German and Dutch Translations of the Artificial-Social-Agent Questionnaire Instrument for Evaluating Human-Agent Interactions

Creation of culture file

Nele Albers

18 June, 2024

Contents

Load packages **1**

Preprocessing the English ratings data from German sample **2**

Preprocessing the data from mixed-international English sample **4**

Combine data from mixed-international English sample and German sample **7**

Add Dutch data **8**
Preprocessing the English ratings data from Dutch sample 8
Combine Dutch data with the remaining ratings 10

Save combined file **11**

References **11**

In this file, we create the culture data file based on our transformed data as well as the mixed-international English data published by Fitrianie et al. (2024).

Required files: Data/ASAQ_ConstructValidity_p532.xlsx, summative_first_half_transformed_dutch.sav, summative_second_half_transformed_dutch.sav, summative_first_half_transformed_german.sav, summative_second_half_transformed_german.sav.

The file `Data/ASAQ_ConstructValidity_p532.xlsx` is to be downloaded from the repository accompanying the paper by Fitrianie et al. (2022), which can be found here: https://doi.org/10.4121/19758436.V6. Specifically, this is the file `ASAQ_ConstructValidity_p532.xlsx` from that repository, saved to our "Data"-folder.

Created files: data_culture_DE_EN_NL.sav

Authored by Nele Albers, Andrea Bönsch, Jonathan Ehret, Boleslav A. Khodakov, and Willem-Paul Brinkman.

Load packages

First, we load the packages that we need.

```
library(dplyr)    # Use select function
library(formatR)  # For formatting
library(haven)    # Use read_sav function
```

```
library(reshape2) # Use melt function
library(xlsx)     # Load excel-file
```

Preprocessing the English ratings data from German sample

First we preprocess the English ratings data from our study with people who had German as primary language (i.e., the data where the people which German as their primary language rated the ASAQ items in English). Therefore, we calculate scores for each English construct/dimension per participant, by averaging the items' scores. And we add a new column 'ConstructID' to represent each construct/dimension, facilitating further analysis of the comparison between different cultural backgrounds when using the English version of the ASAQ.

