Interviewer #1: Okay, ready.

Interviewer #1: So the first thing to do, this page is basically what you saw in your survey, which is, we were showing you the prior probability density the first is the intercept parameter and the next is the [inaudible] parameter. So could you please walk me through what you did with interface and how you decided to choose the priors. The prior you chose was 3.50.2

Participant #2: Sorry, what did I choose?

Interviewer #1: That's written over there, your prior for the intercept.

Participant #2: Okay, right. Okay, so you want me to just walk you through my thoughts?

Interviewer #1: Yes.

Participant #2: Okay. So, let's just start from the top here. There's some taps and it's counting pumps or something, and I really paid attention to the prior studies, and in particular, there's this bit that says, "A meta-analysis of 22 studies, which used [inaudible 00:01:15]..." yadda yadda. So the average number of pumps averaged across conditions varied between 24.60 and 44.10.

Interviewer #1: Okay.

Participant #2: So I took that as some sort of interval estimate. And I said to myself, "Okay, well assuming that maybe this meta-analysis was done, I don't know, they used a normal likelihood or something, maybe the interval is stemetric about the mean.

Interviewer #1: Okay.

Participant #2: So what I did is I computed the midpoint of that interval.

Interviewer #1: Okay.

Participant #2: I think that's 40.0 [inaudible] So that's about 34.35.

Interviewer #1: Okay.

Participant #2: My interpretation of what the information was telling me was that the sort of expected number of pumps was 34-35.

Interviewer #1: Okay.

Participant #2: Okay.

Participant #2: With a standard deviation of 5.3. Cool. So now the regression model. You tell me about the regression model. In this framework, the expected number of pumps would be e^Alpha, right?.

Interviewer #1: Right, yeah.

Participant #2: So I said okay, for the intercepts, my prior should be centered around a value such that E to that value would be around 34-35.

Participant #2: So if I did something like log 35 or 34.5

Participant #2: Oh shit, is that what I did?

Interviewer #1: No that's fine, go ahead.

Participant #2: Sorry I don't remember what I was thinking.

Interviewer #1: What do you mean by that?

Participant #2: Sorry, what did I mean by what?

Interviewer #1: As in, it seems that you realize you made some mistake or something, what was that?

Participant #2: Sure, I just forgot if I had to take the exponential about them or the logarithm about them [inaudible] exponential about them because e^Alpha will give me, if I take the exponential of alpha that will give me the expected number of pumps. We said the expected number of pumps would be like 35 or something.

Interviewer #1: Right.

Participant #2: So if I do e^3.5, 3.6 that gives me somewhere in that range. My prior for Alpha should be between 3.5-3.6. That's where the mean should be.

Interviewer #1: Okay

Participant #2: Okay, so then I went down here and set my mean to about 3.55.

Participant #2: If I do expo 3.55, .58, .59

Participant #2: Yeah, 3.55 sounds good to me. That's good enough. It doesn't have to be perfect.

Interviewer #1: Yeah.

Participant #2: And now comes the issue of how fat the tail should be.

Interviewer #1: Yes

Participant #2: So back in this piece of information, the majority, I interpreted this interval estimate as saying if it was a 95% confidence interval. If it was coming from a frequentest study then I'm going to make sure that the majority of my density is within this area. So log of 24.60 that's 1.3 and 44.10. What did I do? Basically what I have to do is; I want to find an interval in this space which maps this interval, does that make sense?

Interviewer #1: Okay, yeah.

Interviewer #1: So if I'm interpreting you correctly, so you're trying to make sure that the transformation of the prior [inaudible] and the upper bound of that would somewhat match with the upper bounds of the interval of the meta analysis

Participant #2: Yeah that's right. That's right.

Interviewer #1: Okay that makes sense.

Participant #2: Cool. So one thing I didn't comment on was my choice of student T vs. normal.

Interviewer #1: Yes.

Participant #2: I think I would use.. What would I use here? The benefit of using the Student T, is if I'm really wrong. If the alpha parameter is. Basically the fatter tails allow for larger or smaller values than the meta analysis have given. I think if I was very confident that the meta analysis estimated the correct number of pumps, I might be tempted to switch to a normal but if maybe I thought, if my domain knowledge about the experimental design or something. Maybe if I though that it was more variable, I might choose the student T. I'm going to choose a student T anyways just because it will allow for larger smaller values. [inaudible] data support that.

