

Users' needs for a digital smoking cessation application and how to address them: A mixed-methods study

Overview of Steps to Reproduce Analyses

Author: Nele Albers

Date: July 2022

This file is meant to give you an overview of the steps needed to reproduce our analyses from the paper "Users' needs for a digital smoking cessation application and how to address them: A mixed-methods study" by Nele Albers, Mark A. Neerincx, Kristell M. Penforinis, and Willem-Paul Brinkman. All code was written and run by the first author Nele Albers. The fourth author, Willem-Paul Brinkman, checked the code as shown in the file "ChecklistDataRepositoryReview.pdf."

Types of Analyses

Below we describe how you can reproduce our results for the different types of analyses.

Acceptance of the Virtual Coach

In the "Materials"-section of the paper, we report the acceptance of the virtual coach Sam based on the public dataset that can be found here:

<https://doi.org/10.4121/19934783.v1>. Navigate to the "Acceptance"-folder to reproduce the reported acceptance. There is a README-file in the folder with further instructions.

Bayesian Power Analysis

In the "Participants"-section of the paper, we report results from a Bayesian power analysis for the quantitative part of our analysis. Navigate to the "Sample_Size"-folder to reproduce the reported results. The folder contains a README-file with more information.

Data and Preprocessing

The folder "Data" contains the data we used for our analysis. In case the provided data is already preprocessed, we provide the corresponding preprocessing code:

- Have a look at the file "Data/preprocess_activity_experiences.py" to see how we preprocessed the anonymized session data before coding.
- Have a look at the file "Data/preprocess_post-questionnaire.py" to see how we preprocessed the anonymized post-questionnaire data before coding.

The "Data"-folder contains a README-file with further information on the provided data.

Coding Reliability

Run the file "calculate_agreement_between_coders.py" to:

- reproduce our agreement between the two coders reported in the "Materials & Methods"-section of the paper. You can find the values as computed by us in the file "calculate_agreement_between_coders_output.txt."

Cronbach's Alpha Values and Processing of Coded Data

Run the file "process_after_coding_for_analysis.py" to:

- recreate the file "all_coded_data.csv" that forms the basis of our subsequently described analyses. This file also contains the consecutive participant numbers that we use to refer to participants in our paper. Note that we have already run this code, so the data file already exists;
- reproduce our reported Cronbach's alpha values for the three identity measures reported in the "Materials & Methods"-section of the paper. You can find the values as computed by us in the file "process_after_coding_for_analysis_output.txt."

Participant Characteristics and Interaction Scenario Ratings

Navigate to the folder "Scenario_Ratings_and_Participant_Characteristics" to reproduce our reported participant characteristics from the "Participants"-section and the supplemental information as well as the interaction scenario ratings from Figure 3. The folder contains a README-file that provides further information.

Code Frequencies

Navigate to the folder "Code_frequencies" to reproduce Figure 2 and Figure S9 that show the code frequencies. The folder contains a README-file that provides further information.

Correlations between User Characteristics and Activity Efforts and Interaction Scenario Ratings

Navigate to the folder "Correlations" to reproduce our findings on the correlations between user characteristics on the one hand and activity efforts and interaction scenario group ratings on the other hand (Figure 4 and Table 1). The folder contains a README-file that provides further information.

Explanation of Files

This folder contains the following folders and files:

- Acceptance: Code for reproducing our reported acceptance of the virtual coach from the "Materials"-section of the paper.
- Code_Frequencies: Code for reproducing our Figure 2 and Figure S9 that show the code frequencies.
- Correlations: Code for reproducing our Figure 4 and Table 1 with correlations between user characteristics on the one hand and activity efforts and interaction scenario group ratings on the other hand.
- Data: Data files, including the codes assigned to free-text responses by the two coders. Also includes the code for preprocessing the anonymized session and post-questionnaire data before coding.
- Sample_Size: Code for reproducing our reported Bayesian power analysis results from the "Participants"-section of the paper.
- Scenario_Ratings_and_Participant_Characteristics: Code for reproducing our Figure 3 and Table S6 from the paper.

- `calculate_agreement_between_coders.py`: Code for computing the agreement between the two coders.
- `calculate_agreement_between_coders_output.txt`: Output from the above file when run by us.
- `ChecklistDataRepositoryReview.pdf`: Data and analysis code review checklist as filled out by the fourth author, Willem-Paul Brinkman.
- `process_after_coding_for_analysis.py`: Code for processing the data after coding for all analyses from the "Results"-section.
- `process_after_coding_for_analysis_output.txt`: Output of the above file when run by us.
- `README.md/README.pdf`: This README-file.