We begin with the first half of the questionnaire.

```
d1_bilingual <- read_sav("summative_first_half_transformed_german.sav")


# Remove Q_E_ prefix from English items
for (col in 8:51) {
    colnames(d1_bilingual)[col] <- sub("Q_E_", "", colnames(d1_bilingual)[col])
}


# Select English item scores of HLA1, HLA2, HLA3 and HLA4
# for Dimension 'HLA'
HLA_EN <- data.frame(select(d1_bilingual, HLA1:HLA4))
# Select English item scores of each construct/dimension
# for Construct 1-8
HLB_EN <- data.frame(select(d1_bilingual, HLB1:HLB5))
NA_EN <- data.frame(select(d1_bilingual, NA1:NA5))
NB_EN <- data.frame(select(d1_bilingual, NB1:NB3))
AAS_EN <- data.frame(select(d1_bilingual, AAS1:AAS3))
AU_EN <- data.frame(select(d1_bilingual, AU1:AU3))
PF_EN <- data.frame(select(d1_bilingual, PF1:PF3))
AL_EN <- data.frame(select(d1_bilingual, AL1:AL5))
AS_EN <- data.frame(select(d1_bilingual, AS1:AS3))
APP_EN <- data.frame(select(d1_bilingual, APP1:APP3))
UAA_EN <- data.frame(select(d1_bilingual, UAA1:R_UAA3))
AE_EN <- data.frame(select(d1_bilingual, R_AE1:R_AE4))


# Calculate English item score mean of HLA1, HLA2, HLA3 and
# HLA4 as construct score
d1_bilingual$HLA_EN <- rowMeans(HLA_EN)
# Calculate English mean of each construct/dimension per
# participant
d1_bilingual$HLB_EN <- rowMeans(HLB_EN)
d1_bilingual$NA_EN <- rowMeans(NA_EN)
d1_bilingual$NB_EN <- rowMeans(NB_EN)
d1_bilingual$AAS_EN <- rowMeans(AAS_EN)
d1_bilingual$AU_EN <- rowMeans(AU_EN)
d1_bilingual$PF_EN <- rowMeans(PF_EN)
d1_bilingual$AL_EN <- rowMeans(AL_EN)
d1_bilingual$AS_EN <- rowMeans(AS_EN)
```

```
d1_bilingual$APP_EN <- rowMeans(APP_EN)
d1_bilingual$UAA_EN <- rowMeans(UAA_EN)
d1_bilingual$AE_EN <- rowMeans(AE_EN)


# Transform data into the long format
bilingualdata1 <- melt(select(d1_bilingual, agentName:AE_EN),
    id.vars = c("agentName", "AgentID"), variable.name = "Construct",
    value.name = "Rating")


# Add a column 'Culture', 0 is participant group with
# bilingual German cultural background
bilingualdata1$Culture <- 0


# Label ConstructID for English constructs/dimensions
bilingualdata1$ConstructID[bilingualdata1$Construct == "HLA_EN"] <- 1
bilingualdata1$ConstructID[bilingualdata1$Construct == "HLB_EN"] <- 2
bilingualdata1$ConstructID[bilingualdata1$Construct == "NA_EN"] <- 3
bilingualdata1$ConstructID[bilingualdata1$Construct == "NB_EN"] <- 4
bilingualdata1$ConstructID[bilingualdata1$Construct == "AAS_EN"] <- 5
bilingualdata1$ConstructID[bilingualdata1$Construct == "AU_EN"] <- 6
bilingualdata1$ConstructID[bilingualdata1$Construct == "PF_EN"] <- 7
bilingualdata1$ConstructID[bilingualdata1$Construct == "AL_EN"] <- 8
bilingualdata1$ConstructID[bilingualdata1$Construct == "AS_EN"] <- 9
bilingualdata1$ConstructID[bilingualdata1$Construct == "APP_EN"] <- 10
bilingualdata1$ConstructID[bilingualdata1$Construct == "UAA_EN"] <- 11
bilingualdata1$ConstructID[bilingualdata1$Construct == "AE_EN"] <- 12
```

