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: Now we will start with the first block. The goal of this block is to get some general information about you. Let me just check, yeah, it is still running. So, the first question is: Are you a PhD Student?

P: No, I am not. I am a junior researcher. So, I don't have a PhD yet, but I am also not in a track to get one.

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

P: Social psychology.

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

P: Not a Stroop task. I conducted interviews, [REDACTED], and one experimental - one experiment, so in the lab. Do you need to know more about that?

I: No, but could you describe like your knowledge or your experience with the Stroop task a bit more?

P: Oh yeah, ok. So I don't have any experience with the Stroop task other than that I know what it is and those (?) and I have been a participant in it, but I didn't use it for my own research.

I: Ok. But you have learned about during your own studies?

P: Exactly, yes.

I: Ok. And did you also read some papers about it so far or so?

P: Yeah, in the past or I think I read about it in a lot of papers when I was a student. And also for this research, I read some papers about it to be able to fabricate data that is naturally for the Stroop task.

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, just SPSS.

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: Ok and with peers you mean fellow researchers?

I: Yeah, I mean other researchers or scientists in your field.

P: Ok, so I'd say 7.

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: 5.

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, did you fabricate the data in one day or spread the data fabrication over several days?

P: I spread it over several days. I did some things - part of it I did two weeks ago and then another part two days ago.

I: And on how many days did you work on fabricating the data?

P: Like total days or . . . ?

I: No, like . . .

P: I think I did something with it on three different days.

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

P: I think 8 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: 6.

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

P: I prepared by searching for previous research on the Stroop task and, yeah, I mean this was part of fabricating it, I wrote down the means of - and standard deviations of these previous researchers for a lot of them. So, I searched 14 different papers with 14 different means and I read a lot more papers, but a lot of those papers had special conditions with the Stroop task, for example anxiety

disorders or schizophrenia or something. So, I didn't use those. So, that was a lot of preparation, I feel.

I: Ok. And how much time do you estimate you spent on preparing?

<mark>P: 3 hours.</mark>

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

P: No.

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

P: No. I actually haven't even thought of doing that. That would have been good preparation.

I: So, are there other ways in which you prepared? So for instance, did you think about different approaches how to fabricate the data or so?

P: Yeah, not very extensively, but I thought - I just thought about how natural data would look, that there would be some outliers, for example, that it wouldn't be perfect or maybe too significant or too predictable. So, I thought about that and tried to remind that all the time when I fabricated the data.

I: Ok. And did this preparation influence your approach to fabricating the data?

P: Yeah, I think so because I reminded myself of it.

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 think I did all the means I filled in on a different day than I did the standard deviations.

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: Yeah, I think it would look fabricated if it is all too neat and too like wishful thinking or too perfect - if it looks too perfect, then I think it would be suspicious. If there are no cases that go against what you predict, so are (?) maybe the other way around, then I would think maybe that is suspicious, but it could also be the case. So an example - especially with the Stroop task what is very reliable, because it has been replicated so much, I thought it was very difficult to decide how many outliers or weird things I should put in there. And I thought, maybe if you have a lot of - maybe people have a preference for certain numbers

3

when they write it down or try to avoid double numbers or numbers like 100 or 200, because they think, no that is too obvious, but in a real world situation the chances are very likely that a number like that would be there. So, I also tried to correct for that by sometimes putting like 117 or something or even 200, I think, is in the - something like that.

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

P: Yeah, so if there are some mistakes and some things that a researcher wouldn't be happy about, some things that you don't want in your data. If those things are present, then I think that will look more genuine.

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

P: Yes.

I: And how did you take them into account?

P: When I wrote everything down, I tried to correct myself by looking at how many times I had the same end number and starting number and I put in on purpose some outliers and some, yeah, numbers that maybe some other people would try to avoid like 400 or something.

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

P: What do you mean?

I: For instance, the distribution of the scores or other aspects that could be inspected with the dataset?

P: Oh yeah, so after I did all the means, I calculated myself in SPSS what the standard deviation was of these means and if that was corresponding to the standard deviation that could you expect from a Stroop task and I corrected for it. And also I looked at the spreading just in Excel, in a different Excel sheet, I tried to see if all the numbers were on a linear distribution and it wasn't, so I was glad I checked that. So, I also corrected for that.