Interviewer #1: Okay

Participant #2: Okay cool.

Participant #2: So now I actually have to [inaudible 00:08:21]. Sorry, go on.

Interviewer #1: No, that's cool. Let's move on to the [inaudible] parameter

Interviewer #2: Can I ask a question. A couple of things that you went out to do some calculations, are we doing those in a separate window or ...?

Participant #2: Oh yeah, can you see me do the calculation here?

Interviewer #2: No, which is fine. The reason I brought it up [crosstalk 00:08:56]...

Interviewer #2: Sorry?

Participant #2: Nevermind, go on.

Interviewer #2: The reason I ask is it seems like a process that you've gone through before, right? Am I right in asking that? Not necessarily for this study.

Participant #2: This is something familiar to me but not incredibly familiar.

Interviewer #2: Would you say that the way that you've gone through that process, is it similar to what you've done studying these sorts of priors in the past or are there other strategies that you've used in the past for this sort of problem, particularly with this transformation issue?

Participant #2: Honestly, it's sometimes difficult for me to think on the linear predictor scale so sometimes what I'll do is I'll actually simulate draws from my prior and then transform them onto the natural scale. You know what I mean? 3.5, if Alpha is 3.5 I kind of know what that means, but if were to transform it so I'm look instead not at Alpha but at Lambda, then that makes it a little easier for me to choose my prior because I can actually immediately visualize what changes in the prior results in changes to lambda. Does that make sense?

Interviewer #2: That makes perfect sense. So when you do that are you just manually writing a little snippet of code to do that and then, [inaudible] visualization? What sort of tools are you using for that?

Participant #2: No, that's precisely right. I might, in this example, I might sample from a student T distribution with mean 3.5 in standard deviation whatever and then I'll just take the appropriate transform to see how does that manifest into the parameter space, the natural parameter, the Lambda space. Does that make sense?

Interviewer #2: Yeah, that's great. That's really useful, thanks.

Participant #2: No problem. So do you want me to keep going or do you want me to move onto the mean difference parameter?

Interviewer #1: Do you have anything to add about the current one?

Participant #2: No, like I said I would try to find a standard deviations which resulted in the majority of the density being such that the majority of the density resulted in the Lambda between here.

Interviewer #1: Okay.

Participant #2: Does that make sense?

Interviewer #1: Yeah.

Participant #2: Can I move onto this now?

Interviewer #1: Yeah, sure.

Participant #2: According to the regression model this is some sort of setup, experimental setup. We know the effect isn't going to be really really big. In most experiments I would probably say the effect is small, smallish. We aren't going to get a beta of 10. I know that would be incredibly unlikely in my experience. So I know that it's going to be maybe centered around zero, maybe it does nothing. So what I would do for most cases is probably use, again a student T centered around zero. I'm assuming it will do nothing but I'm going to allow myself to be wrong. Does that make sense?

Interviewer #1: Okay. But then in the survey, instead of choosing a student T, you used a normal.

Participant #2: Yeah, that's right. Sorry, yeah.

Interviewer #1: Is there any reason for that?

Participant #2: I guess it depends on...

Participant #2: What was I thinking there

Participant #2: I think like, yeah, the normal. I would choose the normal because the tails are not that fat. I don't think the effect is going to be 5 or 3 so the majority of the density is between 2 and 2 which I think would be where the effect might be found.

Participant #2: Maybe just depends on the type of day. Maybe I feel the effect would be larger or smaller than I originally though and I might be interested in choosing a student T, I guess.

Interviewer #1: Okay.

Participant #2: Does that make sense.

Interviewer #1: Yeah, so when you say about effect and you mean the value of the beta, how do you interpret the effect?

Participant #2: So Beta, thinking on the Lambda scale.

Interviewer #1: Okay

Participant #2: Beta will be a multiplicative effect, right. So if the base line number of pumps was 10 and the people in this experimental group pumped on average 20 times then [inaudible] would be too, right. That would be a multiplicative effect.

Interviewer #1: Okay cool.

Participant #2: Did I answer everything to your satisfaction, do you have any more questions?

Interviewer #1: No I think we got it. More of to get a sense of how you were thinking about this and I think we get a sense of how you came about choosing this prior.

