

## **P10**

### **Comprehensibility**

When I said "good luck tomorrow" I didn't want it to end but the bot ended the conversation.

### **Usefulness**

I get more feedback about how the child felt, and this table would be useful on improving certain points, like I guess I need 100% for everything. So when I talk to the next child I can talk about the belief that was 0% from the previous interaction.

### **Points of Improvement**

I think it would have been nicer as a table (the transcript), with the first column being the Kt or child, second the message and third the overtuiging. It would make it way more readable.

I would say maybe applying another model, because now this 5 phase model might not be perfect for every situation. So some other model might be better suited for another situation. This model is focused on completing every phase, meanwhile the child might just want the bullying to stop and he can jump to any phase. So having a model that revolves around problem solving. Like how well did you solve the problem or something. Also I asked it "why are you sad" and it's intention changed to "Lilo wants to talk about it's problems" even tho it happened because I asked it to, which doesn't seem correct. Also it asked "What do you mean" when I asked why are you sad.

## **P9**

### **Comprehensibility**

I'm not sure why this one jumped to 30%, don't know what kind of response to give it so... I think the understanding part is due to the bot not fully understanding. What does this belief mean? "Ik denk dat de KT vraagt naar een positive wens". Oh, if I say, let's find a solution together, it says that the help line worker cannot solve the problem.

### **Usefulness**

I understand that some input triggers positive and negative responses but if you're someone who doesn't understand that then should you be told that the changes you're making are only two fold? Like boolean increase or decrease. But overall, this feedback would help me because I now know how the bot works so you can work around with it to pass by the next phase quickly. With a real person then it would be more difficult. Like how do you define these phases with a real person? Like phases can overlap, phase 4 can be before phase 2 for example. So how do you know that it's always gonna be stage 1 and going linearly to 5. So phases can be a bit more mixed and swapped.

### **Points of Improvement**

Could be nice to also have a graph instead of the percentages, like overtime from start to finish. Because if you're talking for a long time and the belief change increases or decreases slightly then the graph would be better. If I say let's find a solution together, it says that the help line worker cannot solve the problem. I don't know why.

## **P8**

### **Comprehensibility**

I don't understand why at first when I ask is there someone that you trust she leaves the conversation.

### **Usefulness**

So the scores with percentages are quite useful cuz it tells you which areas you're lacking and which areas you can improve upon. Also the fact that it's differentiated per phase you can see which phases you can focus more on and which you're adequately touched upon. In the transcript it's useful to see how each prompt affect Lilobot but it does seem a little barebone.

### **Points of Improvement**

I'm getting a very analytical overview of the conversation but I think what's missing are suggestions on how to improve and alternatives of what to say during each phase and maybe the style of your responses and prompts during each phase, how that should change. For example in phase 1 you require more empathy and in the other phases you need to be more pragmatic. So far its hard to distinguish between each phase. So for example if it shows a low score for one of the phases or one of the things Lilo is convinced about maybe suggest a response on what to say with that. So you can get an idea of which phase you're lacking in, but you're not getting suggestions on how to improve upon it.

## **P7**

### **Comprehensibility**

The transcript is a bit hard to read, so having them in blocks or u coded would make it clear to separate them. Also the order of sentences shown, and then the belief, makes it clear to make the distinction.

### **Usefulness**

I guess I know if an approach worked cuz I see a bunch of upward arrows, so at least my approach wasn't terrible. The first time around I was really thinking for what to say and now I which beliefs I need to tackle first, so I can repeat that.

### **Points of Improvement**

Maybe a sidebar with pre-made sentences that have been proven to work in the past. So besides the phrasing problem, using these proven working approaches I guess. The transcript is a bit hard to read, so having them in blocks or u coded would make it clear to separate them. Also the order of sentences shown, and then the belief, makes it clear to make the distinction.

## **P6**

### **Comprehensibility**

I can see that it felt safe with me the whole time. It felt like I understood it more at the end. I think that the belief is tied to a particular phase, say I believe I have told my story is in phase 3, which makes sense looking at the 5 phase model. Why does it go down (Ik denk dat de KT mijn probleem kan oplossen) when I say that we can find a solution together? The intention is probably the next step he wants to go.

### **Usefulness**

With Lilobot yes, because I can learn how to manipulate it. It's not a particularly smart chatbot. But with an actual child I think there would be way more uncertainty and emotion when talking.

