.2 Review Use

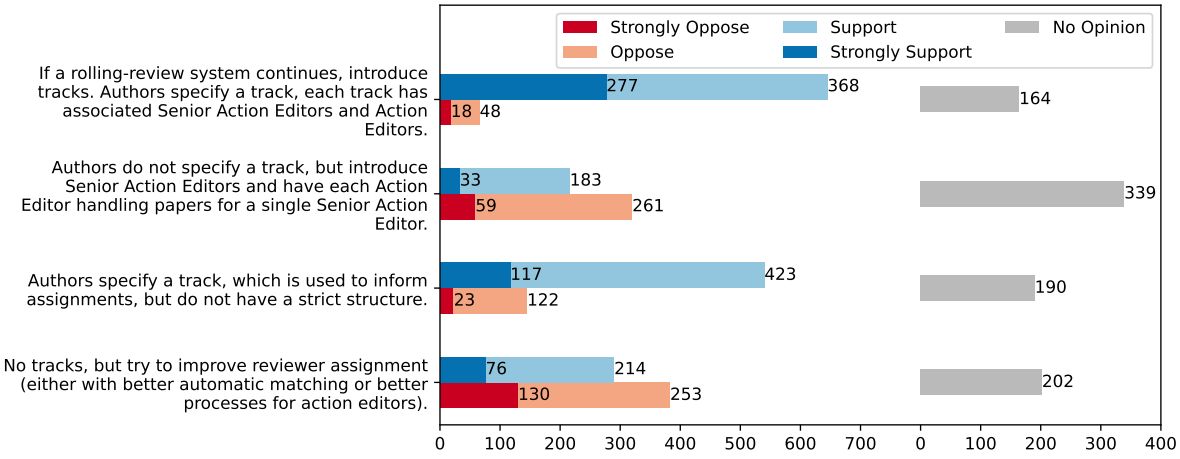
These questions asked for opinions about how reviews should be used. Again, respondents provided four opinions, one per statement.

