

Cultural Differences and Similarities in Perceptions of Artificial Social Agents (ASAs) - German version

Summative round

Author: Boleslav Khodakov **Date:** June 2023

Contributors: Bolek Khodakov, Emma Bokel, Nele Albers, Andrea Bönsch, Jonathan Ehret, and Willem-Paul Brinkman.

OSF form for this part: <https://osf.io/g3729> OSF form for previous part: <https://osf.io/adknw>

This codebase is based on code from: Fengxiang Li

The raw data (with legends) can be obtained from: <https://osf.io/g3729> in the folder "Translation_German".

This file is meant to guide you through the steps to go through the transformation (e.g. ICC values for items/constructs) of raw data for the summative round.

Due to time constraints, the data transformed and analyzed in these files is only based on the first data collected between the 19th and 22nd of June. We are still aiming for 120 participants per questionnaire part for future work. In the first group we collected from 72 bilingual participants with German primary tongue who are fluent German and fluent English speakers. In the second group, human-ASA interaction evaluation data of the last 46 items were collected from 82 bilingual participants with German primary tongue who are fluent German and fluent English speakers. The amount was artificially reduced to 72, since this codebase requires the two halves to have an equal amount of rows. Currently, data of 144 (72 per each half) participants is analyzed.

Out of 144 participants (72 per questionnaire), 141 recommended using their data. We use *all* participants' (144) data for the evaluation, however.

Explanation of what was done

The files in this section were used to calculate item-level ICC values, construct-level ICC values, variation between English and German ASA questionnaires, and to compare human-ASA interactions between different Cultural Backgrounds. Data was acquired from the survey platform Qualtrics, while recruiting participants for the study was done via Prolific. Two surveys

were created (to reduce fatigue of participants) - the first and second halves of the ASA questionnaire items, with their respectful translations. Data from these surveys was first pre-processed. Herewith, the data was firstly redundantly checked (e.g. for passing attention checks), then unnecessary columns (e.g. attention checks) were removed, leaving only columns of question items (English and German). The question items were of both halves were combined, and their analysis was done. A culture file was also created during pre-processing and analyzed during analysis.

There were 14 attention checks (in each of the two surveys), which participants had to pass for their data to be valid. 7 are in English, 7 are in German. In the `Legend.txt`, items can be referred by both their names. Attention checks can also be referenced by their names in the legend file. In the survey, we asked participants on whether they recommend using their data for research purposes. We use *all* participants' data for the evaluation, however.

Examine Output of Code Run by Us

To see the transformation of qualtrics data into the format used for ICC calculation, access the files `Data Transformation Summative first half (.pdf or .rmd)` and `Data Transformation Summative second half (.pdf or .rmd)`. To see the analysis, access the file `ICC calculation Summative (.pdf or .rmd)` in a different folder.

Run Code Yourself

This section is to explain how you can run the code yourself.

Requirements

You need to have the R programming language installed. Version 4.2.3 was used for this study. We recommend using RStudio to view the `.rmd` files. The `.rmd` files contain code chunks which list/install the necessary dependencies. Some of these are commented out (else they install the dependency each time).

Steps to Reproduce Analyses

The reproduction of our code is based on RStudio. Take the following steps:

1. Make sure you have R installed.
2. Make sure you have obtained the raw data from: <https://osf.io/g3729> in the folder "Translation_German".
3. If you wish to arbitrarily remove 10 rows from the second-half data yourself, run the file `data_equalization.rmd`. Else, change the code to delete the exact rows specified in `data_equalization.pdf`. You will overwrite this file, if you run `data_equalization.rmd`.

4. Open the file `Data Transformation Summative first half.rmd` . Run the file's contents. The file accesses the file `Final_ASA_German_Summative_First_Half_v1_2023_06_22_anonym.sav` .
5. Open the file `Data Transformation Summative second half.rmd` . Run the file's contents. The file accesses the file `Final_ASA_German_Summative_Second_Half_v1_2023_06_22_anonym_reduced.sav` .
6. In the directory where the `.rmd` files were run, 2 new files should appear. By default, they will be named `summative_first_half_transformed.sav` , `summative_second_half_transformed.sav` . This is the preprocessed data.
7. Open and run the file `Transformation Culture.Rmd` . This will merge the German-English and mixed-English cultures into one file. The file should access `summative_first_half_transformed.sav` and `summative_second_half_transformed.sav` and `data_culture.sav` .

A note on knitting R Markdown

Knitting the `.rmd` files into `.pdf` files requires `.tex` and `.bib` files. These are present in the current directory. Data files needed for the `.rmd` code chunks are also present in this directory.

You may want to install dependencies present in the first code chunk of the `.rmd` file you wish to knit. If you have these installed, pressing the `knit` button in Rstudio should knit your file to PDF.

Explanation of Files

Please note: The `.rmd` files contain extensive comments.

This directory contains (or references) the following files and folders:

- `README.md` and `README.pdf` : Short explanations of how to use files in current folder, and what those files are.
- `data_equalization.rmd/pdf` : File for removing 10 arbitrary rows from second-half survey.
- `Transformation Culture.Rmd` : File for merging German-English answers with mixed-English answers to compare cultures.
- `Final_ASA_German_Summative_First_Half_v1_2023_06_22_anonym.sav` : Qualtrics data for the "first half" survey. This survey contains the first 44 English items of the ASA questionnaire and their translations.
- `Final_ASA_German_Summative_Second_Half_v1_2023_06_22_anonym.sav` : Qualtrics data for the "second half" survey. This survey contains the last 46 English items of the ASA questionnaire and their translations.

- `Data Transformation Summative first half.rmd/pdf` : Pre-processing of the Qualtrics data into usable data for calculation for the "first half" survey.
- `Data Transformation Summative second half.rmd/pdf` : Pre-processing of the Qualtrics data into usable data for calculation for the "second half" survey
- `summative_first_half_transformed.sav` : Pre-processed data for the "first half" survey. Based on data of all participants. Ready for calculations.
- `summative_second_half_transformed.sav` : Pre-processed data for the "second half" survey. Based on data of all participants. Ready for calculations.
- `Legend.txt` : Legend for all columns of the questionnaire.
- `prolific_export_final_eval_german_first_half_v1_2023_06_22_anonym.csv` : General data on participants for the "first half" survey obtained from Prolific.
- `prolific_export_final_eval_german_second_half_v1_2023_06_22_anonym.csv` : General data on participants for the "first half" survey obtained from Prolific.
- `data_culture_EN_DE.sav` : German-English, mixed-English culture file.
- `data_culture.sav` : Original culture file, containing mixed-English culture.
- `Prolific Data.rmd/pdf` : Statistical analysis of participants' data.
- `Prolific Legend.txt` : Legend for all columns of Prolific files.