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

P: What criteria you would check it on, you mean?

I: Yeah like or like criteria that you used to determine whether you think that you had a good chance that your fabricated data would go undetected?

P: Yeah, so the approximately linear spreading, but not too linear. And that at least the mean and the standard deviation overall were very much similar to the means and the standard deviations I have seen in previous research. So, at least I know that those are realistic, that was the most important thing that I tried

to keep track off that outcome - so if my fabricated data were in correspondence with realistic data.

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

P: Yes, so I used - I think I used a standard deviation to calculate the spreading of the means. Of course, I did and that's why I checked it after I did the means in SPSS to see if the standard deviation was alright, so I am very confident because of that that my means have a realistic spreading, but for the standard deviations I thought it was very difficult, because it was really hard to find in the literature standard deviations of individual cases, so that was - and I really had to dive into statistics again to see how individual standard deviations deviate from the overall standard deviation. And in the end, I was able to find some data from ==google from individual cases that I used to find a mean individual standard deviation and from that mean individual standard deviation I created all the other individual standard deviations and here I tried to correct myself where my - wanting to put random numbers in by just going to a website like it was called randomnumbergenerator.org and== I asked it to generate 30 - because I thought there were 30 cases, but there were 25 at the time I did this - 30 random numbers between 150 and 250 and then I just use those for the congruent condition. So, the first 25 of those random numbers I used for the congruent condition and then from there on I noticed that usually in the incongruent condition standard deviations were a little bit higher, so then I adjusted it myself that some of those standard deviations were somewhat higher, but some also somewhat lower. So here I also created some outliers and some difference between the numbers and that was random in my head, but probably there is not completely random, so I might go wrong there.

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

P: ==Yeah, so when you mentioned that I could have - when you asked if I researched data detection articles, I think I== could have paid more attention to those and tried to find out how you were going to do the detection instead of just thinking about the things I can come up with.

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, I don't think so.

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.

So, could you indicate what steps you took to fabricate the means for the participants?

P: Yeah, I think some of this I already said, but it might be good to repeat it. So at first, I looked at other studies and wrote down the 12 or 13 means of other studies. And then I used those means and I - then I just came up with new means for myself that were around those means. And then I did - I calculated the standard deviation in SPSS to see if that standard deviation was corresponding to a realistic one. And then I adjusted the means again and again until it was realistic standard deviation or the one I decided for myself that I wanted it to be. And then I looked at all the means again and I tried to put the means in random case rows so it wasn't sanding (?) all. And then I think I adjusted some numbers again and again and again and hoping that it would be sort of random, which obviously probably is not. But, yeah, I did it like that.

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

P: Yeah, so at first, I didn't know about individual standard deviations, so I inserted wrong ones. And then when I looked at it another day, I thought, no, this cannot be right, so I googled for Stroop with a Excel S between brackets so I would get Excel sheets with data hopefully or I tried like a 1000 different google searches, but this one in the end was one that really got me data. So, but most of those data sheets didn't include individual cases, only one. So, I used that one and I had to put them all in SPSS and calculate the standard deviations, because those were standard deviations even per trial. So, I had to first calculate the standard deviation for a set of trials. This was the part where I really don't know if I did it right or not. But I thought that in the end, the standard deviations were somewhere between 150 and 250, but there were a lot of outliers on reaction time in that file which was very difficult because in the instructions of the study it wasn't said if we could remove outliers or not. And in most studies where they do the Stroop task they remove outliers over 2000 milliseconds or they even already present people with a new one. So, I didn't know from your instructions what the Stroop task would exactly be performed like -if participants would be presented with a new word after 2000 milliseconds or not. So, I just decided for myself to remove the outliers for from above 2000 milliseconds. And then I found the standard deviations were between 150 and 250 on average. So, that's when I used the random number generator to generate random standard deviations for the congruent condition. And then from there, I did myself somewhat random numbers for the incongruent condition, where I made sure that most of the time the standard deviations in the incongruent condition were higher than in the congruent condition.

I: Ok. And did you also look for - or like did you do the fabrication of the means completely separately from the fabrication of the standard deviations or?

