

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 now. The goal of this block is to get some general information about you. So, the first question is: **Are you a PhD Student?**

P: I am.

I: Ok. And what is your field within psychology? With field, we mean for instance social or cognitive psychology.

P: Yeah, so I am in developmental psychology.

I: Ok. And did you conduct any experiments including a Stroop task in your career so far?

P: **Eh, no, never.**

I: And could you describe your knowledge or experience with the Stroop task a bit?

P: Well, I knew about the Stroop task, but I - well, I might have done one sort of funny internet thingy (?) setting but never in an experiment here at psychology.

I: Ok. And you learned about it during your studies or ..?

P: Yeah.

I: Ok. 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: **Yeah, R.**

I: R, 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: Probably **something like 8.**

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 6 or s... - yeah **h 6, I guess.**

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

I: Ok. Then this is the end of the first block about general information.

P: Alright.

I: 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 did **it in one day.**

I: Ok. And how much time do you estimate that it took you to fabricate the data in their entirety?

P: I think maybe **4 or 5 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: Well, I would say **6, quite** some effort.

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

P: Yes. So, should I elaborate how?

I: Ehm yeah sure.

P: So, what I - so, I did not really have any knowledge of the Stroop task. So, I wanted to get some feel for data like the Stroop task. So, first, I went online and googled something like what's the Stroop task, can I do it online, find a site where I could play, checked my own reaction times, get a bit of a feel. And then I didn't feel that was really good enough. So, I thought, well, there is probably data online somewhere of the Stroop task. **So, then I went to the Open Science Framework and I found a large data set with Stroop task data and I used that to get some feel for that kind of data you get.**

I: Ok, let me just check whether it is still running. yeah ok. And how much do you estimate you spent on preparing?

P: So, many 2 hours or something. **2 to 3 hours, yeah.**

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

P: Ehm, no.

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

P: Well, not really, but I did think about this. And I have heard - so I have heard the most obvious cases in which, well, data fraud is detected is because lack of variance between studies or yeah. So, I did consider that.

I: Ok. and did the preparation that you mentioned influence your approach to fabricating the data?

P: Sorry, can you repeat the question?

I: Did you use what you did during the preparation time to - like when you actually fabricated the data?

P: Yeah, yeah, yeah, for sure.

I: And how did that preparation inform your approach to fabricate the data?

P: So, I looked at the distributions of this data and tried to fit distributions to the data to get a feel for that so that I could later use to simulate the data. I checked correlations between - yeah - the different measurements, so for instance between the mean reaction time and the standard deviation and stuff like that.

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, not really, I guess.

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, could you name specific characteristics that would make data look fabricated or more fabricated in your opinion?

P: So, I think low variance probably and ... yeah, so, I wasn't really sure (?), I think it's - obviously if maybe effects are too large or maybe if there aren't correlations between for instance the mean reaction time and the standard deviation that are likely to be there in real data.

I: Ok, you said that one characteristic could be the variance. Could you like describe in a bit more detail what you mean by the variance?

P: Well, I think, people - most people are like onerous (?) in estimating variance. If you ask them to draw a random sequence of heads or tails or something like that, typically variance is lower than you would observe in real data.

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

P: Yeah, so what I thought was if the correlation structure that is there in real data is preserved in the simulated data.

I: Ok. And did you take ...

P: That was sort of my approach to the problem.

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

P: Well, yeah, I tried.

I: Ok. And how did you do this?

P: So, first - so, we had to simulate 4 measurements per person. So, the mean and standard deviation of the congruent trials and the mean and standard deviation of the incongruent trials. So first, I thought, well, there is probably correlations between the mean and the standard deviation. And then there is also probably correlations between the reaction times on the congruent trials and on the incongruent trials and also for the standard deviations. So, in this data set that I found on the Open Science Framework, I checked - I checked these correlations whether there were present and how large they were and - yeah, starting from there, I started to simulate data with that correlational structure.

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

P: So, for instance?

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

P: Yeah, yeah. So, I did. I checked for the structure for the difference scores and for the distribution of reaction times, which I knew to be mostly non-normal but more gamma(?) - distributed. So yeah, I did consider these properties.

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

P: Hm. Well, I just tried to reproduce the similar correlational structure that I saw in real data. And I thought that would make it look real.

I: Ok. And did you use different criteria for the means and standard deviations?

P: Well, I thought, because they are dependent I used this dependency to simulate the data. So, I started off with the - I th... - yeah, what I did was I checked for the difference scores between the reaction times on the congruent and the incongruent trials. And then - so then I had one difference score for each participant. And then I also checked for difference scores for the standard deviations. And I saw that these difference scores were more or less normally distributed - and correlated. So, large differences in reaction times also meant large differences in standard deviations. So, I thought I (?) should incorporate that. So, that was sort of my starting point. And then I simulated from two

correlated normal distributions and I saw (?) like mean differences between these reaction times were something like 50, 60 milliseconds. And I think for the standard deviations something similar. So, I simulated data with those means and large variances. I think something like - I don't know - standard deviations of 100 to 200 milliseconds. And I correlated - or I simulated correlated data so then I had the difference score data and then I simulated using gamma distributions reaction time data and standard deviation data on the congruent trials and then I added the difference scores that I simulated earlier for the - to come up with the incongruent trial measures. Yeah, so that's what I did.