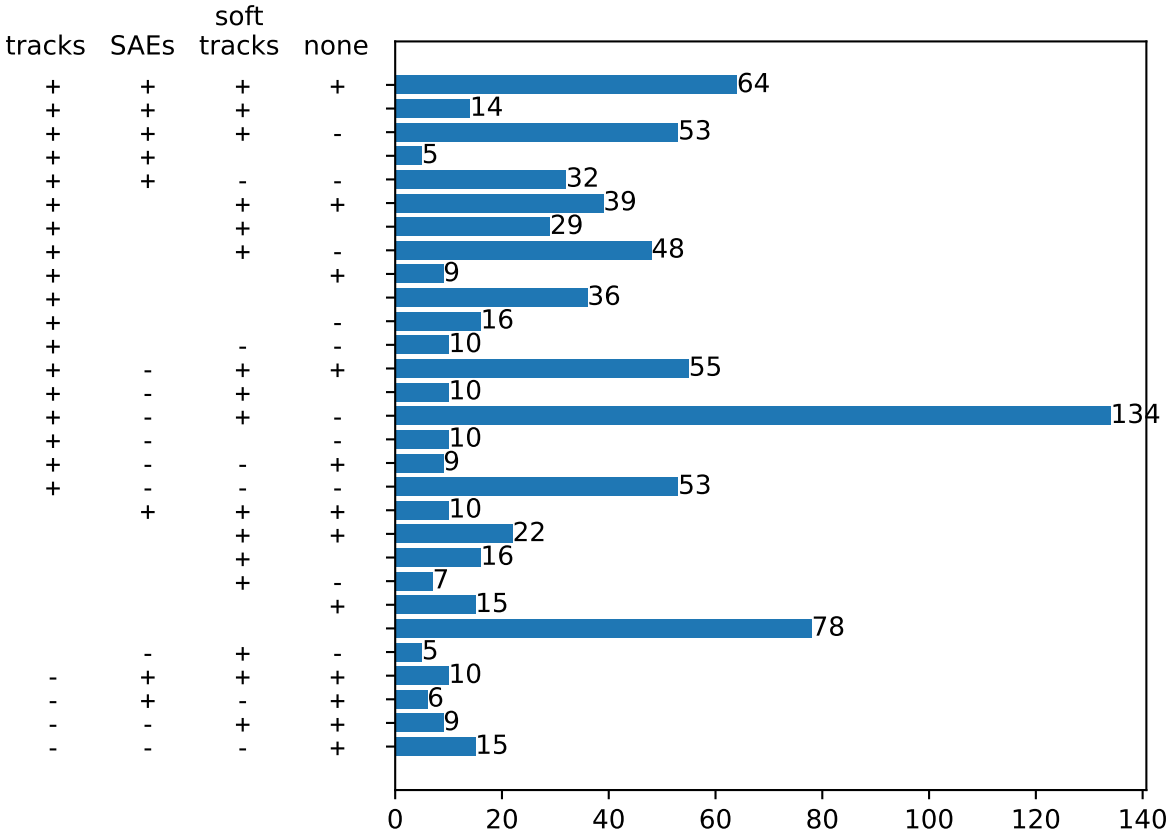




1.3 Tracks

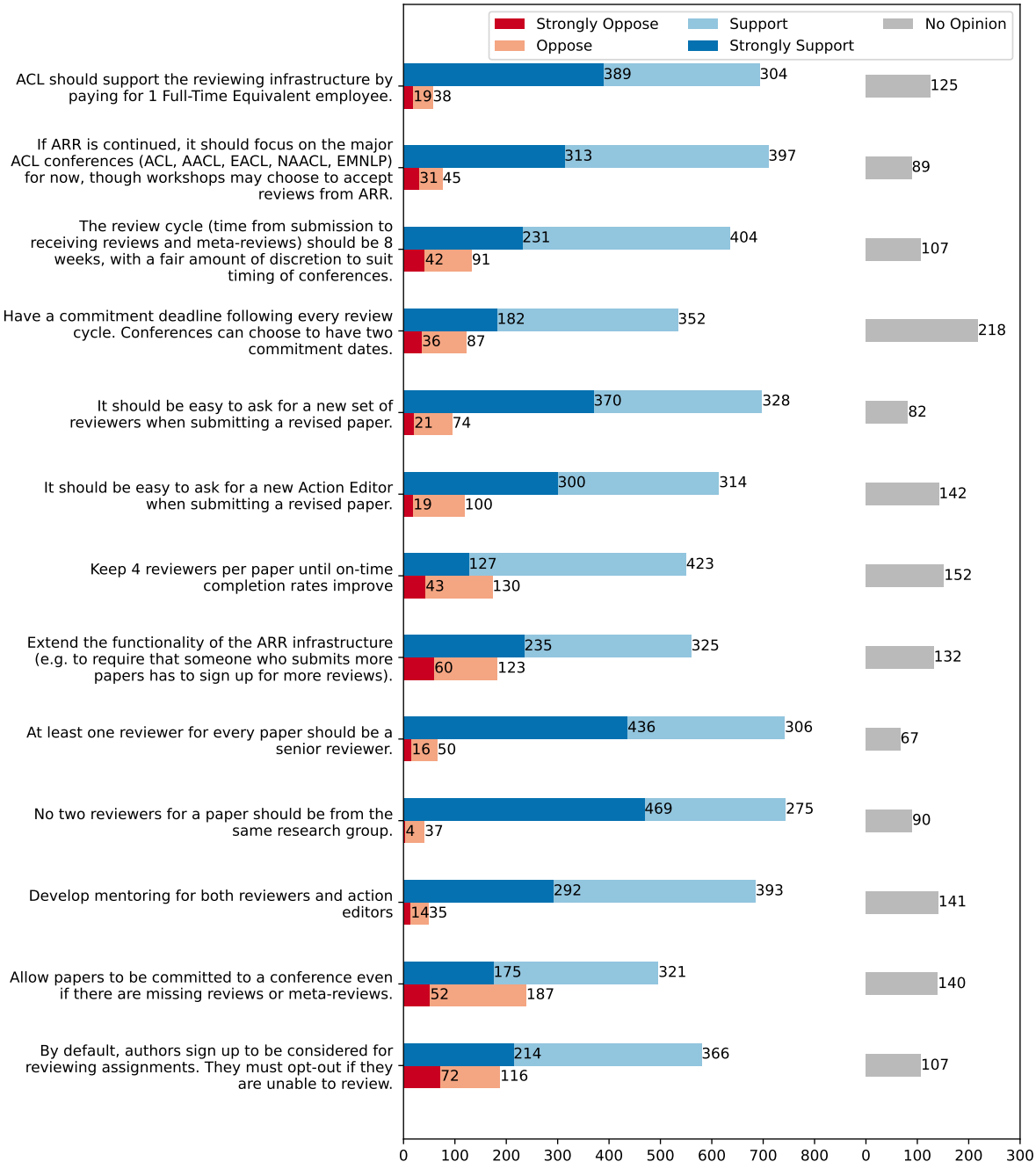
These questions asked for opinions about the use of tracks or related mechanisms. Again, respondents provided four opinions, one per statement, with no constraints on their pattern of answers.

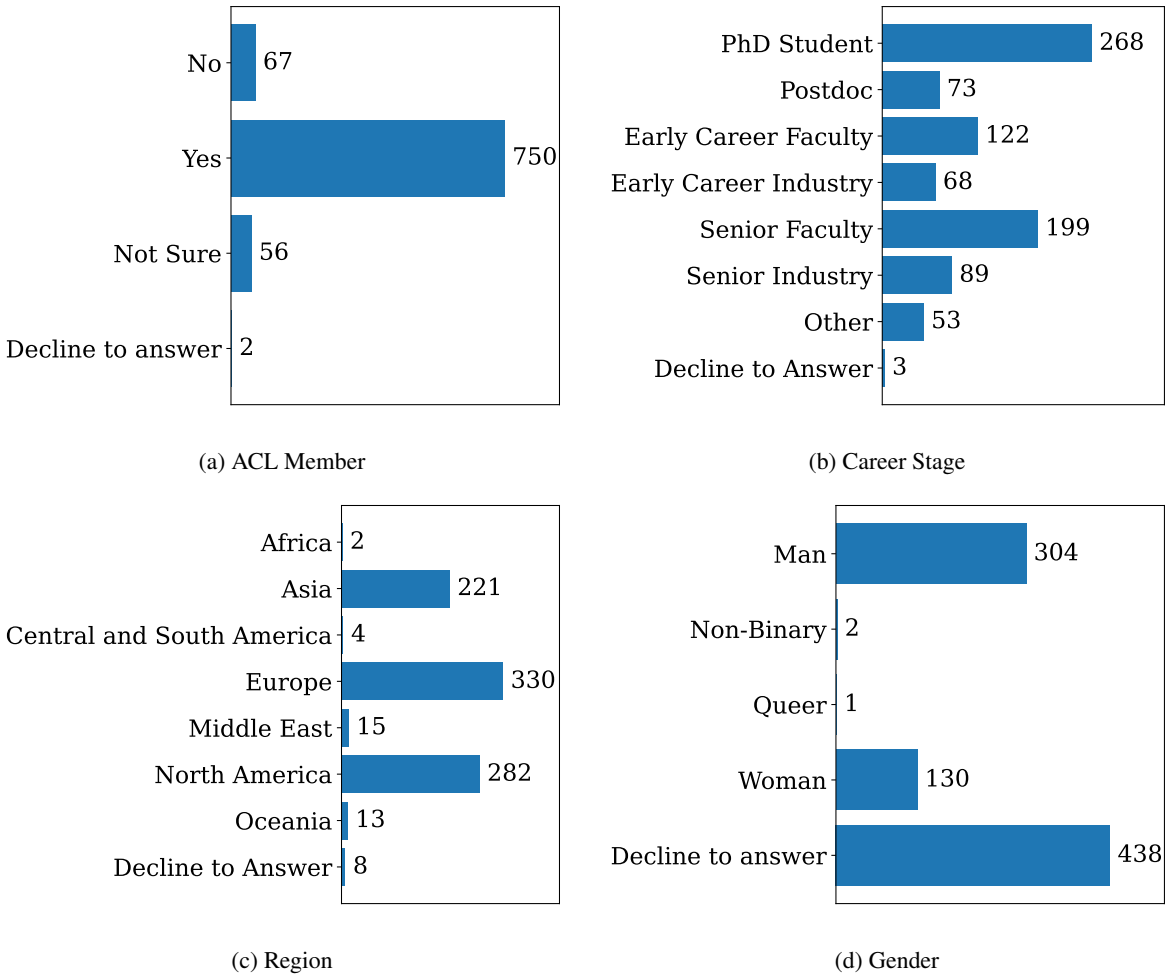




1.4 Other Questions

These are all the other opinion questions in the survey.