P: Yeah, I did it almost separate. But when I allocated the standard deviations to the cases, I looked at the means so that I gave low means a lower standard

deviation than really high means, because when the means are higher then there is probably an outlier and an outlier has a really big effect on the standard deviation. So I then - those higher means I gave a higher standard deviation. But also sometimes I turned this around, because some people are just slower, so they can have a high mean but a low standard deviation, because they are quite accurate, but just slow. So, yeah, I didn't really know how to handle that, but I tried to look at that.

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

P: Yeah, I twitched (?) it a lot until I thought, yeah, now it looks good or now I am happy with the spreading and with the mean and with the SPSS statistics.

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

P: Yeah so with SPSS and the Excel I tried to see if I - if those things were ok and then if necessary, adjusted and re-check it.

I: Ok so, yeah it may be a bit repetitive, but could you say again the criteria that you used to like determine whether you were satisfied with the fabricated data? Like what did you check in SPSS and Excel?

P: I checked the standard deviation of the means. I used SPSS to find standard deviations of the data I found online, so from there I didn't do any checks anymore, because I wouldn't even know how to do that. And I checked the spreading of all values in a - just a scatterplot to see if they were linear.

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

P: Yeah, but just with - just how I perceive it and the plots.

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

P: Yeah, but with the same measures (?).

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

P: What do you mean exactly?

I: Like, so, in the final Excel file, you had the 25 . . .

P: 25, oh, if I did more than 25?

I: So, you first created the data that you could have transferred to the Excel file and whether you then like changed the combinations or whether you just like had the first final thing that you created and you directly transferred it or not?

P: Yeah, I immediately put in 25 from - in the Excel sheet and I worked actually constantly in the Excel sheet changing the means when I thought that it was necessary.

I: Ok. And how often would you say that did you make some changes?

P: So, probably every mean was changed twice or three times.

I: Ok. And the same for the standard deviations?

P: The standard deviations were four times changed, I think.

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

P: SPSS and, yeah, Excel in a different sheet.

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

P: Yeah.

I: With the one online, right, that we talked . . . ?

P: Yeah, yeah.

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

P: Yes. So, the data I found online from the standard deviations that were really individual cases and the data from 14 previous studies.

I: Ok and did you use these data more as an inspiration or did you really copy data from some cases in your - in the Excel file?

P: I used it as an inspiration mostly, but I think I chose three means that I thought were very reliable because the studies had a really big n. So, I used those three means and those - from the congruent and incongruent condition as a standard to do the other 22.

I: Ok. And did you change these means in the process or did you keep them as they were?

P: I probably adjusted one number, so instead of 679 I did 678 or something.

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 just - when I look at my notes now, I see that I also before - actually the first step I took was doing an online test myself and write down my own personal score on the Stroop test - task.

I: Ok. And you also compared those scores then to the other scores or did you do anything . . . ?

P: Yeah, I was very slow. I was very slow in comparison but I did see the difference between congruent and incongruent, so it still gave me a little bit more feeling with the task.

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, it was more difficult than I expected. The means I didn't think were so difficult but the standard deviations were really difficult, I think.

I: And what was difficult about them?

P: That I really had to dive into statistics and how it actually outworked to come up with realistic numbers. And even though I used real data as an inspiration, I am still not sure if the spreading of the standard deviations is right. So, like the standard deviation of the standard deviation. I do not know if I did that correctly.

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

P: No, I think probably you will figure it out.

I: But could you think of ways how your fabricated data could be detected? Is there like a hunch that you have or so?

P: I don't really have a hunch. I just think there is a lot more than I came up with just by myself and probably you have a real - a good program thing with different algorithms that maybe my test is even too significant or something, that was something I found hard to decide because yeah that is difficult. So, I don't know, maybe my data is too significant and I don't know what exactly, but I am pretty sure there are things that I haven't thought of, that will come up.

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

P: Because I think it is very important that people don't come in the temptation to fabricate data. Because I usually don't like it that people don't take social science seriously and it hurts me as a part of the group social scientists that sometimes there are some bad apples and then you know the other scientific fields think of us psychologists as not to be taken serious or fraudulent, which I think is very bad. So, if I can help in any way to prevent that further, then yeah ... And of course because I get monetary compensation.

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

P: I discussed it with my colleagues, but not the content of my work so I didn't discuss with them how I should do it, but just that this study was there.

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

P: No, no.

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.

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: I am just looking through my notes, but I think we have discussed everything about it, so yeah.