# Step-by-step protocol for interview [Including full list of interview questions]

## Step 1: Introduction

First of all, on behalf of all researchers involved in this study, thank you very much for participating in our study. We highly appreciate the time and effort that you spent to create the fabricated dataset.

My name is Jan. I am a research assistant and I would like to conduct the interview with you today. The interview will take around 45 minutes and it will be semi-structured. That is, I have prepared some questions to understand how you fabricated data.

Before we begin with the actual interview, I would like to ask you to hand over the spreadsheet with the fabricated data you created. You can either give me the spreadhsheet on a USB stick or you can use this usb stick.

[If participant has spreadsheet on their own stick, Jan copies spreadsheet to his laptop]

[If participant does not have their own stick, participants copies fabricated data to EMPTY transfer stick]

Thank you very much. Now, we come to the interview. I will start the recording now.

[Turn on recording]

[CLICK ON RECORD BUTTOM FOR THE FIRST TIME TO CHECK WHETHER SOUND IS OK (Bars should be in the middle)]

[CLICK ON RECORD BUTTOM FOR THE SECOND TIME ACTUALLY START RECORDING]

All personally identifying information will be deleted as soon as possible, except for the recording of this interview. However, the recordings will be stroed on an encrypted USB stick to which only two researchers have access to.

## Step 2: Interview

I will ask you a set of questions. Please feel free to elaborate beyond the actual question. Just to let you know, I will take notes while you speak to make sure that all parts of the fabrication process are addressed. These notes will be deleted after the interview is over.

### Step 2a: Broad question

Please tell me how you fabricated your data set, beginning after you downloaded the instructions from the qualtrics page.

### Step 2b: Follow-up questions (If this was not already covered)

Part 1: Preparations

- Did you prepare in any way before starting to fabricate data?

\* [if yes] how did this influence your approach to fabricating data?

Part 2: Actually fabricating

- What criteria did you use to determine if the data fabrication approach was successful?

- How did you fabricate the data for the Stroop task?

- Did you apply a specific strategy to fabricate data?

\* [if a specific strategy was used] Please describe these strategies

\* [if a specific strategy was used] Why did you choose for this specific strategy (or these specific strategies)? \* [anyway] Did you use a trial-and-error approach?

- Were there things you particularly focused on during the data fabrication in order to make them look more genuine?

\* [if yes] what aspects did you spend particular attention to?

\* [if no] in hindsight, are there things you think you should have paid specific attention to?

- Did you use real data during the fabrication process?

\* [if yes] how much real data did you use?

\* [if yes] how did you use these real data?

Part 3: Test and retest – Improving the data set

- Did you test the quality of your fabricated data set?

\* [if yes] How did you test the quality of the fabricated data?

\* [if unclear] did you try to verify whether these results would be seen as genuine?

- How many different mean-sd combinations did you fabricate before getting to the final, fabricated dataset?

- Did you take into consideration relations in the data other than those specified in the hypotheses?

- Are you familiar with statistical methods to detect data fabrication?

\* [if yes] can you describe or name the methods you are familiar with?

\* [if yes] did you use these methods and if so, how?

Part 4: Others

- What was the timeline of your work? (Did you work constantly on the dataset or did you do everything on one day)?

- How would you rate your knowledge of statistics? (1 = extremely poor; 10 = excellent)

- Which statistical analysis programs do you use at least once a week? [multiple answers possible]

\* SPSS R Stata SAS Matlab Python

\* Other, namely: \_\_\_\_\_\_\_\_\_\_\_\_\_

\* None

- How many experiments including a Stroop task have you conducted?

- How long do you estimate that it took you to fabricate the data?

- Besides the supplied spreadsheet, did you use any other computer programs to fabricate data?

\* [if yes] Did you use a random number generator to simulate data during this study?

Part 5: Final questions

- How do you rate the quality of your fabricated data set? (1 = extremely poor; 10 = excellent)

- How likely do you think is it that your fabricated data set can be detected by existing methods to detect data fabrication? Please answer with a probability between 0 % and 100 %

- Is there anything else you can recall about the data fabrication process that you think is worth mentioning?

Ok, then this is the end of the interview.

[Turn off recording]

## Step 3: Post-Interview

If you sign the receipt we will transfer the 100 euros to you in the coming weeks. The identifying, bottom part of the receipt will be destroyed once the prizes are awarded to the datasets that were hardest to detect as fabricated.

[Hand out receipt]

Note that for the additional 100 euros that you may receive, we first have to determine which dataset is hardest to detect. If you are one out of those three, you will receive an additional 100 euros. We will let you know if you will receive an additional 100 euros.

[Participant signs receipt]

Thank you very much for participating in our study. As a final reminder, I have to emphasize that fabricating data is only allowed within this experiment and should not be taken as an invitation to fabricate research results elsewhere. Data fabrication is a serious form of scientific misconduct and is widely and explicitly condemned by professional organizations, institutions, and funding agencies alike.

If you have any questions about this project, you can contact the Principal Investigator

[Hand over contact slip]

Thanks again for your participation! Have a nice day!