Interviewer #2: I think this was really helpful. One thing I would ask is, as you were going through this process, were there challenges that you had in doing it? Anything that didn't work quite the way you wanted it or could have been different?

Participant #2: I think my biggest challenge was really the interpretation of the prior density on the scale that it's on. Like I spoke before, I think it would be more natural to think on the Lambda scale rather than on this linear predictive scale. But I've already spoken about that. I think it's quite easy. This little [inaudible] here is interesting but I probably would have wanted more [inaudible 00:16:25]. It would have been more useful to have two slider bars instead of it be on this plane.

Interviewer #1: Okay.

Participant #2: Maybe that's just me, I'm not sure.

Interviewer #2: Is that because you want to be able to set something more precisely?

Participant #2: Yeah.

Interviewer #2: That's helpful, thank you.

Participant #2: No, thank you. Cool.

Interviewer #1: Should we move onto the next page?

Participant #2: Sure.

Interviewer #1: So this will take a while to load. We can ignore the first thing now if you scroll down a bit you'll see the prior parameter probability density on your response scale.

Participant #2: Ah there we go.

Interviewer #1: Okay, so this is another way of visualizing the prior probability density, right. In the response scale we basically do the log transformation of the prior probability density.

Participant #2: Right.

Interviewer #1: So essentially what you are seeing is this is the scale of lambda. Then you'll see what the, basically the density shows on the mean number, sorry the number of pumps.

Participant #2: Right.

Interviewer #1: Okay, in this, do you think the information presented in this visualization would effect your choice of priors? How would you, if at all, use this information that was presented in this visualization.

Participant #2: How would I use this one?

Interviewer #1: Yeah.

Participant #2: I think it would be in the same way I described earlier. So if the majority of the density is somewhere in between, what was it, 22 and 44, or something. So what I might do is I might move it here and say well, here's 50. So I know the majority of the density has to be somewhere in this area.

Interviewer #1: Okay.

Participant #2: So what happens if I make it more variable. It might be useful to see where 20 is on this scale. I think, that might be too much. Right, and the mean has to be like 30 something so it has to be farther this way. Assuming that this was centered appropriately.... This looks like it has the majority of the density but I'm really worried about this tail. This is giving credence to values that are outside of that interval that the meta analysis provided. Lets see what a normal would look like. I feel like that would be more, I would be more comfortable with this. I think that the normal would be probably the better prior because it doesn't give as much credence to things beyond 50.

Interviewer #1: Okay.

Interviewer #2: A little interesting there because on the other scale you used the T and here you are using the normal. I wonder if you can reflect on that maybe.

Participant #2: On the other scale it's sort of easy to justify; Oh well, you know know, fatter tails, it might allow for larger effects that I think is right. When you see how that manifests, how those long tales manifest on the scale of lambda. I think what I was thinking that's too much density there. Do I really have the belief that the prior belief that that much density is beyond 50. It's a tricky thing because pictures aren't. Pictures are good but I don't know how much density is beyond 50 here. If I were to switch back to a student T....

Participant #2: Somewhere around there. If I were to have a normal, if I were to see both of them on the same scale, maybe I'd be better able to judge if there was too much density here.

Interviewer #2: Would it even be helpful to be able to see maybe the CDF or some ways of quantiles to be able to judge more accurately how much density beyond certain points.

Participant #2: Yeah, absolutely. I think the CDF would be really good. I think if there was a shaded region within the alpha. This is the 95, the [inaudible] Where 95% of the density lies, I think that would be helpful.

Interviewer #2: Yeah, cool.

Interviewer #1: So the neutral, You'd change your prior to a neutral density, based on this information? [crosstalk]

Participant #2: Seeing how much density, how much probability density was given to values over 50 might make me change from a normal to a T. Sorry, the other way around. From a student T to a normal, rather.

Interviewer #1: So I guess we can move on to the next visualization which is the prior predictor probability. For some reason you are only [inaudible]

Participant #2: Is it rendered correctly now?

Interviewer #1: You still don't have the drop down which is kind of annoying. I guess we can still move on.

Interviewer #2: We can do this. I'm just, that's okay that we only get the T version, that's the only difference. I think we should still do this one.

Interviewer #1: Again, the same thing as. This is the prior predictive distribution so basically it just draws from the prior assuming that the prior is the actual data generating process. It shows you what the predictive distribution would look like.

