

## Appendix A: Pre-Test

### Pre-test

We conducted a pre-test among 74 participants to test the perceived level of personalization of the conversational agent's recommendations, as well to assess the sensitivity level of a battery of 24 sleep-related and sleep-unrelated questions to identify a variety of questions the conversational agent can ask to prompt self-disclosure (see Measurements in manuscript). Participants in the pre-test were on average 34 years old ( $M = 34.03$ ,  $SD = 8.93$ , range = 21 – 61), predominantly resided in the United States ( $n = 61$ , 82.4%), and 66 percent of the participants were male ( $n = 49$ , 66.2%). About half of the participants completed a middle level of education (i.e., some college, no degree; Bachelor's degree in college, 3/4 years;  $n = 38$ , 51.4%), followed by a higher level of education (i.e., Master's degree; doctoral degree;  $n = 28$ , 37.8%), and a small group completed a lower level of education (i.e., less than a high school diploma; high school diploma or equivalent;  $n = 8$ , 10.8%). Participants were, on average, quite interested in their health ( $M = 5.82$ ,  $SD = 1.24$ ) and sleep ( $M = 5.58$ ,  $SD = 1.45$ ).

The pre-test consisted of (1) a conversation with the conversational agent Robin and (2) an online survey. The procedure of the pre-test was similar to that of the main experiment (see Procedure in manuscript). With regard to assessing perceived personalization of the conversational agent's recommendations, we used three items asking whether conversational agent Robin personalized the conversation based on participants' personal information (generally and related to sleeping) and whether it targeted participants as a unique individual (measured on a 7-point Likert scale, see Measurements). Participants who were assigned to the personalized condition experienced the conversation with the conversational agent as significantly more

personalized ( $M = 4.89$ ,  $SD = 1.59$ ) than participants assigned to the non-personalized condition ( $M = 3.37$ ,  $SD = 1.70$ ),  $t(72) = -3.96$ ,  $p < .001$ , 95% CI [-2.29, -0.75].

With regard to identifying sleep-related and -unrelated questions for the conversational agent to ask during the main experiment, we provided participants with a battery of 24 questions: 12 sleep-related and 12 sleep-unrelated, with half of each expected to be rated as sensitive and half as non-sensitive. Example questions are: “Can you describe a personal experience that made you talk or consider talking with a doctor, therapist or sleep professional about your sleep? (sleep-related, sensitive),” “What causes you to feel tired when you wake up? (sleep-related, non-sensitive),” “What have you done in your life that you feel most guilty about? (sleep-unrelated, sensitive),” and “What are your favorite things to do in your free time? (sleep-unrelated, non-sensitive)” (a complete list of questions can be found in Table A1). Sleep-related questions were based on several non-academic sources, such as the Sleep foundation (Suni, 2020). Sleep-unrelated questions were based on both academic and non-academic sources. Non-sensitive questions were based on websites with conversation starters (see Conversation Starters World, n.d.) and sensitive questions were based on prior research on intimate exchanges (Moon, 2000). All questions were open-ended questions.

To assess the perceived sensitivity of the questions, participants were asked to imagine that a conversational agent would ask them these questions. They were asked to indicate on 7-point semantic differential scales the extent to which they felt disclosing information to a conversational agent regarding each question would be very uncomfortable (1) to very comfortable (7) and whether this presented a low risk (1) to a high-risk (7). In addition, they indicated whether this information was not at all intimate (1) to very intimate (7) and not at all sensitive (1) to very sensitive (7) to them. Finally, participants indicated how very unwilling (1)

to very willing (7) they were to disclose information to a chatbot regarding each question. These five items represented sensitivity (Mothersbaugh et al., 2012), and were computed into mean scales. Based on their mean scores, three sleep-related sensitive questions and three sleep-related non-sensitive questions, as well as three sleep-unrelated sensitive questions and three sleep-unrelated non-sensitive questions were selected for the conversation content of the chatbot in the main experiment.

