

Legend

1. [REDACTED] means that the original word/fragment was deleted to ensure the anonymity of the participants.
2. [?] is a placeholder for words/fragments that could not be transcribed.
3. (?) means that the transcriber was not completely sure what the last word/fragment was, but had a guess.
4. Sentences that begin with “I:” were said by the interviewer
5. Sentences that begin with “P:” were said by the participant

Block 1: General Information

I: Ok. Then we will start with the first block. The goal of this block is to get some general information about you. So, the first question is: Are you a PhD Student?

P: Yes, I am.

I: Ok. And what is your field within psychology? For instance social or cognitive psychology.

P: It's medical psychology and my research focuses on [REDACTED].

I: Ok ... still running ok ... Did you conduct any experiments including a Stroop task in your career so far?

P: Yeah, in my clinical internship in my masters I had some patients on which I conducted a Stroop task, yeah. I think it might be 2-4 patients, I think, yeah.

I: Ok and could you describe a bit more your experience with the Stroop task?

P: Yeah, it is not a hard task to conduct, not for me, not for the patient. I've had a patient with dementia, so it was quite hard for him, but for the other patients not that hard. And we have had the Stroop task with 100 trials, so it is a bit more than the patients in this fabricated data set.

I: Ok, thank you. And which statistical analysis programs do you use at least once a week? Multiple answers are possible. For instance, SPSS, R, Stata, SAS, Matlab, Python, or any other?

P: I mainly use SPSS, yeah.

I: Ok. And how would you rate your knowledge of statistics relative to your peers on a scale from 1, extremely poor, to 10, excellent?

P: I think it might be 5 - not poor, not excellent.

I: Ok. And how confident are you that your fabricated data will go undetected as fabricated? Again on a scale from 1 to 10, where 1 means extremely insecure and 10 means extremely confident.

P: I think 4. It is not that bad but I am sure that there are a lot of statistical methods to detect the fabricated data, yeah.

Block 2: Timeline of Data Fabrication Process (When?)

I: Ok. Then this is the end of the first block about general information. Now, we will start with the second block. The goal of this block is to get some information about the timeline of the data fabrication process. So, the first question is: Did you fabricate the data in one day or spread the data fabrication over several days?

P: I spread it over several days. I think, it might be 4 or 5 days actually.

I: 4 or 5 days, ok. And how much time do you estimate that it took you to fabricate the data in their entirety?

P: It took me a lot more than expected. I think it might be around 2 to 3 hours.

I: Ok. And how much effort do you feel you invested in fabricating the data on a scale from 1 (no effort at all) to 7 (a lot of effort)?

P: Yeah, it might be 5, I think. 5 or 6.

I: Ok. Did you prepare in any way before starting to fabricate the data?

P: Yeah. I searched the internet for trials maybe, for means, standard deviations on the Stroop task so that I have got an idea about what to fabricate. And I have looked into my results of the former patients of my clinical internship, what they did on the Stroop task before, yeah.

I: And how much time do you estimate you spend on preparing?

P: I think, half an hour.

I: Ok, and did you read literature on detecting data fabrication?

P: No, I didn't, no.

I: And did you look into previous cases of data fabrication and how they had been detected?

P: No, neither (?).

I: And did the preparation that you did influence your approach to fabricating the data?

P: Yeah, because I saw articles in which they stated that the congruent - that task - patients generally needed less time than (?) for the incongruent task. So, yeah, it influenced my fabrication a lot, yeah.

I: Ok. Then this is the end of the second block. Do you have any other comments about the timeline of the data fabrication process that you think could be interesting for us to know?

P: No, I don't think so.

Block 3: Broad Framework of Data Fabrication Process (What?)

I: Ok. Then, we will now start with the third block. The goal of this block is to get some information about the broad framework of the data fabrication process. So, the first question is: Could you name specific characteristics that would make data look fabricated or more fabricated in your opinion?

P: I think, yeah, that we have got some - that we have all got some patterns in which we like to work. And I think that if you find some patterns in fabricated data that you might see that it is fabricated. For example, if participant 1, 2, 4 is a mean time of 10, 20, 30, 40, yeah, that is a pattern and I think that might be fabricated because I think people like to work in patterns.

I: Ok. Other characteristics you can think of that would make data look more fabricated or?

P: Yeah, I think if data is highly significant it might be fabricated because, yeah, in real life data is not that significant most of the time.

I: And could you name specific characteristics that would make data look genuine or more genuine in your opinion?

P: Yeah, I think if data looks like - yeah, like average data. That's what I have tried to do. To make my participants as average as possible with maybe 1 or 2 outliers. Yeah, and it should look random, I think.

I: Ok. And did you take these characteristics you just mentioned into account when fabricating the data?