Participant #2: Right.

Interviewer #1: Right. So, how would you use. Do you think this information presented here would affect your choice of priors in some way. Would you use this information at all?

Participant #2: This is exactly what I was talking about previously. I like that the scale changes here, that's really good. I know that the majority of them should be between 30 and 40. So I would maybe do it like, that looks like the majority is around, the mean is around 30 but maybe it's too variable because we're getting lots of stuff around 50 so maybe I scale it back a little bit.

Participant #2: Yeah, the majority. Is this for Beta?

Interviewer #1: No, this is for the intercept.

Participant #2: So [crosstalk]

Interviewer #2: You're changing the intercept. Beta is fixed, what you are seeing are densities [inaudible] response distribution, right?

Participant #2: Right. I think this is really useful because now I see the majority of the draws coming around maybe this is 30. There's not a ton beyond 50. There's some but not a lot. It would be interesting to see in comparison with the normal but what I'm really looking for is the mode around what the middle of that interval and are we getting lots of observations beyond, outside of that interval.

Interviewer #1: Cool. How would you contrast the information in the prior predictive distribution versus the prior distribution in the response scale?

Participant #2: How would I contrast these? It's useful to see what's being drawn. We go to this previous one, I have to think like, okay. There's some amount of density over here, outside the interval. How does that actually manifest, I'm not sure but in the prior predicative probability density, I can actually see I'm not getting... In comparison to all of these, there's only a handful of observations out here which might seem unrealistic or something.

Interviewer #1: Okay cool. I don't think I have any more questions about this right now.

Participant #2: Great, can we go to the next?

Interviewer #1: Yeah.

Interviewer #2: This is the last page, right?

Interviewer #1: Yeah, this is the last page. Again this takes a while to load.

Interviewer #1: Again, you can skip the first one and you can go to the second visualization.

Participant #2: Okay.

Interviewer #1: So again this is the prior parameter probability density and your response scale for the mean difference parameter.

Participant #2: Right. So now I got to pick priors, right?

Interviewer #1: I mean, again the question is, do you think this information presented in this visualization.. Based on what you used to choose the prior on the survey, How would you use this information if you used it at all.

Participant #2: Being on the scale. I'm interpreting this as the [inaudible] if there were this will mean if it's centered at one then the majority will be... 1X means people in the experimental group will pump just as many times as the people in the control group, right?

Interviewer #1: Yeah.

Participant #2: Okay, so, how would I use this? Do I find this useful?

Participant #2: Yeah, I think I might. I think it would depend on how much I know about the experimental group but it's very nice to see how much weight is given to two times, pressing it half as many times. It's nice to see. It's very tough to say, what does this actually manifest as. You know what I mean?

Interviewer #1: Interesting. Would you change the prior that you set if you saw this visualization?

Participant #2: I think...

Interviewer #1: I think you said zero and ...

Participant #2: I definitely wouldn't change the mean. I still think that the hypothesis that the intervention does nothing is still a good one. The question is, do I make it more variable or do I switch it to the student T? I think. Maybe I would change it. I think I said normal before, right?

Interviewer #1: Yes.

Participant #2: Maybe I would change it because it gives a little more credence to stuff beyond two. The optimal number, I think I remember the optimal number of pumps was something like 64. So if the average is 35, that's almost two times...

Participant #2: Depending on the intervention, it might be great, people might, I don't know. Maybe I would go with Student T after seeing this because it just gives a little more credence to stuff on this scale, I mean, down here.

Interviewer #1: Okay.

Participant #2: As compared to a normal.

Interviewer #1: Okay, the other thing I was, like I said, in the survey you choose the standard deviation of this prior as one. Right now, you basically went to the other end of the scale. You chose the least diffused prior that you could have chosen, could you elaborate on that?

Participant #2: [chose student's t(0, 0.2)] I think, if choosing standard deviation of one, now that I see that it results in very unrealistic effects. Are people really going to push it a quarter of the time? Probably not. Especially if the inter... you know what I mean? Are they going to push it? This is giving credence to 0.1, which would mean would they push it 3 times? Probably not? I think I would, this visualization definitely makes me think a bit more about how the standard deviation of the prior manifests as an effect on the natural scale.

Interviewer #1: That's interesting. Cool.

Interviewer #1: I don't have any more questions on this one. Matt, do you have any?