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

P: Well, not now. But probably, in later hindsight, yes.

I: Ok, but at the moment you can't think of anything?

P: No, at the moment I can't think of anything.

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 that you think could be interesting for us to know?

P: No.

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

I: Ok. 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. Could you indicate what steps you took to fabricate the means for the participants?

P: Yeah, well, I think I just elaborated on that. So, I started with - I started from the difference scores - and I started by simulating difference scores. And then I used those scores to simulate the means of the data.

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

P: Yeah, so that was similar. So, I looked at the data I had of the Stroop task. And I - I checked these difference scores so it was approximately normal. So, I simulated a normal distribution of these difference scores. And then from a gamma distribution, I added - I simulated 25 observations and added the difference for the incongruent trials.

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

P: Well, not really. I think I did like fabricate the data multiple times to check like - more or less eyeballing the data - if the correlational structure that I

observed in real data was found in the simulated data. And then I don't know maybe after 10 or 15 times - obviously it is a simulation so it comes out different but sort of similar every time - so then I thought, well, that is ok. And then I just copy-pasted the data I had then into the Excel file and it gave me a p-value that supported it so I thought ok, that is ok.

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

P: Well, I didn't really - I couldn't really think of anything else to make it better or more real. So, I decided it was enough.

I: But you mentioned earlier that, for instance, you checked the correlation structure, right. So like did you do this after each simulation - you checked this and then based on this you like decided whether you would simulate another round or not?

P: No, not really. Yeah, I am not really sure. Probably, [?] I saw it and I was like this is reasonable or ... but not really like - I didn't go through a - how do you say it - thorough process of like simulating many, keeping them and then using some of sort selection criteria to determine which one would be the best or most real data set. Just ..

I: So, you had sort of gut feeling ...

P: Yeah.

I: ... that told you this is good ...

P: Yeah.

I: ... or I should do this again.

P: Exactly.

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

P: Yeah, well, I did. So, this is a point. Because of the method I used I somehow sometimes ended up with negative standard deviations which is a bit odd. So then I just some of sort of last (?) trick just to put that - always be a positive number.

I: Ok. So how did you change the negative numbers to the positive numbers? By another simulation or just by hand?

P: No, well, I just set it to the absolute value. So, I just checked if it is negative and (?) set to the absolute value.

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

P: I think I did, ja. And then I started wondering how can you ever tell from these 25 data points if they are derived from a gamma distribution or from a normal distribution or ... yeah.

I: Ok. So, like you had sort of a gut feeling whether this looks genuine and - but you didn't apply like specific criteria about what to check ...

P: No. I did plot histograms of the data and I checked for negative values.

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

P: I think maybe 15, 20.

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

P: Yes, so I used R.

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

P: Yeah, I think so. So, I used the R random number functions of the - I think the rgamma and the rnorm functions to simulate the data.

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

P: I used only real data to - not to fabricate the data but only to get a feel for the kind of distributions these data have, yeah

I: Ok. And this was based on a large data set that you found online, right?

P: Yeah, exactly.

I: Ok. And did you use specific numbers from this data set or just to get a feeling from it?

P: No, more to just get a feeling. So I think what I did was - yeah, it was a large data set so I just looked at the distributions to see what it was. And then I draw like samples of 25 observations from that - from that data a couple of times to get a sort of feel how 25 data points would look like.

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: No, I think I kind of told everything of how I fabricated the data.

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: Not really, no. So, I have a lot - well, I don't know if I have a lot but I have some experience in simulating data. And if - yeah sort of fabricating data. And I use it in my work to get some estimates of power or to get a feel of what my

actual data would look like - to, well, plan analyses in advance and stuff like that. So, it is a common task for me to do.

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

P: Well, I am a bit uncertain about that. So, I know I wouldn't spot it as fabricated data but I am not aware of any detection methods or anything other (?). So, it is kind of hard to get a feel for that, yeah.

I: But could you think of methods that like that - or like things that you think might suggest that the data set is fabricated?

P: Yeah, so these sort of things that I took in consideration when simulating the data. So, I thought how would I try to detect if data is fabricated or not. And then, well, the sort of obvious things are then that there is low variances or weird scores are not in the data or there is no correlation between dependent measures that you would expect or - yeah, so I tried to incorporate that as much as possible.

I: Ok. Then why did you decide to participate in this study?

P: Well, I thought it is fun. So, I think it is important to get a bit of a feeling for - if you can detect these things if it is real data or fabricated data and, well, for the reasons to prevent fraud and etc. etc. But I also think it is important to - or I think it can be very useful to fabricate data for the reasons I just mentioned to get some power estimates, to plan your analysis ahead and stuff like that, yeah.

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

P: I only mentioned it with my supervisor who also wanted to participate but he didn't have the time - or he was too late, I don't know.

I: Ok, but ...

P: So ...

I: Did you receive any help in fabricating the data?

P: Nonono, not at all. I did it all by myself.

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, not really.

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

P: No, I think I told you everything of how I did it. So, yeah.