GroupReasoning

Embargoed registration

Metadata

Preregistration Template from AsPredicted.org

Data collection
Have any data been collected for this study already? Note: 'Yes' is a discouraged answer for this preregistration form.

No, no data have been collected for this study yet.

Hypothesis

Two dimensions (among others) may incline an agent to accept a proposed belief: (A) the proportion of respondents independently supporting it, or (B) the manner in which respondents arrived at their decision. When these dimensions produce conflicting intuitions, agents must decide whose testimony is most reliable on other grounds. One way to do this is to consider the kind of question being asked. Discussion may be particularly beneficial in complex problem solving and rational inference, despite the fact that discussion sacrifices independence of testimony, an important criterion in determining the value of majority support.

Prior results suggest that even children's preference for group discussion is specific to "reasoning" questions. Here we test the effect of a question's difficulty by contrasting 4 pre-tested "easy" reasoning questions with 4 difficult perceptual judgment questions. If participants' preference for group discussion in earlier studies was due to the difficulty of the question, then they will prefer group discussion for the difficult perceptual judgment questions over the easy reasoning questions. If their preference was due to an intuition that group deliberation positively impacts reasoning performance, they will prefer group discussion for the easy reasoning questions over the difficult perceptual judgement questions.

H1: Adults will prefer group discussion answers for questions that require reasoning or problem solving, and independent "crowd poll" answers for perceptual judgement questions.

H2: By 9-10, children will understand the difference between question domains and recognize the value of discussion for complex problem solving. If this is the case, then children's responses will be qualitatively similar to adults. We predict that the older children will show the adult pattern; however,

younger children (7-8) may show a general preference for group reasoning for both question types.

Dependent variable

In a within-subjects design, each participant will be told that "Jack" (depicted as a silhouette) has some questions, and that there are many people (depicted as a large crowd) available to help him. They will be told that Jack can ask for help in two ways — through "Talking Together" (5 person group discussion), or by "Answering Alone" (independent crowd poll of 50 people). Each participant will see Jack's 8 questions: 4 questions that require reasoning or problem solving, 4 perceptual judgement questions.

For each of Jack's questions, each participant will be asked to decide whether a group discussion of 5 people or a crowd poll of 50 people will be more helpful for Jack to answer the question.

Responses will recorded on a 4-point "probably/definitely" scale from definitely answering alone to definitely talking together

Adults will use the scale directly. Children's answers will be staggered (binary preference followed by definitely/probably degree of preference).

The DV will be the mean of the ratings for the 4 reasoning questions and the 4 perceptual judgement questions.

Conditions
How many and which conditions will participants be assigned to?

We will use a within-subjects design: each subject will see all 8 questions, 4 of each type (reasoning, perceptual judgement).

Children ages 7-8 (40)
Children age 9-10 (40)
Adults (40)

Analyses

The mean rating by question type will be computed for each participant.

We will conduct a repeated measures ANOVA, using AgeGroup and QuestionType as predictors of the mean rating for each kind of question. One sample t-tests will be conducted to compare mean ratings for each domain to chance for each age group. Multiple comparisons will be conducted to test the difference in mean ratings by domain for each age group.

Outliers and Exclusions

Child participants will be excluded and replaced in case of connectivity issues with the platform, such as extreme lag, poor audio-connection, etc. Children will also be excluded and replaced if parents interfere.

Adult participants on MTurk will be required to pass an attention check after reading the initial instructions. Participants who fail the attention check twice will be screened out of the study.

Sample Size

Our prior studies suggested that n=40 per age group showed large effects of question type. Based on

piloting results, we expected the main effect of QuestionType to be be large, while an AgeGroup*QuestionType interaction for the youngest group would be moderate if observed. Cohen

(1992) suggested .1, .25, and .40 as small, moderate, and large effect sizes.

```
library(WebPower)
cohens_f<-.3
#### Power to detect f=.3 interaction at n=120: 83.6%
wp.rmanova(n = 120, ng = 3, nm = 2, f = .3, nscor = 1, alpha = .05, power = , type = 2)
#power=0.8358527
#### Main effect detectable at 80% power with n=120:  f=.258
wp.rmanova(n =  120, ng = 3, nm = 2, f =  , nscor = 1, alpha = .05, power =.8 , type = 1)#0.2578694
```

Other

Responses to each question will also be analyzed separately for item-level differences, but only as an exploratory measure and if necessary.
For converging evidence, we will also compare responses in this study to responses in our previous study.

Name

Groups vs Crowds: Difficult Perceptual judgment

Finally

Experiment

Other

No response