We now perform the same preprocessing for the second half of the questionnaire.

```
d2_bilingual <- read_sav("summative_second_half_transformed_german.sav")


# Remove Q_E_ prefix from English items
for (col in 8:53) {
    colnames(d2_bilingual)[col] <- sub("Q_E_", "", colnames(d2_bilingual)[col])
}


# Select English item scores of UE1, UE2 and UE3 for
# Construct 'UE'
UE_EN <- data.frame(select(d2_bilingual, UE1:UE3))
# Select English scores of the remaining 11
# constructs/dimensions
UT_EN <- data.frame(select(d2_bilingual, UT1:UT3))
UAL_EN <- data.frame(select(d2_bilingual, UAL1:UAL6))
AA_EN <- data.frame(select(d2_bilingual, AA1:AA3))
AC_EN <- data.frame(select(d2_bilingual, R_AC1:R_AC4))
AI_EN <- data.frame(select(d2_bilingual, AI1:AI4))
AT_EN <- data.frame(select(d2_bilingual, AT1:R_AT3))
SP_EN <- data.frame(select(d2_bilingual, SP1:SP3))
IIS_EN <- data.frame(select(d2_bilingual, IIS1:IIS4))
```

```r
AEI_EN <- data.frame(select(d2_bilingual, AEI1:R_AEI5))
UEP_EN <- data.frame(select(d2_bilingual, UEP1:UEP4))
UAI_EN <- data.frame(select(d2_bilingual, UAI1:UAI4))


# Calculate English item score mean of UE1, UE2 and UE3 as
# construct score
d2_bilingual$UE_EN <- rowMeans(UE_EN)
# Calculate English mean of each construct/dimension per
# participant
d2_bilingual$UT_EN <- rowMeans(UT_EN)
d2_bilingual$UAL_EN <- rowMeans(UAL_EN)
d2_bilingual$AA_EN <- rowMeans(AA_EN)
d2_bilingual$AC_EN <- rowMeans(AC_EN)
d2_bilingual$AI_EN <- rowMeans(AI_EN)
d2_bilingual$AT_EN <- rowMeans(AT_EN)
d2_bilingual$SP_EN <- rowMeans(SP_EN)
d2_bilingual$IIS_EN <- rowMeans(IIS_EN)
d2_bilingual$AEI_EN <- rowMeans(AEI_EN)
d2_bilingual$UEP_EN <- rowMeans(UEP_EN)
d2_bilingual$UAI_EN <- rowMeans(UAI_EN)


# Transform data into the long format
bilingualdata2 <- melt(select(d2_bilingual, agentName:UAI_EN),
    id.vars = c("agentName", "AgentID"), variable.name = "Construct",
    value.name = "Rating")


# Add a column 'Culture', 0 is participant group with
# bilingual German cultural background
bilingualdata2$Culture <- 0


# Label ConstructID for English constructs/dimensions
bilingualdata2$ConstructID[bilingualdata2$Construct == "UE_EN"] <- 13
bilingualdata2$ConstructID[bilingualdata2$Construct == "UT_EN"] <- 14
bilingualdata2$ConstructID[bilingualdata2$Construct == "UAL_EN"] <- 15
bilingualdata2$ConstructID[bilingualdata2$Construct == "AA_EN"] <- 16
bilingualdata2$ConstructID[bilingualdata2$Construct == "AC_EN"] <- 17
bilingualdata2$ConstructID[bilingualdata2$Construct == "AI_EN"] <- 18
bilingualdata2$ConstructID[bilingualdata2$Construct == "AT_EN"] <- 19
bilingualdata2$ConstructID[bilingualdata2$Construct == "SP_EN"] <- 20
bilingualdata2$ConstructID[bilingualdata2$Construct == "IIS_EN"] <- 21
bilingualdata2$ConstructID[bilingualdata2$Construct == "AEI_EN"] <- 22
bilingualdata2$ConstructID[bilingualdata2$Construct == "UEP_EN"] <- 23
bilingualdata2$ConstructID[bilingualdata2$Construct == "UAI_EN"] <- 24
```