(a) ACL Member

(b) Career Stage

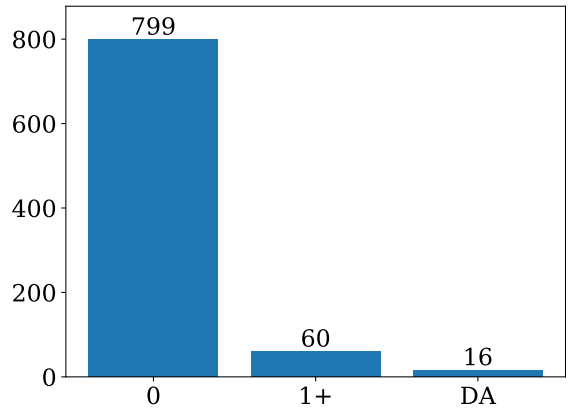
(c) Region

(d) Gender

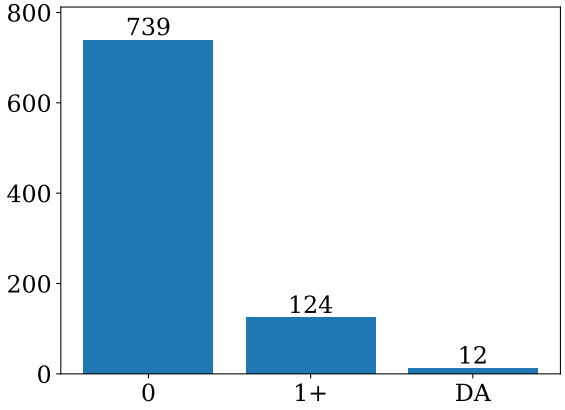
Figure 1: General demographic information about respondents.

2 Demographics

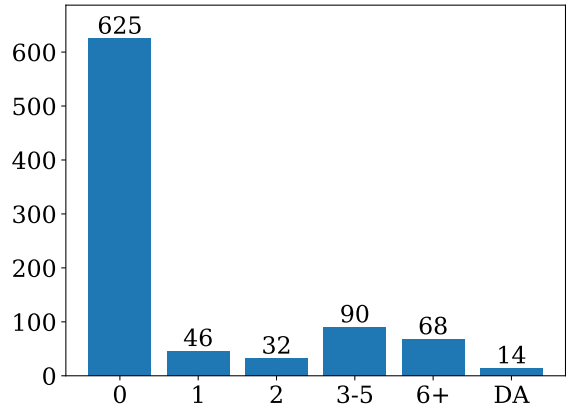
This section provides a summary of responses to the demographics questions.



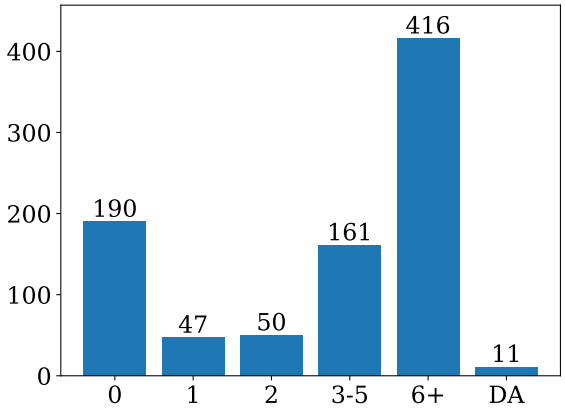
(a) Program Chair



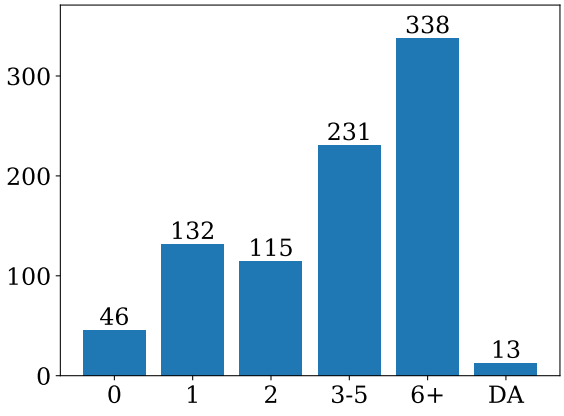
(b) Senior Area Chair



(c) Area Chair

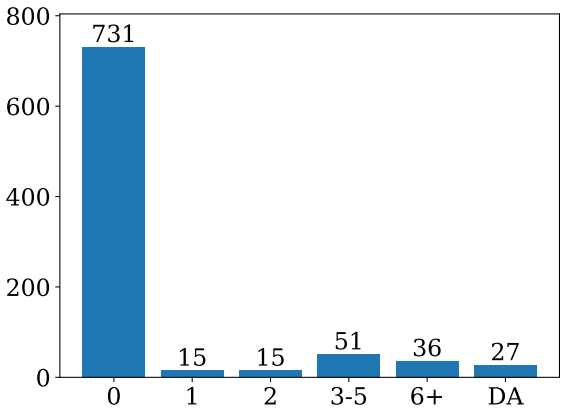


(d) Reviewer

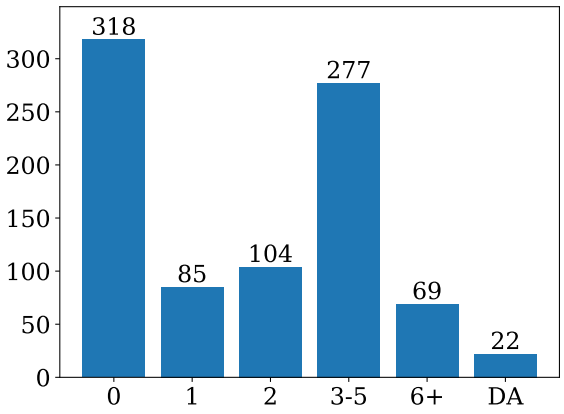


(e) Author

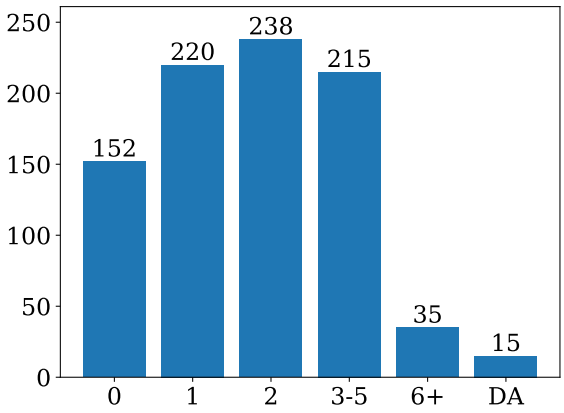
Figure 2: Conference experience.