## References

- Conversation Starters World. (n.d.). *250 conversation starters*. Retrieved on July 19, 2021, from <https://conversationstartersworld.com/250-conversation-starters/>
- Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research*, *26*(4), 323-339. <https://doi.org/10.1086/209566>
- Mothersbaugh, D. L., Foxx, W. K., Beatty, S. E., & Wang, S. (2012). Disclosure antecedents in an online service context: The role of sensitivity of information. *Journal of Service Research*, *15*(1), 76–98. <https://doi.org/10.1177/1094670511424924>
- Suni, E. (2020). *Healthy sleep tips*. Retrieved on July 19, 2021, from <https://www.sleepfoundation.org/sleep-hygiene/healthy-sleep-tips>

Table A1

*List of Sleep-Related and Sleep-Unrelated, Sensitive and Non-Sensitive Questions*

Question	<i>M</i>	<i>SD</i>	alpha
<i>Sleep-related, sensitive questions</i>			
<b>Q1.</b> Can you describe a personal experience that made you talk or consider talking with a doctor, therapist or sleep professional about your sleep?	<b>4.19</b>	1.77	.94
<b>Q2.</b> In what ways does your (lack of) sleep affect your performance at school or work?	2.56	1.27	.91
<b>Q3.</b> When was the last time you were too stressed to sleep? Can you tell me about the source of the stress?	<b>3.58</b>	1.61	.94
<b>Q4.</b> In what ways does your lifestyle, such as alcohol use or lack of physical activity, cause difficulties with sleeping?	3.09	1.48	.93
<b>Q5.</b> Can you describe your worst nightmare that causes you trouble sleeping?	<b>3.67</b>	1.75	.94
<b>Q6.</b> What do you think about using medication to help you sleep?	2.85	1.59	.93
<i>Sleep-related, non-sensitive questions</i>			
<b>Q7.</b> In which situations do you have trouble falling asleep?	2.40	1.25	.94
<b>Q8.</b> For what reasons do you take naps during the day?	<b>2.28</b>	1.32	.94

Question	<i>M</i>	<i>SD</i>	alpha
<b>Q9.</b> Can you describe your schedule for going to bed every night and waking up every morning?	2.51	1.27	.91
<b>Q10.</b> In which ways do you make up for lost sleep during the week?	<b>2.15</b>	1.14	.89
<b>Q11.</b> Under which circumstances do you sleep without waking up at night?	2.40	1.38	.93
<b>Q12.</b> What causes you to feel tired when you wake up?	<b>2.13</b>	1.19	.91
<i>Sleep-unrelated, sensitive questions</i>			
<b>Q13.</b> What do you dislike about your physical appearance?	4.05	1.53	.88
<b>Q14.</b> What have you done in your life that you feel most guilty about?	<b>5.01</b>	1.72	.93
<b>Q15.</b> What are some of the things that really hurt your feelings?	3.86	1.57	.93
<b>Q16.</b> What has been the biggest disappointment in your life?	<b>4.51</b>	1.74	.91
<b>Q17.</b> What are some of the things that make you furious?	2.94	1.43	.92
<b>Q18.</b> What are some of the things you hate about yourself?	<b>4.46</b>	1.67	.91
<i>Sleep-unrelated, non-sensitive questions</i>			
<b>Q19.</b> What do you do when you hang out with your friends?	2.54	1.37	.93

Question	<i>M</i>	<i>SD</i>	alpha
<b>Q20.</b> What do you do to improve your mood when you are in a bad mood?	2.59	1.36	.90
<b>Q21.</b> What is the best thing about your school or work?	<b>2.34</b>	1.29	.88
<b>Q22.</b> What did you do on your last vacation?	<b>2.31</b>	1.33	.90
<b>Q23.</b> What is the most annoying habit someone can have?	2.36	1.39	.93
<b>Q24.</b> What are your favorite things to do in your free time?	<b>2.17</b>	1.18	.89

*Note.* For sensitive questions, the three most sensitive questions based on the mean scores are presented in bold. For non-sensitive questions, the three least sensitive questions based on the mean scores are presented in bold. Based on  $N = 59$  due to missing data.