Preprocessing the data from mixed-international English sample

Let's preprocess the data from the mixed-international English sample. For this, we need to create an agentID column, compute the construct means, transform the data into a long format, create a culture column, and create IDs for the constructs.

```r
# Load data
data_culture_mixed <- read.xlsx("Data/ASAQ_ConstructValidity_p532.xlsx",
    sheetName = "LONG VERSION")


# Rename GROUP column to agentName
names(data_culture_mixed)[names(data_culture_mixed) == "GROUP"] <- "agentName"


# Add a column 'AgentID' to facilitate analysis for
# comparison between different cultural backgrounds in the
# main markdown file No specific meaning for 14 numbers,
# just a code for each agent
data_culture_mixed$AgentID <- 0
# Label AgentID for 14 ASAs
data_culture_mixed$AgentID[data_culture_mixed$agentName == "iCAT"] <- 1
data_culture_mixed$AgentID[data_culture_mixed$agentName == "DEEPBLUE"] <- 2
data_culture_mixed$AgentID[data_culture_mixed$agentName == "AMY"] <- 3
data_culture_mixed$AgentID[data_culture_mixed$agentName == "FURBY"] <- 4
data_culture_mixed$AgentID[data_culture_mixed$agentName == "POPPIE"] <- 5
data_culture_mixed$AgentID[data_culture_mixed$agentName == "SIRI"] <- 6
data_culture_mixed$AgentID[data_culture_mixed$agentName == "HAL 9000"] <- 7
data_culture_mixed$AgentID[data_culture_mixed$agentName == "SIM SENSEI"] <- 8
data_culture_mixed$AgentID[data_culture_mixed$agentName == "CHAPPIE"] <- 9
data_culture_mixed$AgentID[data_culture_mixed$agentName == "AIBO"] <- 10
data_culture_mixed$AgentID[data_culture_mixed$agentName == "SARAH"] <- 11
data_culture_mixed$AgentID[data_culture_mixed$agentName == "NAO"] <- 12
data_culture_mixed$AgentID[data_culture_mixed$agentName == "MARCUS"] <- 13
data_culture_mixed$AgentID[data_culture_mixed$agentName == "DOG"] <- 14
attr(data_culture_mixed$agentName, "label") <- c("iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI,
HAL 9000, SIM SENSEI, CHAPPIE, AIBO, SARAH, NAO, MARCUS, DOG")
# Add label to 'AgentID'
attr(data_culture_mixed$AgentID, "label") <- c("1=iCAT, 2=DEEPBLUE, 3=AMY, 4=FURBY, 5=POPPIE, 6=SIRI,
7=HAL 9000, 8=SIM SENSEI, 9=CHAPPIE, 10=AIBO, 11=SARAH, 12=NAO, 13=MARCUS, 14=DOG")


# Select English item scores of HLA1, HLA2, HLA3 and HLA4
# for Dimension 'HLA'
HLA_EN <- data.frame(select(data_culture_mixed, HLA1:HLA4))
# Select English item scores of remaining constructs
HLB_EN <- data.frame(select(data_culture_mixed, HLB1:HLB5))
NA_EN <- data.frame(select(data_culture_mixed, NA1:NA5))
NB_EN <- data.frame(select(data_culture_mixed, NB1:NB3))
AAS_EN <- data.frame(select(data_culture_mixed, AAS1:AAS3))
AU_EN <- data.frame(select(data_culture_mixed, AU1:AU3))
PF_EN <- data.frame(select(data_culture_mixed, PF1:PF3))
AL_EN <- data.frame(select(data_culture_mixed, AL1:AL5))
AS_EN <- data.frame(select(data_culture_mixed, AS1:AS3))
APP_EN <- data.frame(select(data_culture_mixed, APP1:APP3))
UAA_EN <- data.frame(select(data_culture_mixed, UAA1:R_UAA3))
AE_EN <- data.frame(select(data_culture_mixed, R_AE1:R_AE4))
UE_EN <- data.frame(select(data_culture_mixed, UE1:UE3))
UT_EN <- data.frame(select(data_culture_mixed, UT1:UT3))
```

```
UAL_EN <- data.frame(select(data_culture_mixed, UAL1:UAL6))
AA_EN <- data.frame(select(data_culture_mixed, AA1:AA3))
AC_EN <- data.frame(select(data_culture_mixed, R_AC1:R_AC4))
AI_EN <- data.frame(select(data_culture_mixed, AI1:AI4))