P: Yeah, yeah I did. Yeah.

I: And how did you take these into account?

P: Yeah, I tried to avoid patterns in my fabricated data. And as I said before, I have looked up the average time for one trial and I have tried to take the average for the majority of my participants with 1 or 2 outliers, yeah.

I: Ok. And did you take into consideration relations in the data other than the Stroop effect itself?

P: Aehm ...

I: For example, the distribution of the scores or other aspects that could be inspected with the data set.

P: Yeah, I did. I found a site - a website online in which there was a chart which has specified the proportion of participants who scored on average between 0 and 5 seconds for 30 trials, between 5 to 10 seconds for 30 trials. And I have tried to incorporate that into my data so that, for example, 30 % of my participants scored between a particular range.

I: Ok and what criteria did you use to determine whether you thought your fabricated data would go undetected?

P: I am not sure I understand what you mean.

I: So, like did you use any criteria to see whether your fabricated data looked like - or like would be difficult to detect as fabricated?

P: Yeah, if I used specific criteria?

I: Yeah.

P: No, I don't think I did. I looked at it and I thought, yeah, this might not be fabricated. And that is my only, yeah, justification.

I: Ok. And in hindsight, are there things you think you should have paid specific attention to while fabricating the data?

P: Yeah, as I said before that it should not look too much in patterns and that there is some variety but not too much variety. That there are average participants with some outliers. It might be that.

I: Ok, then this is the end of the third block. Do you have any other comments about the broad framework of the data fabrication process that you think could be interesting for us to know?

P: No.

Block 4: Specific Steps of Data Fabrication Process (How?)

I: Then, we will now start with the fourth block. The goal of this block is to get some information about the specific steps of the data fabrication process. So, could you indicate what steps you took to fabricate the means for the participants?

P: Yeah. I have - actually, I understand the instructions wrong the first time. So, I thought I had to fabricate the total amount of time for the entire congruent Stroop task. And I thought, now (?), in my former experience there were 100 trials and so I fabricated the total amount of time for each participant for 100 trials and then I thought, no, that's not ok, I have to do it for 30 trials. So, yeah, I have divided it so that it is total amount for 30 trials. And then I thought, no, I have to take the mean of the 30 trials. So, I have divided it by 30 and, yeah, I think, maybe these mistakes are actually good because my numbers are divided

and multiplied and maybe that's something which is good in a fabricated data so that it might not look that fabricated.

I: Ok. But could you describe like how you did it from step to step? So, you said you started first with like looking up information and ..

P: Ja, I have started up looking like information about the mean total amount of time for the entire Stroop task in the congruent condition. And I think it might be - I have [?] it here somewhere - oh yeah, yeah, it was between 6 seconds and 15 seconds or something, so I thought, yeah, that might be great. Then I have looked into my patients but they were patients of 70, 80 so I thought they may be a bit too high for an average participant and then I have went to a website in which you could yourself do a Stroop task of 30 trials and I did it 4 times and I came like 12 seconds, 13 seconds, 15 seconds for 13 trials. And then I divided it by 30 and, yeah, that's when I tried to - well, that is an average of - I don't know exactly but maybe half a second per trial. And then I decided, well, then I am going to randomly type some numbers between 300 milliseconds and 800 milliseconds for each participant so that I've got a lot of mean, yeah, second trial things (?) and then I went to SPSS because I saw that there were actually a lot of participants scored between - how could I say - 300 milliseconds and 500 milliseconds in a large population. So, I have changed a couple of numbers so that the proportions were as equal as the website displayed for the general population.

I: Ok. And could you indicate what steps you took to fabricate the standard deviations for the participants?

P: Yeah, I have took the first participant and, yeah, for this participant I say, yeah, he is roughly about - what is it - 600 milliseconds on average, so in SPSS I have typed 30 - yeah, I don't know how to call it - 30 numbers which are around 600 milliseconds and then I have let SPSS decide about the standard deviation would be. It was around 42 and so, yeah, I have decided myself that 40 seconds might be 32 and for 30 it might be 46. So, for the first one, I actually did fabricate all the trials and for the other participants I have decided what might look good.

I: Ok and like for the variation of the means and standard deviations, then like you just - like you eyeballed and then like you used different numbers for these or did you like create random deviations or so?

P: No, I just thought, well, this might look great and when participants had low mean time for each trial I tried to keep the standard deviation also low. And for high mean times, high standard deviations.

I: Ok. And did you repeatedly fabricate data until you were satisfied with the results?

P: Yeah, I did. And this is especially true for the mean times because I - as I said - I first just typed some random numbers between range which I intended but then I saw the chart with the proportions and the general population and

I have randomly changed some numbers to let it look closer to the chart. So, yeah, I might think I have hit each number multiple times, yeah sure.