(a) Action Editor

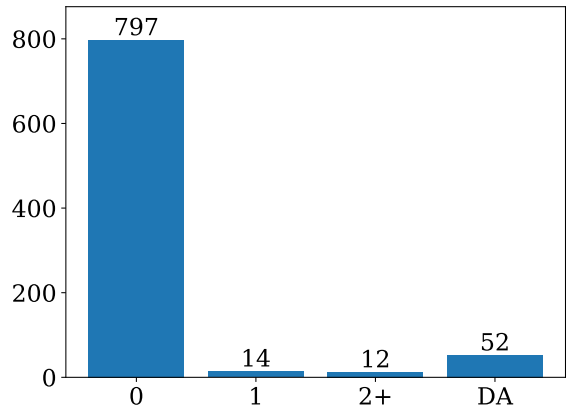


(b) Reviewer

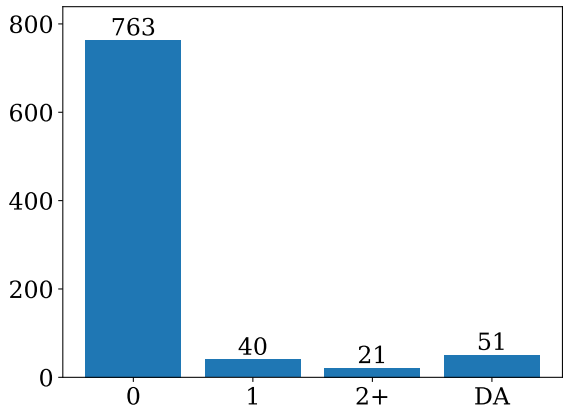


(c) Author

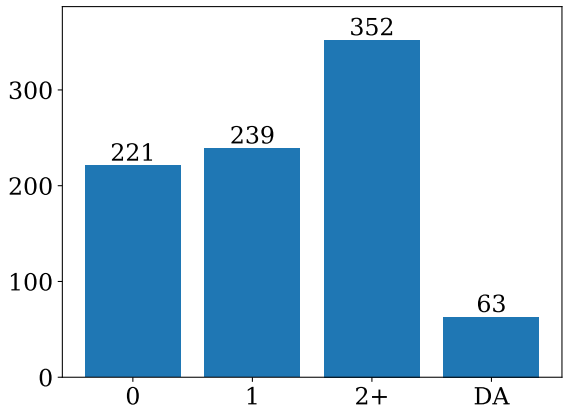
Figure 3: ACL Rolling Review experience.



(a) Program Chair



(b) Senior Area Chair



(c) Author

Figure 4: Commitment process experience.

3 Questions with Substantial Variations in Answer Distributions

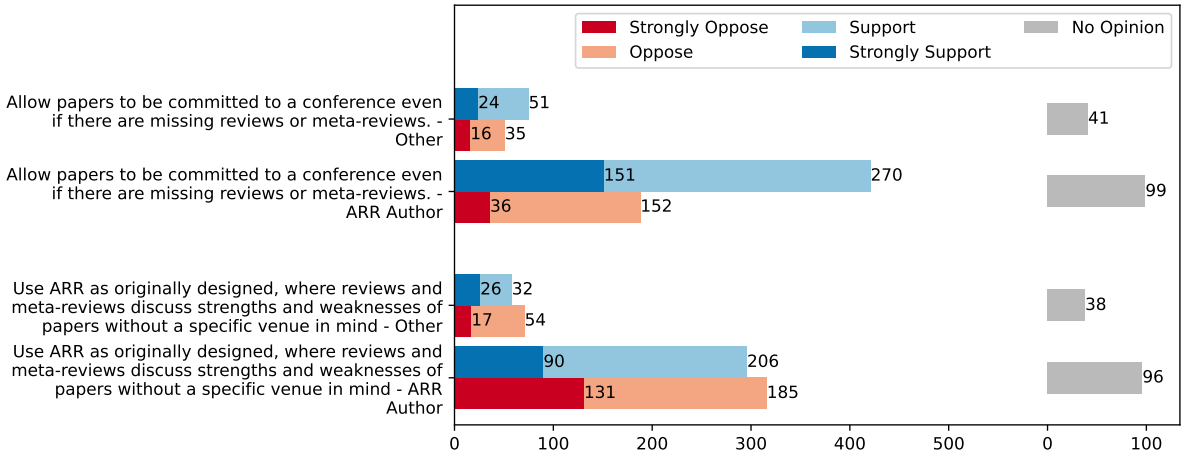
There may be important variations in opinion between different demographic subgroups. To identify these, we considered several ways of splitting the data collected. In each case, we compared responses from the group with the rest of the responses[1]

- Experience as an author for ARR
- Experience as a reviewer for ARR
- Experience as an author and reviewer for ARR
- Experience as an action editor for ARR
- Experience as a reviewer for conferences or ARR
- Experience as an SAC or PC
- Senior, defined as respondents who have reviewed for at least 6 conferences, taken on some sort of senior role (AC, AE, SAC, PC), and are not PhD students or postdocs.
- Late career, either senior faculty or senior industry
- PhD students
- Industry, either early or senior
- Location is North Amercia
- Location is Europe
- Location is Asia
- Men and women

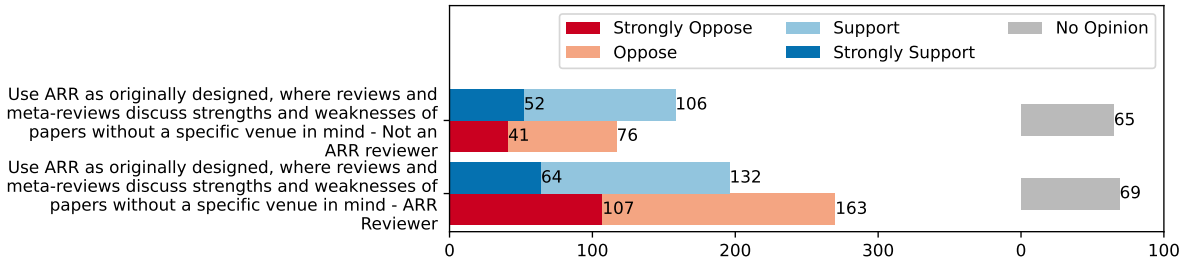
For each comparison, for each question, we measured (1) the l_1 and l_2 distances between the distributions, and (2) the difference in overall support percentage. The top 20 cases for each metric were selected. Many of these overlapped, giving the 33 cases shown below. These 33 questions come from all subgroupings except (a) action editors, and (b) industry.