Interviewer #2: I think this is good. Yeah, thank you.

Participant #2: Okay.

Interviewer #1: The drop down is working here. So I don't screw up everything.

Interviewer #2: Research prototypes, it happens.

Participant #2: Yeah it happens.

Interviewer #1: Now this is the prior predictive distribution. There's a small typo. The prior on the intercept is fixed and you're just changing the prior on the mean difference parameter.

Participant #2: Right. Oh I see. The blue and the orange are different. This means that... Oh I see.

Interviewer #1: Okay. IF I remember correctly, the orange is the experimental condition and the green is the control condition.

Participant #2: Right.

Interviewer #1: And then I don't know where the labels went but there were labels. I guess again the question is, how do you think this information presented. Do you think this information would change your choice of priors or how would you use this information, if at all?

Participant #2: Right. So, like my hypothesis in this sort of experiment was that the intervention did nothing. This visualization doesn't affect where I would place the mean of the prior. It's very clear what moving in this direction would be like. But now comes the question of the standard deviation of the predictor or even the [inaudible] one. This would be minimally variable, right?

Interviewer #1: Yeah.

Participant #2: So now, in this visualization. Let me get that back to where it was. So in this visualization I can see the Student T is resulting in a lot of draws over 50.

Interviewer #1: Okay.

Participant #2: Again, that's.... well maybe that's not that bad. No, that's outside the interval, so I would probably maybe chose a normal. If I can get it. Oh, that one is getting a lot of over 50 too. Here maybe...

Participant #2: Oh right, then I'd get lots of them outside. Lots of people pressing the button lots. Don't want to allow for that. [inaudible] believe that might happen. They might press it like....

Participant #2: I think I remember the setup saying the optimal number of times is 60. Do I really think that they're going to press it like 160 times? They might learn not to so I might stick at being not as variable, having a thinner prior.

Interviewer #1: Okay.

Interviewer #2: So you paused there for a second, right before you came to that. I wonder what were you thinking about and how you came to that conclusion?

Participant #2: I was just trying to interpret what the visualization was really showing me. You saw that I dragged this up and down, making the how having a more variable prior manifested in the prior predictive probability distribution. Having a very variable prior down here resulted in, you know, people hitting the button 120 times, 100 times, much more than would be optimal. Under the assumption that people would learn the optimal strategy in this design. That they would not be pressing it 120 times as optimal, maybe they would press it 120 times maybe once or maybe they'd only press it like 80 times and then they would learn. What I'm saying is, I was playing with this, I took some time to realize how increasing the prior manifested in, what that meant for the prior predictive, that's what I was really trying to [inaudible]

Interviewer #2: Right.

Participant #2: Does that answer your question?

Interviewer #2: Yeah, that's perfect, that makes a lot of sense. Thank you.

Participant #2: Have I answered whatever you wanted to know about this one?

Interviewer #1: I think so yeah. This is great.

Participant #2: [inaudible] Go on.

Interviewer #1: There's nothing next. Don't click on the next page, it will just give you an error. I think that should have been removed, but anyways. I guess the other questions we have is, Is there other information that you consider when choosing priors in general, not for this specific example?

Participant #2: Is there other information I consider? To be honest, I'm really new to [inaudible] inference so let me think here about the handful of times I've done it.

Participant #2: I think the other thing I consider is how much data am I going to have. If I only have N data points then my priors really really matter but if I have 10,000 then my priors will likely be dominated by the likelihood assuming the model structure, is like this one, it's relatively simple and not super hierarchical.

Participant #2: I think I would be more comfortable choosing more variable and weakly informative priors if I knew that I'd have a lot of data but if I knew that I wasn't going to have a lot of data then I might pay more attention to what priors I'm choosing and chose a little more carefully. Does that make sense?

Interviewer #1: It's interesting you mentioned a hierarchical model. How do you think your decision changes if you are doing a more complicated, like a hierarchical model or something?

Participant #2: How does my decision change about the data in particular?

Interviewer #1: When you go about choosing priors.

Participant #2: When I go about choosing priors for a hierarchical model, I think when the model is really hierarchical, the choice of priors and how you choose them is really important because high dimensional space is weird. When you enter high dimension space, really unintuitive things happen. If I knew that my model was going to be very big and very hierarchical, then I'd have to think really really careful about my priors.