

‘The crowd within’ effect

Dr Wolf Vanpaemel, Professor of Psychology, KU Leuven, Belgium

The Dataset

Vul and Pashler (2008) conducted an elegant experiment in which participants were asked to answer eight general knowledge questions. Immediately upon completion, participants were unexpectedly asked to provide a second, distinct answer to each question. The results demonstrated that the mean of the two responses from a single participant was more accurate than either response taken alone. Furthermore, a second group of participants was instructed to provide the second answer three weeks after the first session, rather than immediately afterwards. The improvement in accuracy from averaging the two responses was found to be even greater in this delayed condition compared to the immediate condition.

Steegeen, Dewitte, Tuerlinckx, and Vanpaemel (2014) sought to replicate this phenomenon, known as the *crowd within effect*. The current set of exercises is based on their pre-registered, high-powered, replication study, with a focus on the immediate condition only, in which the second guess was collected immediately following the first.

ACTIVITY 1: Loading the dataset

To begin, load the file `rawdata_immediate.txt` into R.

ACTIVITY 2: Filtering based on self-reported cheating

Steegeen et al. (2014) asked participants to indicate whether they had looked up any answers and excluded participants who reported doing so. Retain only those who answered “Nee” (Dutch for “No”) on the `cheated` variable.

ACTIVITY 3: Verifying the filtered sample size

How many participants remain after applying the exclusion criterion?

ACTIVITY 4: Checking the variables

There seem to be, for each participant, 54 variables. Which ones?

ACTIVITY 5: Extracting the answers from each round

The dataset contains several variables we are not interested in. We are only interested in the two sets of responses to the eight trivia questions each participant provided. They are contained in the variables starting with `q` and ending with `a` (for the first guess) or `b` (for the

second guess), so that, for example, q6a refers to the first answer to the sixth question and q3b refers to the second answer to the third question. Extract the relevant columns into separate data frames, one for the first guess and another one for the second guess.

ACTIVITY 6: Building the correct answers matrix

To evaluate the accuracy of the responses, we will use the Mean Squared Error (MSE). To know the error, we of course need to know the correct values. They are reported in Table 3 of Steegen et al. (2014), taken from The World Factbook (Central Intelligence Agency, 2013), and given below:

```
correct_vec <- c(
  6.3, # The area of the USA is what percent of the area of the Pacific
  Ocean?
  43.3, # What percent of the world's population lives in either China,
  India, or the European Union?
  32.3, # What percent of the world's airports are in the United States?
  13.4, # What percent of the world's roads are in India?
  53.6, # What percent of the world's countries have a higher fertility rate
  than the United states?
  54.8, # What percent of the world's telephone lines are in China, USA, or
  the European Union?
  26.4, # Saudi Arabia consumes what percentage of the oil it produces?
  22.4 # What percentage of the world's countries have a higher life
  expectancy than the United States?
)
```

To streamline the computation of the MSE, create a matrix that repeats the correct answer vector `correct_vec` for every participant.

ACTIVITY 7: Computing the MSE for each response set

Use the `correct_mat` matrix to compute the MSE for the first and second guesses separately.

ACTIVITY 8: Averaging the two guesses and computing the MSE

To investigate the crowd within effect, we also need, for each participant, the average answer. Compute the mean of the two guesses for each question and participant. Next, compute, for each participant, the associated MSE of the average response.

ACTIVITY 9: Storing the MSEs for further analysis

Add the computed MSE values to the cleaned dataset.

ACTIVITY 10: Visualizing the results

To provide a visual comparison, plot the average MSEs in a barplot along with error bars representing standard errors. Hint: use `barplot()` and `arrows()`.

ACTIVITY 11: Effect size

Steege et al. (2014) computed the effect size required for a statistical power of 33 per cent, based on Vul and Pashler's (2008) original sample size ($n = 255$). Compute that effect size for a one-sided paired-sample t-test, assuming $\alpha = .05$.

ACTIVITY 12: Statistical comparison of accuracies

Conduct a one-sided paired-sample t-test to assess whether the averaged responses are significantly more accurate than the first responses (i.e., whether the MSE of the average response is lower than the MSE of the first response). Do the same for the second guess. What do you conclude?

Concluding Remarks

We find that the mean MSE of the first guess is significantly larger than the mean MSE of the average of the two guesses. We also find that the mean MSE of the second guess is significantly larger than the mean MSE of the average of the two guesses.

Together, this means that we can follow Steege et al. (2014) in concluding that 'averaging two guesses within one person provides a more accurate answer than either guess alone' (7).

References

Steege, S., Dewitte, L., Tuerlinckx, F., & Vanpaemel, W. (2014). Measuring the crowd within again: A pre-registered replication study. *Frontiers in Psychology*, 5 (786), 1–8. <https://doi.org/10.3389/fpsyg.2014.00786>

Vul, E., and Pashler, H. (2008). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, 19 (7), 645–7. <https://doi.org/10.1111/j.1467-9280.2008.02136.x>