Each graph below shows the questions from one group comparison. For each question, we show the answer distribution for the group and for the rest of the responses.

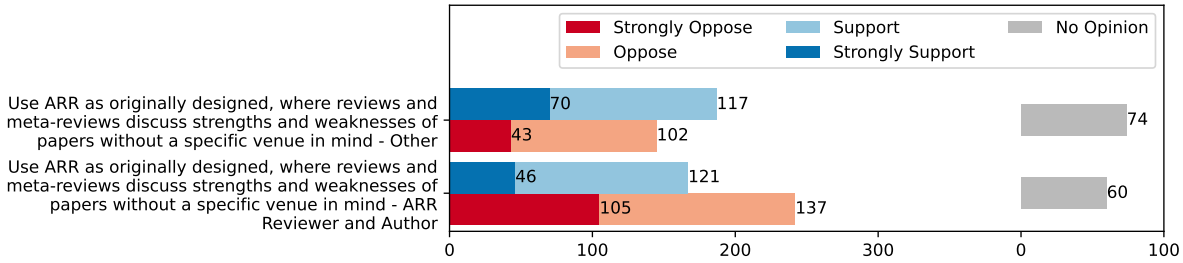
[1] One exception is the gender based comparison. We only compare men with women. This comparison does not consider people who identify in other ways (as there were only three such responses), and people who did not provide a response.



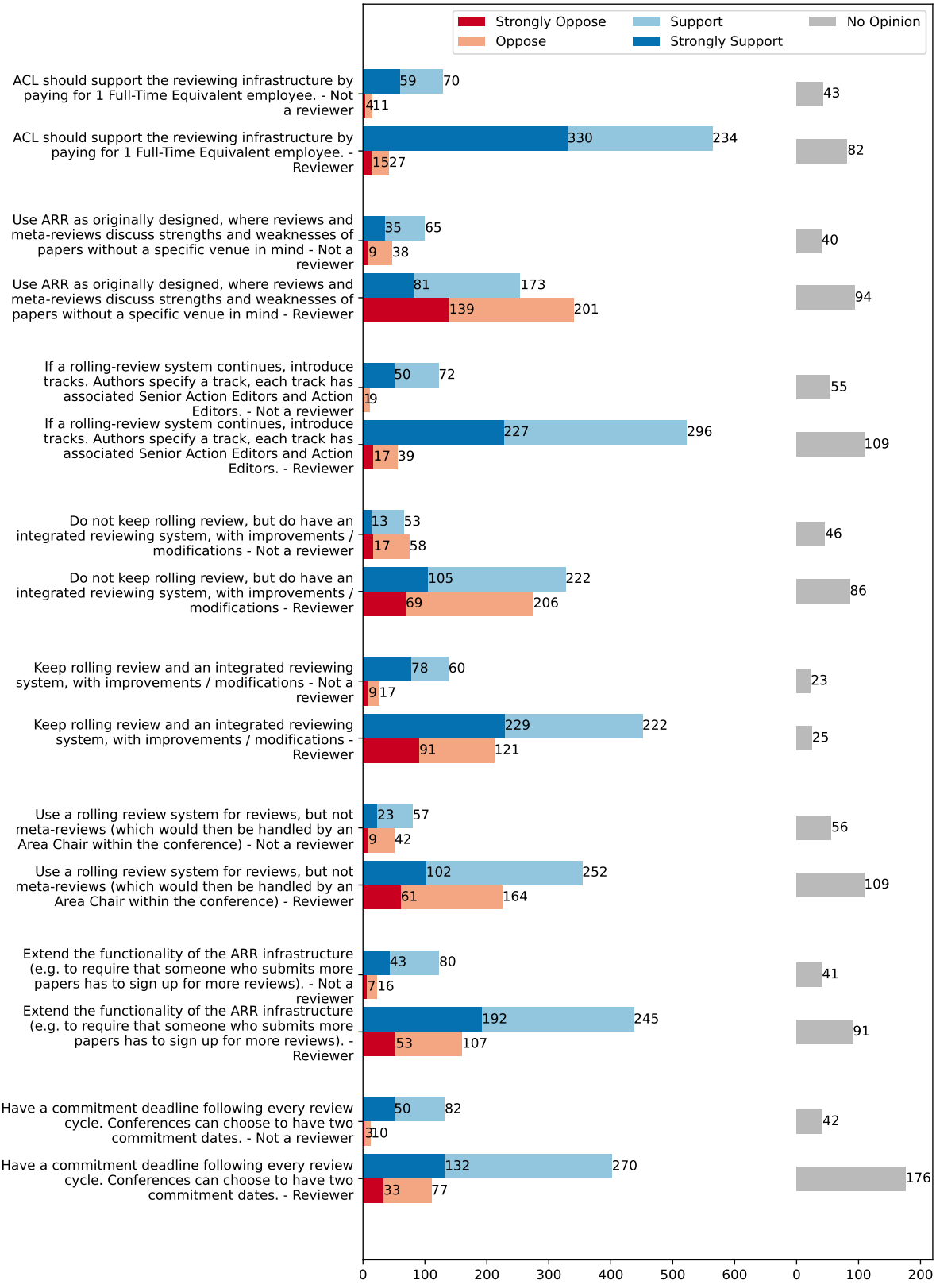
This graph considers respondents who have authored papers submitted through ARR. Their overall support for being able to submit without all reviews or the meta-review is 15% higher than respondents without experience as an author. On whether reviews should be written without a certain venue in mind, there is a shift from no opinion (-9%) and oppose (-6%) to support ($+10\%$) and strongly oppose ($+8\%$).