AT_EN <- data.frame(select(data_culture_mixed, AT1:R_AT3))
SP_EN <- data.frame(select(data_culture_mixed, SP1:SP3))
IIS_EN <- data.frame(select(data_culture_mixed, IIS1:IIS4))
AEI_EN <- data.frame(select(data_culture_mixed, AEI1:R_AEI5))
UEP_EN <- data.frame(select(data_culture_mixed, UEP1:UEP4))
UAI_EN <- data.frame(select(data_culture_mixed, UAI1:UAI4))


# Calculate item score mean of HLA1, HLA2, HLA3 and HLA4 as
# construct score
data_culture_mixed$HLA_EN <- rowMeans(HLA_EN)
# Calculate mean of each construct/dimension per
# participant
data_culture_mixed$HLB_EN <- rowMeans(HLB_EN)
data_culture_mixed$NA_EN <- rowMeans(NA_EN)
data_culture_mixed$NB_EN <- rowMeans(NB_EN)
data_culture_mixed$AAS_EN <- rowMeans(AAS_EN)
data_culture_mixed$AU_EN <- rowMeans(AU_EN)
data_culture_mixed$PF_EN <- rowMeans(PF_EN)
data_culture_mixed$AL_EN <- rowMeans(AL_EN)
data_culture_mixed$AS_EN <- rowMeans(AS_EN)
data_culture_mixed$APP_EN <- rowMeans(APP_EN)
data_culture_mixed$UAA_EN <- rowMeans(UAA_EN)
data_culture_mixed$AE_EN <- rowMeans(AE_EN)
data_culture_mixed$UE_EN <- rowMeans(UE_EN)
data_culture_mixed$UT_EN <- rowMeans(UT_EN)
data_culture_mixed$UAL_EN <- rowMeans(UAL_EN)
data_culture_mixed$AA_EN <- rowMeans(AA_EN)
data_culture_mixed$AC_EN <- rowMeans(AC_EN)
data_culture_mixed$AI_EN <- rowMeans(AI_EN)
data_culture_mixed$AT_EN <- rowMeans(AT_EN)
data_culture_mixed$SP_EN <- rowMeans(SP_EN)
data_culture_mixed$IIS_EN <- rowMeans(IIS_EN)
data_culture_mixed$AEI_EN <- rowMeans(AEI_EN)
data_culture_mixed$UEP_EN <- rowMeans(UEP_EN)
data_culture_mixed$UAI_EN <- rowMeans(UAI_EN)


# Transform data into the long format
mixeddata <- melt(select(data_culture_mixed, agentName, AgentID:UAI_EN),
    id.vars = c("agentName", "AgentID"), variable.name = "Construct",
    value.name = "Rating")


# Add a column 'Culture', 1 is participant group with
# mixed-international English background
mixeddata$Culture <- 1
```

```
# Label ConstructID for constructs/dimensions
mixeddata$ConstructID[mixeddata$Construct == "HLA_EN"] <- 1
mixeddata$ConstructID[mixeddata$Construct == "HLB_EN"] <- 2
mixeddata$ConstructID[mixeddata$Construct == "NA_EN"] <- 3
mixeddata$ConstructID[mixeddata$Construct == "NB_EN"] <- 4
mixeddata$ConstructID[mixeddata$Construct == "AAS_EN"] <- 5
mixeddata$ConstructID[mixeddata$Construct == "AU_EN"] <- 6
mixeddata$ConstructID[mixeddata$Construct == "PF_EN"] <- 7
mixeddata$ConstructID[mixeddata$Construct == "AL_EN"] <- 8
mixeddata$ConstructID[mixeddata$Construct == "AS_EN"] <- 9
mixeddata$ConstructID[mixeddata$Construct == "APP_EN"] <- 10
mixeddata$ConstructID[mixeddata$Construct == "UAA_EN"] <- 11
mixeddata$ConstructID[mixeddata$Construct == "AE_EN"] <- 12
mixeddata$ConstructID[mixeddata$Construct == "UE_EN"] <- 13
mixeddata$ConstructID[mixeddata$Construct == "UT_EN"] <- 14
mixeddata$ConstructID[mixeddata$Construct == "UAL_EN"] <- 15
mixeddata$ConstructID[mixeddata$Construct == "AA_EN"] <- 16
mixeddata$ConstructID[mixeddata$Construct == "AC_EN"] <- 17
mixeddata$ConstructID[mixeddata$Construct == "AI_EN"] <- 18
mixeddata$ConstructID[mixeddata$Construct == "AT_EN"] <- 19
mixeddata$ConstructID[mixeddata$Construct == "SP_EN"] <- 20
mixeddata$ConstructID[mixeddata$Construct == "IIS_EN"] <- 21
mixeddata$ConstructID[mixeddata$Construct == "AEI_EN"] <- 22
mixeddata$ConstructID[mixeddata$Construct == "UEP_EN"] <- 23
mixeddata$ConstructID[mixeddata$Construct == "UAI_EN"] <- 24
```