### **Points of Improvement**

Well it would help to have a dictionary and guidelines on how to use this system. Often I have the right intention of the message but then I didn't phrase it the right way so it leaves the conversation or it doesn't respond. I understand most of this feedback but how the bot got to that conclusion no. For example, why does me saying hello instantly make me more trustworthy?

## **P5**

### **Comprehensibility**

The table is quite clear. Also the instruction in the beginning. Intention was mistaken to be intensity, not the action the child is gonna take.

### **Usefulness**

The table is quite practical, cuz it directs you to the steps you have to take (the phases). You can navigate towards what you need to say to change the bot's beliefs. The transcript is not that useful I think, I wouldn't know what to do with the change in intention or the belief.

### **Points of Improvement**

I didn't know how to connect the table and the intentions. But it's a human interaction so it would be better to talk to a human about this, so going through the transcript with someone who is experienced with De Kindertelefoon, so you get idea from their experienced perspective. Usually text is not the best way to get a message across, we experience this with daily texting also, so real human involvement would be good.

## **P4**

### **Comprehensibility**

Started with transcript first. Didn't read the guiding text. It says who they are and they believe they told their story. That doesn't sound right. So in this regard, the intent and the beliefs aren't really accurate. I understand the table and that certain beliefs are tied to a certain phase.

### **Usefulness**

I like that it gives feedback with every message you give it. You see if it believes you more or less. My conversation went pretty bad, on multiple occasions it just said goodbye because it didn't trust me. Getting the feedback with the very short transcripts, I didn't get a sense of where I could improve and what I could say next time. Hence with that little information there's not much I can do to learn.

### **Points of Improvement**

Would be useful to know when the bot doesn't understand a message as opposed to taking time to reply. Also the phases don't have to be tightly tied to a phase, they can be intertwined. Thinking you can help it is only applicable to phase 3, but why not 1 or 2. This feels a bit unnatural, phases can be a bit more intertwined.

## **P3**

### **Comprehensibility**

The first paragraph that explains it all is a lot of text. This can be left out maybe and add an info button for when you need help. The phases are numbered so I wouldn't know what they are. The table is clear. Surprising is that Lilobot thinks I can't solve the problem in the end, 0%. I like the transcript that it shows the arrows and how it's influencing the bot. So you know this statement was good and this one horrible. The content is definitely fine overall.

### **Usefulness**

I feel better prepared from the overtuiging, seeing how the beliefs change is useful. But you sort of just know how to work with the bot. So giving this response makes it trust me less so I will avoid that.

### **Points of Improvement**

One thing would be to say why lilobot asks or says a specific thing. As in sometimes it throws in random piece of info without you collaborating and maybe knowing the intention behind those would be helpful. So for example why does it smile and what does that mean.

## **P2**

### **Comprehensibility**

This looks so long and a lot of numbers. The table is kinda too much text and numbers thrown at you. It's cool to see how the bot felt in the beginning and in the end. It only thought I understood it 20%, but I understood it way more. The table looks interesting to look at. Oh you can see what the bot is thinking in a way with every prompt. To me it wasn't giving the responses that I thought it would give but according to it, it seems like he said his entire story. The opinions that it has are very repetitive, even tho it's given one line of the story.

Feedback could be easier to read, it took me a solid minute to understand what it said. So like these percentages could be shown as a graph. Or if stuff was highlighted in different colours it would have been easier to read. For example if value is "I don't trust" KT and it's low then turn it red.

### **Usefulness**

Yes cuz it's useful to see what it thinks and what exactly I said to make it trust me more so I know how to work around it in the future as well. And if you see that some prompt of yours made it trust you less you learn to avoid saying it again.

### **Points of Improvement**

I guess in the table you can see which phase its in the discussion. So it would be useful to see which phase you're in with each message also, to see how fast/slow you're progressing or if you're staying in one phase for too long. But I also don't know if 70% is a good or bad thing.

## **P1**

### **Comprehensibility**

### **Usefulness**

The table is quite nice cause you can see which beliefs are touched on the most and you can see which one you need to tackle next. The transcript is a bit hard to follow. I'm not sure how I can link it with the table since the table only has beginning and end percentages of the beliefs.

### **Points of Improvement**

The phases don't really make sense to me. Interacting with the kid just seems very mechanical. Like I don't think that jumping from one phase to another is gonna be the same with a real child.

The transcript is a bit hard to comprehend at first. I can't easily distinguish which text is from who and when the belief value is shown.