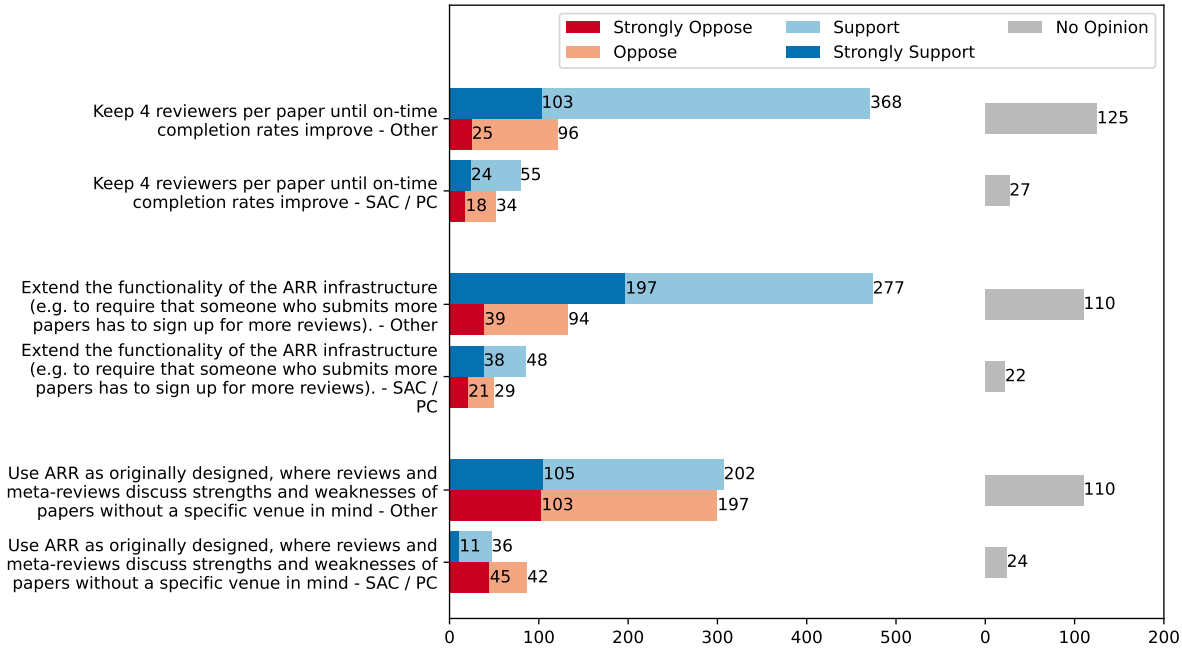
This graph considers respondents who have been a reviewer for ARR. On whether reviews should be written without a certain venue in mind, there is a shift from no opinion (-6%), support (-7%), and strongly support (-3%) to oppose ($+8\%$) and strongly oppose ($+8\%$).



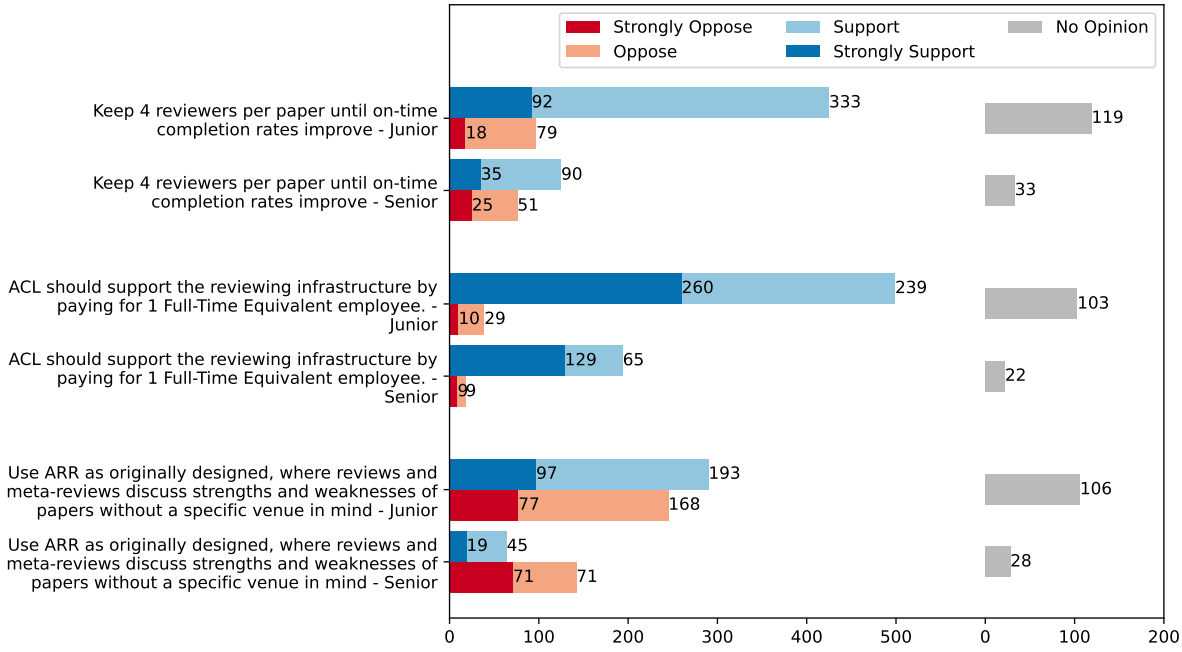
This graph considers respondents who have been an author and reviewer for ARR. The same trend occurs here, a shift from no opinion (-5%), support (-3%), and strongly support (-7%) to oppose ($+4\%$) and strongly oppose ($+12\%$).



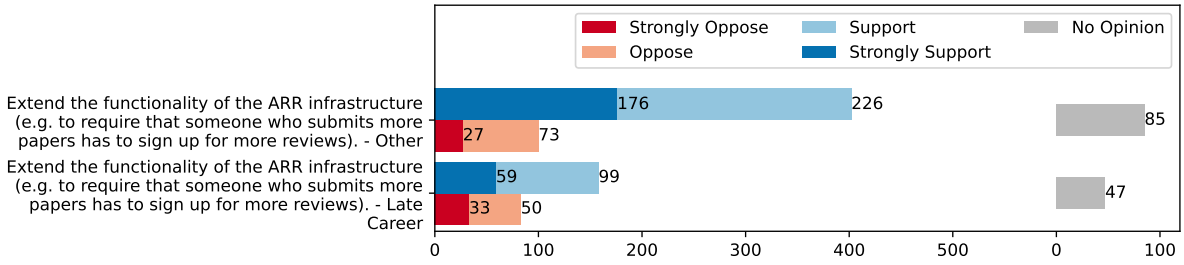
This graph considers respondents who have been a reviewer for either a conference or ARR. Most of these questions had shifts in the level of support, but the same overall trend in the distribution of answers. Exceptions are: (1) reviews should not be venue specific, which had 15% higher strongly oppose and 9% higher oppose, leading to more opposition than support, (2) greater support for an integrated reviewing system (12% higher).