Combine data from mixed-international English sample and German sample

Now we combine our preprocessed English ratings data from a German sample with the data from a mixed-international English sample.

```
# Remove construct column from German-English data and from
# the mixed-international English data
bilingualdata1 <- select(bilingualdata1, -Construct)
bilingualdata2 <- select(bilingualdata2, -Construct)
mixeddata <- select(mixeddata, -Construct)

# Combine our English data and the previously collected
# mixed-English data
data_culture_EN_DE <- rbind(bilingualdata1, bilingualdata2, mixeddata)

## Add labels
attr(data_culture_EN_DE$agentName, "label") <- c("iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI,
HAL 9000, SIM SENSEI, CHAPPIE, AIBO, SARAH, NAO, MARCUS, DOG")
# Add label to 'AgentID'
attr(data_culture_EN_DE$AgentID, "label") <- c("1=iCAT, 2=DEEPBLUE, 3=AMY, 4=FURBY, 5=POPPIE,
6=SIRI, 7=HAL 9000, 8=SIM SENSEI, 9=CHAPPIE, 10=AIBO, 11=SARAH, 12=NAO, 13=MARCUS, 14=DOG")
# Add label to 'ConstructID'
attr(data_culture_EN_DE$ConstructID, "label") <- c("1=HLA, 2=HLB, 3=NA, 4=NB, 5=AAS, 6=AU,
7=PF, 8=AL. 9=AS, 10=APP, 11=UAA, 12=AE, 13=UE, 14=UT, 15=UAL, 16=AA, 17=AC, 18=AI,
19=AT, 20=SP, 21=IIS, 22=AEI, 23=UEP, 24=UAI")
```

Add Dutch data

Preprocessing the English ratings data from Dutch sample

Now we preprocess the English ratings data from our study with people who had Dutch as their first and primary language. Therefore, we calculate scores for each English construct/dimension per participant, by averaging the items' scores. And we add a new column 'ConstructID' to represent each construct/dimension, facilitating further analysis of the comparison between different cultural backgrounds when using the English version of the ASAQ.

We begin with the first half of the questionnaire.

```
d1_bilingual <- read_sav("summative_first_half_transformed_dutch.sav")


# Remove Q_E_ prefix from English items
for (col in 8:51) {
    colnames(d1_bilingual)[col] <- sub("Q_E_", "", colnames(d1_bilingual)[col])
}


# Select English item scores of HLA1, HLA2, HLA3 and HLA4
# for Dimension 'HLA'
HLA_EN <- data.frame(select(d1_bilingual, HLA1:HLA4))
# Select English item scores of each construct/dimension
# for Construct 1-8
HLB_EN <- data.frame(select(d1_bilingual, HLB1:HLB5))
NA_EN <- data.frame(select(d1_bilingual, NA1:NA5))
NB_EN <- data.frame(select(d1_bilingual, NB1:NB3))
AAS_EN <- data.frame(select(d1_bilingual, AAS1:AAS3))
AU_EN <- data.frame(select(d1_bilingual, AU1:AU3))
PF_EN <- data.frame(select(d1_bilingual, PF1:PF3))
AL_EN <- data.frame(select(d1_bilingual, AL1:AL5))
AS_EN <- data.frame(select(d1_bilingual, AS1:AS3))
APP_EN <- data.frame(select(d1_bilingual, APP1:APP3))
UAA_EN <- data.frame(select(d1_bilingual, UAA1:R_UAA3))
AE_EN <- data.frame(select(d1_bilingual, R_AE1:R_AE4))


# Calculate English item score mean of HLA1, HLA2, HLA3 and
# HLA4 as construct score
d1_bilingual$HLA_EN <- rowMeans(HLA_EN)
# Calculate English mean of each construct/dimension per
# participant
d1_bilingual$HLB_EN <- rowMeans(HLB_EN)
d1_bilingual$NA_EN <- rowMeans(NA_EN)
d1_bilingual$NB_EN <- rowMeans(NB_EN)
d1_bilingual$AAS_EN <- rowMeans(AAS_EN)
d1_bilingual$AU_EN <- rowMeans(AU_EN)
d1_bilingual$PF_EN <- rowMeans(PF_EN)
d1_bilingual$AL_EN <- rowMeans(AL_EN)
d1_bilingual$AS_EN <- rowMeans(AS_EN)
d1_bilingual$APP_EN <- rowMeans(APP_EN)
d1_bilingual$UAA_EN <- rowMeans(UAA_EN)
d1_bilingual$AE_EN <- rowMeans(AE_EN)
```

```r
# Transform data into the long format
bilingualdata1 <- melt(select(d1_bilingual, agentName:AE_EN),
    id.vars = c("agentName", "AgentID"), variable.name = "Construct",
    value.name = "Rating")


# Add a column 'Culture', 2 is participant group with Dutch
# cultural background
bilingualdata1$Culture <- 2


# Label ConstructID for English constructs/dimensions
bilingualdata1$ConstructID[bilingualdata1$Construct == "HLA_EN"] <- 1
bilingualdata1$ConstructID[bilingualdata1$Construct == "HLB_EN"] <- 2
bilingualdata1$ConstructID[bilingualdata1$Construct == "NA_EN"] <- 3
bilingualdata1$ConstructID[bilingualdata1$Construct == "NB_EN"] <- 4
bilingualdata1$ConstructID[bilingualdata1$Construct == "AAS_EN"] <- 5
bilingualdata1$ConstructID[bilingualdata1$Construct == "AU_EN"] <- 6
bilingualdata1$ConstructID[bilingualdata1$Construct == "PF_EN"] <- 7
bilingualdata1$ConstructID[bilingualdata1$Construct == "AL_EN"] <- 8
bilingualdata1$ConstructID[bilingualdata1$Construct == "AS_EN"] <- 9
bilingualdata1$ConstructID[bilingualdata1$Construct == "APP_EN"] <- 10
bilingualdata1$ConstructID[bilingualdata1$Construct == "UAA_EN"] <- 11
bilingualdata1$ConstructID[bilingualdata1$Construct == "AE_EN"] <- 12
```

