

Vote Equity Project // Voting and Ranking Methodology

The Vote Equity Project used an innovative process for voting and ranking ideas. This section describes this process and the reasons for its unique design.

The Voting Process

The Ideas

- Ideas were submitted by over 100 community organizations as well as by individual Chicago residents via the Chicago United for Equity (CUE) website. Ideas submitted via CUE's website had to meet their criteria of relevance, feasibility, and potential for impact to be included in the vote.
- 186 ideas were included in the vote. Of these, 107 were categorized by CUE as 'Policy' ideas and 79 were categorized by CUE as 'Structural' ideas.

Voting method

- The method of pairwise comparison was used for the vote. Voters were presented with two randomly selected ideas from the full list of ideas and voted for the idea they felt best responds to the question "*Which idea would best help build a Chicago that works for ALL of us?*". Voters could also select 'I can't decide.' Here is an example from the print ballot:

Which idea would best help build a Chicago that works for ALL of us?

Fully fund neighborhood public schools so they can provide the education and supports students need to thrive.	OR	Invest in expanded and more frequent CTA bus service, including Bus Rapid Transit routes
<input type="radio"/>		<input type="radio"/>
I can't decide		
<input type="radio"/>		

- This process is based on allourideas.org which was developed by Princeton Professor Matt Salganik. The pairwise comparison voting method has many benefits:
 - It allows for the comparison of a large quantity of ideas. This enables a more democratic process where voters collectively prioritize a large list compared to voting on a much shorter list of ideas selected by the organizers of the vote.
 - It resists 'gaming' or vote manipulation because respondents cannot choose which ideas they will evaluate.
 - It resists snowball and order effects that occur because voters are more likely to choose ideas that are presented at the top of a list (e.g. the top ideas).
 - It provides a simple and engaging format. Selecting between two ideas is much easier than ranking a long list. Rating is also harder than pairwise comparison --

for example, asking “how big is this vehicle on a scale from 1-10” is harder and a less reliable foundation for ranking than “which of these two vehicles is bigger?”

- It requires tough choices and prioritization like those faced by policy makers.

The ballots

- Each ballot contained a unique mix of randomly paired ideas. This means that no two people voted on the same set of paired ideas.
- ‘Policy’ ideas (about what our government should do) were only paired with other ‘policy’ ideas and ‘structural’ ideas (about how government should function) were only paired with other ‘structural’ ideas.
- Ballots were available both online and in print, in English and in Spanish. Print ballots included 20 pairwise comparisons and extra ballots were provided for people who wanted to vote more than 20 times. Online voters were able submit only one ballot and vote up to 100 times. Vote limits were used in order to balance engagement with disproportionate influence. A ‘trap’ question was used online to catch fraudulent votes.
- Voting was voluntary and limited to residents of Chicago.

Voters and Votes Summary

- **A total of 2,126 ballots were submitted**
 - Ballots were collected from mid-December through the end of January.
 - Voters who provided a non-Chicago ZIP code were excluded from the results.
- **Language**
 - English ballots submitted = 2,010
 - Spanish ballots submitted = 116
- **Print vs. Online**
 - Online ballots submitted = 1,442
 - Print ballots submitted = 684
- **Number of votes**
 - All = 52,271
 - On policy ideas = 26,193
 - Each of the 107 policy ideas were voted on an average of 490 times
 - On structural ideas = 26,078
 - Each of the 79 structure ideas were voted on an average of 660 times
- **On average each voter cast 25 votes**

Ranking Methodology

How the ideas are ranked

Each idea ended the voting period with a record of wins and loses. We calculate each idea’s final score based on the formula used by allourideas.org, which resembles a simple winning percentage: $(\text{wins} + 1) / ((\text{wins} + 1) + (\text{losses} + 1))$.¹ This score represents “the probability that

¹ Salganik MJ, Levy KEC (2015) Wiki Surveys: Open and Quantifiable Social Data Collection. PLoS ONE 10(5): e0123483, Appendix 1, p. 2. <https://doi.org/10.1371/journal.pone.0123483>. Note that the +1 term in the numerator (α) and the +1 terms in the denominator (α , β) are the beta distribution parameters. This means that the expected win probability of a new idea is $(1 / (1+1)) = .5$, which is the best guess because the mean win probability for all ideas will always be .5.

[the idea] will beat a randomly chosen [idea] for a randomly chosen respondent.”² For example, the top ranked idea has a score of 80 which means that we expect this idea to win 80 times out of 100. Based on their final score, ideas are then ranked from highest (the idea with the highest score) to lowest (the idea with the lowest score).

We chose to calculate the score using this formula because it is used by allourideas.org and to be as transparent as possible (it is a simple calculation that is easy to replicate). This calculation does not take into account the ‘strength of schedule’ of each idea, so all wins and losses are treated equally regardless of the strength of the opponent. However, we compared our results to results using the Glicko rating method that takes ‘strength of schedule’ into account and found a very high level of correlation (>.99) between our results and the Glicko results.

Accuracy of the results

The more times that you flip a coin, the closer the percentage of heads over all flips will be to its true probability of .5. Just like flipping a coin, the more times an idea takes part in a pairwise comparison, the closer its score will be to its true score. When looking at the full set of 186 ideas and based on all votes the average credible interval is $\pm 4.2\%$.³ This means that for this set of ideas and this group of voters there is a 95% probability that the true scores are within an average of $\pm 4.2\%$ of the observed scores.⁴

Representativeness of the results

Participants were a self-selected sample, not a randomly selected sample of Chicago residents. Accordingly, results cannot be statistically generalized beyond the group of participants. It is accurate to say, “Chicagoans who voted in the Vote Equity project prioritize...”. While it’s possible that these results reflect the priorities of Chicago residents more generally, we cannot know this with any measure of statistical certainty because the sample is non-random. However, these results do represent the collective effort of over 2,000 engaged Chicagoans to prioritize 186 ideas and reforms to increase racial equity in the city.

If you have questions about the voting and ranking methodology, please contact Matt Sweeney at mesweene@uic.edu

² Salganik MJ, Levy KEC (2015) Wiki Surveys: Open and Quantifiable Social Data Collection. PLoS ONE 10(5): e0123483, p. 5. <https://doi.org/10.1371/journal.pone.0123483>

³ Since each pairwise comparison included two randomly chosen ideas, the total number of times that each idea was included in a pairwise comparison is not the same. Due to this randomness, some ideas were presented more than others, and accordingly, the scores for ideas that took part in more pairwise comparisons have a smaller credible interval.

⁴ Note that since policy ideas only battled policy ideas (107 ideas) and structural ideas only battled structural ideas, the average credible intervals are different because each of the 79 structural ideas battled more times on average. The average credible interval for policy ideas is 4.5% and for structural ideas its 3.9%.