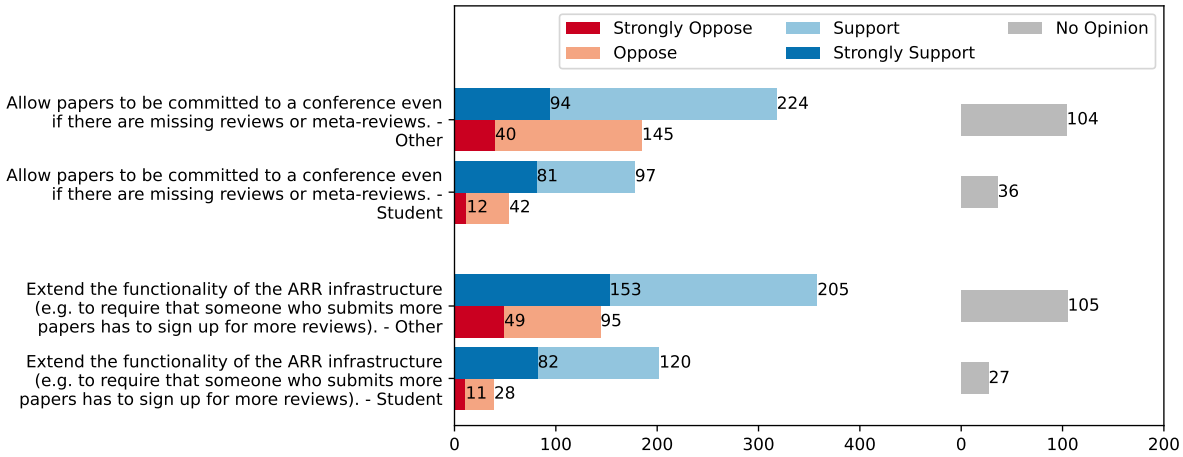
This graph considers respondents who have been a senior area chair or a program chair. In the first two questions, respondents with SAC and PC experience split more evenly between support and opposition. The third shows a trend with much higher opposition (strongly oppose is 14% higher).



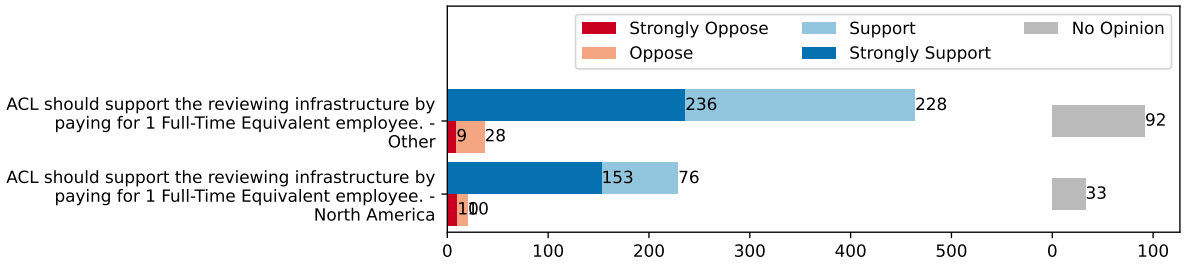
This graph considers respondents who are a senior member of the community: reviewed for at least 6 conferences, taken on some sort of senior role (AC, AE, SAC, PC), and are not PhD students or postdocs. The patterns for the first and third questions here are the same as in the previous figure. The idea of ACL support for ARR has a shift in two directions, with strong support higher by 15% and strong opposition higher by 2%.



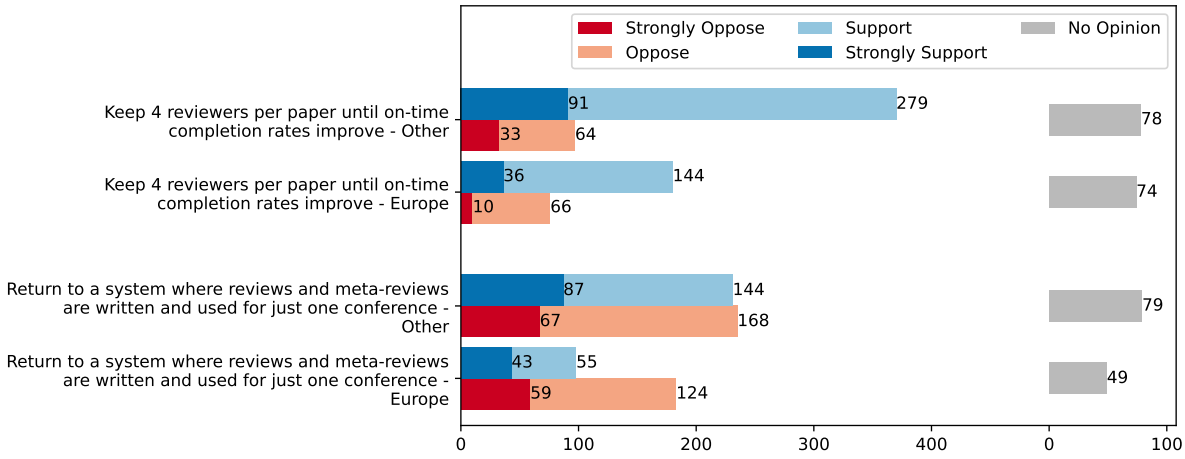
This graph considers respondents who are either senior faculty or senior industry. This group had higher opposition to this question (-14% overall support).