We now perform the same preprocessing for the second half of the questionnaire.

```r
d2_bilingual <- read_sav("summative_second_half_transformed_dutch.sav")


# Remove Q_E_ prefix from English items
for (col in 8:53) {
    colnames(d2_bilingual)[col] <- sub("Q_E_", "", colnames(d2_bilingual)[col])
}


# Select English item scores of UE1, UE2 and UE3 for
# Construct 'UE'
UE_EN <- data.frame(select(d2_bilingual, UE1:UE3))
# Select English scores of the remaining 11
# constructs/dimensions
UT_EN <- data.frame(select(d2_bilingual, UT1:UT3))
UAL_EN <- data.frame(select(d2_bilingual, UAL1:UAL6))
AA_EN <- data.frame(select(d2_bilingual, AA1:AA3))
AC_EN <- data.frame(select(d2_bilingual, R_AC1:R_AC4))
AI_EN <- data.frame(select(d2_bilingual, AI1:AI4))
AT_EN <- data.frame(select(d2_bilingual, AT1:R_AT3))
SP_EN <- data.frame(select(d2_bilingual, SP1:SP3))
IIS_EN <- data.frame(select(d2_bilingual, IIS1:IIS4))
AEI_EN <- data.frame(select(d2_bilingual, AEI1:R_AEI5))
UEP_EN <- data.frame(select(d2_bilingual, UEP1:UEP4))
UAI_EN <- data.frame(select(d2_bilingual, UAI1:UAI4))
```

```
# Calculate English item score mean of UE1, UE2 and UE3 as
# construct score
d2_bilingual$UE_EN <- rowMeans(UE_EN)
# Calculate English mean of each construct/dimension per
# participant
d2_bilingual$UT_EN <- rowMeans(UT_EN)
d2_bilingual$UAL_EN <- rowMeans(UAL_EN)
d2_bilingual$AA_EN <- rowMeans(AA_EN)
d2_bilingual$AC_EN <- rowMeans(AC_EN)
d2_bilingual$AI_EN <- rowMeans(AI_EN)
d2_bilingual$AT_EN <- rowMeans(AT_EN)
d2_bilingual$SP_EN <- rowMeans(SP_EN)
d2_bilingual$IIS_EN <- rowMeans(IIS_EN)
d2_bilingual$AEI_EN <- rowMeans(AEI_EN)
d2_bilingual$UEP_EN <- rowMeans(UEP_EN)
d2_bilingual$UAI_EN <- rowMeans(UAI_EN)


# Transform data into the long format
bilingualdata2 <- melt(select(d2_bilingual, agentName:UAI_EN),
    id.vars = c("agentName", "AgentID"), variable.name = "Construct",
    value.name = "Rating")


# Add a column 'Culture', '2' is participant group with
# Dutch cultural background
bilingualdata2$Culture <- 2


# Label ConstructID for English constructs/dimensions
bilingualdata2$ConstructID[bilingualdata2$Construct == "UE_EN"] <- 13
bilingualdata2$ConstructID[bilingualdata2$Construct == "UT_EN"] <- 14
bilingualdata2$ConstructID[bilingualdata2$Construct == "UAL_EN"] <- 15
bilingualdata2$ConstructID[bilingualdata2$Construct == "AA_EN"] <- 16
bilingualdata2$ConstructID[bilingualdata2$Construct == "AC_EN"] <- 17
bilingualdata2$ConstructID[bilingualdata2$Construct == "AI_EN"] <- 18
bilingualdata2$ConstructID[bilingualdata2$Construct == "AT_EN"] <- 19
bilingualdata2$ConstructID[bilingualdata2$Construct == "SP_EN"] <- 20
bilingualdata2$ConstructID[bilingualdata2$Construct == "IIS_EN"] <- 21
bilingualdata2$ConstructID[bilingualdata2$Construct == "AEI_EN"] <- 22
bilingualdata2$ConstructID[bilingualdata2$Construct == "UEP_EN"] <- 23
bilingualdata2$ConstructID[bilingualdata2$Construct == "UAI_EN"] <- 24
```