I: Ok. And how did you determine whether you were satisfied with the fabricated data or that they needed to be adjusted?

P: Yeah, that's when I put it in SPSS and I have looked at the variation between the numbers and the proportions between - say 100 milliseconds and 200 milliseconds was 10 % of my population and it was similar to the general population, I thought, well, that might be great.

I: Ok and did you try to inspect whether the fabricated data looked weird?

P: Yeah, I've tried to look at if there were similar numbers which might be ok, I think, because in real life there might be some participants who on average have the same mean time, but I tried to have not too many similar numbers. Yeah, as I said, I tried to make it look like an average data set with some outliers and some double numbers, but not too much.

I: And did you try to inspect whether the fabricated data looked genuine?

P: Yeah, more or less, as I said I have tried to make it look like the general population with the proportions but that's it, I think, yeah.

I: Ok, and how many different mean-sd combinations did you fabricate before getting to the final fabricated dataset?

P: It might be 5 or 6, I think, yeah.

I: And besides the supplied spreadsheet, did you use any other computer programs to fabricate the data?

P: Yeah, SPSS, yeah.

I: And did you use a random number generator to simulate data during this study?

P: No, I intended to do that but it was too complex and I thought, well no, I randomly type in some numbers myself.

I: Ok and did you use real data during the fabrication process?

P: Yeah, I have used my own data as I said before and I have tried the Stroop task online multiple times. And I have used the data of my former patients although they were 70-80, yeah.

I: And how much - or no - and how did you use these real data?

P: Yeah, to get an idea of what actual data is about - to get an idea of the mean time of an actual participant, yeah.

I: Ok, then this is the end of the fourth block. Do you have any other comments about the specific steps of the data fabrication process that you think could be interesting for us to know?

P: Yeah. Maybe, I have. I saw that my p-value was very low, very significant. But I decided to let it be as it was because I think the Stroop task and the Stroop effect is a highly significant effect. So, I think in a real data set, yeah, it might look like this, ja.

I: So, like when you created the means and standard deviations for the two different conditions, did you start out with different like baselines for the two different conditions? Or did you like adjusted afterwards or?

P: Yeah, I have adjusted the incongruent condition based on the results of the participant on the congruent condition so that it might not be that a participant scores an average of 500 milliseconds on the incongruent and then 1.5 seconds on the congruent. So, I have adjusted it, yeah.

Block 5: Underlying Rationale of Data Fabrication Process (Why?)

I: Ok. Then, we will now start with the fifth block. The goal of this block is to get some information about the underlying rationale of the data fabrication process. So, the first question is: Did you consider fabricating these data a difficult task to complete?

P: Yeah, yeah, it is very difficult. I talked - yeah, before my colleague [?] said that it might be easier to actually find 25 participants and conduct the Stroop task than fabricate the data, yeah. It's difficult, yeah.

I: Can you describe what you found difficult about it?

P: Yeah. I have invested a lot of time into thinking what might be a good data set, what might be a good mean, what might be the average time in which people could complete the Stroop task. I have invested a lot of time in determining what proportions could be right. You have to think a lot about a lot of things in order to make a data set look genuine and not fabricated, ja.

I: Ok. And do you think that your approach to data fabrication will be difficult to detect as fabricated?

P: Oh well, I don't know. As I said before, I think there are a lot of statistical programs and analyses that (?) could probably detect my fabricated data. But I think, maybe because of the mistakes I have made - that I thought it might be 100 trials, the total amount of times for 30 trials - that it might be more difficult because there is less of human - how do I call it - yeah, human investment in the numbers than when I originally came up with these numbers. But I am sure it will be detected.

I: Why did you decide to participate in this study?

P: Oh well, I am a PhD student myself so I know how hard it is to find participants for research. Yeah, and I thought, well, it might be nice to try something like

fabricating research. Yeah, it looked interesting. I thought, well, why not, let's give it a try.

I: And did you discuss this study or the fabrication of the dataset for this study with other people?

P: Yeah, I did. I have discussed it with [REDACTED]. She is also a participant in this study and we have discussed how hard it is to actually fabricate data.

I: Ok, and did these people help you in fabricating the data?

P: No, no, because she had to fabricate data and I had to fabricate data. So, we have discussed general things but not how we fabricated.

I: Ok. Then this is the end of the fifth block. Do you have any other comments about the underlying rationale of the data fabrication process that you think could be interesting for us to know?

P: No, no. I just think it is hard to fabricate data, yeah. Yeah, it is.

I: Ok, then this is the end of the interview or is there anything else you can recall about the data fabrication process that you think is worth mentioning?

P: No, I don't think so. I think I have said everything.