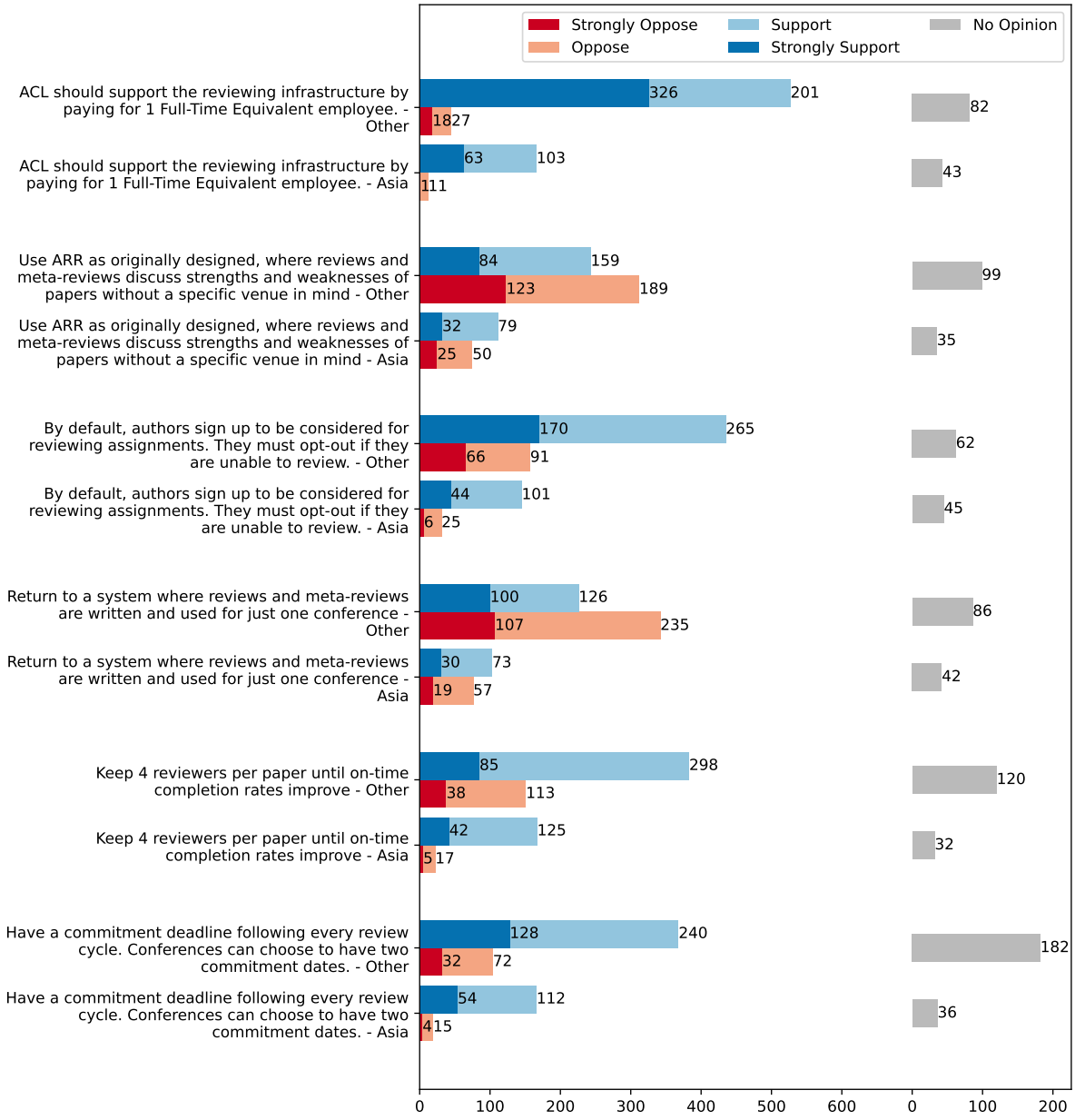
This graph considers respondents who are PhD students. Both questions had higher overall support. In the first, this was almost entirely due to higher strong support.



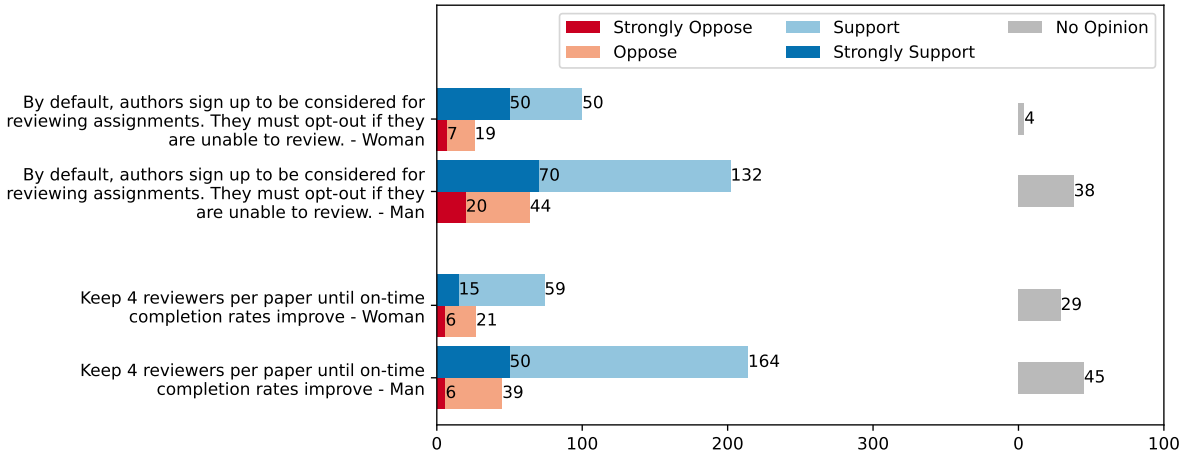
This graph considers respondents whose location is North America. The differences are similar to those in the graph for senior community members.



This graph considers respondents whose location is Europe. Both questions show lower overall support by 13%, which in the second case means there is more opposition than support.



This graph considers respondents whose location is Asia. In most questions, the overall trend in opinions is the same. The exceptions are the second and fourth questions, where the higher overall support (12-13%) is enough to change the majority opinion.



This graph considers respondents who are men and women. Note, this figure does not include all responses (unlike the ones above). It does not consider people who identify in other ways as there were only three such responses. It does not include respondents who chose not to answer the question about gender. The overall opinion is similar in both questions. In the first, women have higher overall support ($+10\%$) and fewer no opinion answers (-9%). In the second, women have lower overall support (-13%).