Combine Dutch data with the remaining ratings

Now we combine our preprocessed English ratings data from a Dutch sample with the remaining ratings.

```
# Remove construct column from Dutch-English data
bilingualdata1 <- select(bilingualdata1, -Construct)
bilingualdata2 <- select(bilingualdata2, -Construct)

# Add Dutch data to remaining ratings
data_culture_DE_EN_NL = rbind(bilingualdata1, bilingualdata2,
```

```
    data_culture_EN_DE)

## Add labels Set label for 'Culture'
attr(data_culture_DE_EN_NL$Culture, "label") <- c("0=bilingual German cultural background,
1=mixed international English-speaking cultural background",
    "2=bilingual Dutch cultural background")
# Set label for the rating
attr(data_culture_DE_EN_NL$Rating, "label") <- c("Rating scores of 24 constructs/dimensions by 240 bili
```

Save combined file

And we save the resulting file.

```
write_sav(data_culture_DE_EN_NL, "data_culture_DE_EN_NL.sav",
          compress = c("byte", "none", "zsav"))
```

References

Fitrianie, Siska, Merijn Bruijnes, Fengxiang Li, Amal Abdulrahman, and Willem-Paul Brinkman. 2024. "Data and Analysis Underlying the Research into the Artificial-Social-Agent Questionnaire: Establishing the Long and Short Questionnaire Versions." 4TU.ResearchData. https://doi.org/10.4121/19758436.V6.

Fitrianie, Siska, Merijn Bruijnes, Fengxiang Li, Amal Abdulrahman, and Willem-Paul Brinkman. 2022. "The Artificial-Social-Agent Questionnaire: Establishing the Long and Short Questionnaire Versions." In *IVA '22: ACM International Conference on Intelligent Virtual Agents, Faro, Portugal, September 6 - 9, 2022*, edited by Carlos Martinho, João Dias, Joana Campos, and Dirk Heylen, 1–8. ACM. https://doi.org/10.1145/3514197.3549612.