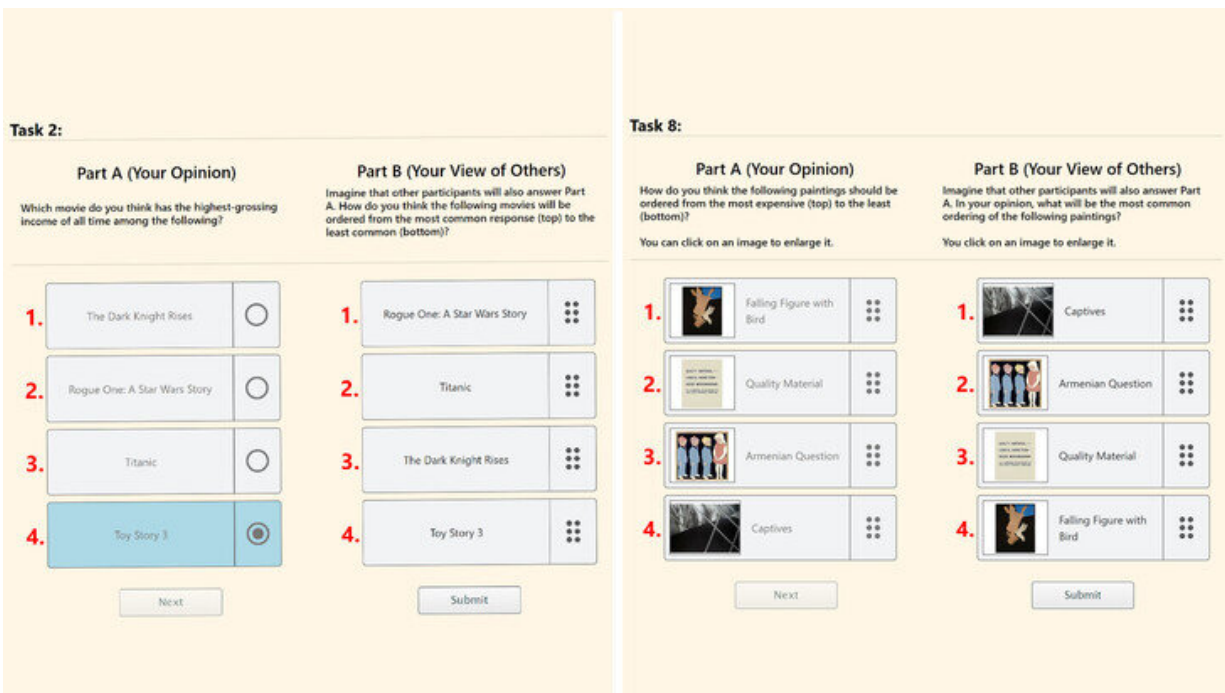# Surprisingly popular voting algorithm developed to recover ranked choices

August 26 2021, by Jessica Hallman

A view of the interface that Hosseini and his team used to test their model, asking crowdsourced workers questions about geography, movies and artwork to choose, for example, which movie from a selection of four had the highest-grossing box office income of all time. The researchers used six different formats to elicit responses from participants, such as selecting the top answer from a list of four choices (as shown in Task 2) or ranking those four choices in order (as shown in Task 8). Some participants were also asked to predict which top answer others would choose or how others would rank their order, as shown in both examples. Credit: Pennsylvania State University

Imagine you are asked to rank the counties in Pennsylvania in terms of number of COVID-19 infections. Or you may be asked to rank the following cities in Pennsylvania based on their populations: Harrisburg, Allentown, Erie and State College.

What would your answer be? How do you think others would answer these questions?

A new algorithm developed by a team of researchers led by Hadi Hosseini, assistant professor at the Penn State College of Information Sciences and Technology, can reach the true answer to this and similar questions by combining a respondent's own vote or opinion with their prediction about how others will answer.

Hosseini explained that a person knowledgeable of Pennsylvania would most likely know the answer to the above question. They may also predict that other less-informed participants, on average, would provide the incorrect ranking. Conversely, uniformed participants are most likely not aware of the correct answer and may give a wrong ranking of these cities.

"This is the crux of our algorithm: using the extra information that informed participants have to correct for such errors," he said.

The researchers' method expands on an existing approach of soliciting wisdom of a crowd, called surprisingly popular method, which has been used in scenarios like political pollsters predicting the outcome of elections and researchers predicting the winners of NFL games. Like Hosseini's model, surprisingly popular method asks respondents to provide two answers per question: what their own opinion or vote is, and how they predict that others will vote. The technique takes advantage of the knowledge of a small group of experts in a larger crowd to point to the correct answer.

However, the surprisingly popular method has been limited to predicting a single correct answer to a posed question, such as "What is the capital of Pennsylvania?" or "How much money did the movie "Titanic' make at the box office worldwide?" Hosseini's model extends this concept to ranked choices or alternatives.

"One does not require to fully elicit complete rankings and predictions of others," said Hosseini. "We are able to recover the ground truth by combining both the vote and predictions without eliciting full distributions over all possible n! rankings. And this is true for recovering either only the top choice or the full ranking."

The method could potentially be applied to improve forecasting with ranked choices, such as in exit polls for predicting outcomes of political elections. According to Hosseini, by asking voters secondary questions through his method, fewer samples would be required compared to standard exit polls that rely on random sampling.

Hosseini's approach also significantly outperforms conventional voting methods, such as the simple majority rule, that do not ask respondents to predict how others will respond.

"It turns out that prediction of others' votes is more important than the actual votes," Hosseini said. "This is very crucial because it posits that asking what you think about other people's opinion is a more critical question than asking their own opinion."

To test their model, Hosseini and his team asked 720 crowdsourced workers questions about geography, movies and artwork to choose, for example, which movie from a selection of four has the highest-grossing, box-office income of all time. The researchers used six different formats to elicit responses from participants, such as selecting the top answer from a list of four choices or ranking those four choices in order. Some

participants were also asked to predict which top <u>answer</u> others would choose or how others would rank their order.

"Our algorithmic methods and findings can have a significant impact on how we see and deal with elections, whether on the national or local levels," said Hosseini. "More importantly, our technique shows that we can predict a ground-truth ranking with high accuracy without the need for massive data collection."

Additionally, he said, the method could be applied in areas outside of forecasting the objective outcomes of events such as political elections and sports games.

"Surprisingly, this technique works even for predicting the price ranking of abstract paintings, which typically is more challenging to speculate by the crowd," he added.

They presented their paper, "Surprisingly Popular Voting Recovers Rankings, Surprisingly!" this week at the International Joint Conference on Artificial Intelligence (IJCAI-21), held virtually Aug. 19-26. The work was partially funded by the National Science Foundation.

More information: Surprisingly Popular Voting Recovers Rankings, Surprisingly! <u>faculty.ist.psu.edu/hadi/papers/spvoting.pdf</u>

Provided by Pennsylvania State University