Mandarin Chinese translation of the Artificial-Social-Agent questionnaire instrument for evaluating human-agent interaction

Transformation from raw data to the input files

Fengxiang Li

23 November, 2022

Contents

1 Introduction **1**

2 Data files **2**

2.1 File ASA_1.sav, ASA_1.1.sav and ASA_1.2.sav 3

2.2 File ASA_2.sav, ASA_2.1.sav and ASA_2.2.sav 12

2.3 File result_all.csv 21

3 Transformation results as input data files for further analysis **26**

3.1 File data01.sav 26

3.2 File data02.sav 28

3.3 File data_culture.sav 30

Reference **36**

1 Introduction

In this document, we transform seven raw data files into three input data files. We used these 3 data files for the analysis we reported. Because of privacy reasons we did not release for public access the seven raw data files. Instead, we created the 3 anonymous data files for public release. First, we downloaded six raw data files 'ASA_1.sav', 'ASA_1.1.sav', 'ASA_1.2.sav', 'ASA_2.sav', 'ASA_2.1.sav' and 'ASA_2.2.sav' from Qualtrics, which are the raw data collected from 242 bilingual participants in the ASA questionnaire translation study. We divided the 90 ASA items and 242 bilingual participants into two groups to control fatigue effects.

In the first group, human-ASA interaction evaluation data of the first 44 items (Construct 1-8: the first 12 constructs/dimensions) were collected from 121 bilingual participants with Chinese mother tongue

who are native Chinese and fluent English speakers. Bilingual participants rated human-ASA interaction on 44 English items and corresponding Chinese translations plus 15 attention control questions. Due to the difficulty of recruiting 121 bilingual participants within a single batch, we launched the questionnaire survey with the same setup in three subsequent batches. Each participant could only take part in one batch. Thus we have evaluation data of the first 44 items and their Chinese translations from three raw data files, i.e., 'ASA_1.sav', 'ASA_1.1.sav', and 'ASA_1.2.sav' in chronological order.

Similarly, in the second group, evaluation scores of the remaining 46 English items (Construct 9-19: the second 12 constructs/dimensions) and corresponding Chinese translations, plus 15 attention control questions were collected from another group of 121 bilingual participants. We also obtained the remaining 46 item scores and Chinese translation scores in three subsequent batches resulting in three raw data files, i.e., 'ASA_2.sav', 'ASA_2.1.sav' and 'ASA_2.2.sav', due to practical limitations.

Besides, the seven raw data files, the 'result_all.csv' file is the evaluation data of 532 mixed international English-speaking participants, which we collected as part of our previous study, which can be downloaded at https://osf.io/hxpsg. In that study, participants rated the same set of 14 human-ASA interaction video clips using the English ASA questionnaire (Fitrianie et al. 2022).

We transformed the seven raw data files into three input data files for further analysis, which we describe in 'Questionnaire translation and validation paper analysis'. This is underlying analysis which we report in the paper:

Mandarin Chinese translation of Artifcial-Social-Agent questionnaire instrument for evaluating human-agent interaction

Authored by Fengxiang Li, Siska Fitrianie, Merijn Bruijnes, Amal Abdulrahman, Fu Guo, Willem-Paul Brinkman.

We use the following packages:

```
library(foreign)  # Open various data files
library(haven)    # Use read_sav function
library(dplyr)    # Use select function
library(reshape2) # Use melt function
library(knitr)    # Get markdown file
library(tinytex)  # Use TeX environment
library(rticles)  # Use CTeX documents template
library(pander) # For pandering tables
panderOptions("table.alignment.default","left")
```

2 Data files

2.1 File ASA_1.sav, ASA_1.1.sav and ASA_1.2.sav

We read the three raw data files consisting of the first 44 English ASA item scores and corresponding Chinese translation scores. As mentioned above, data were collected from three independent questionnaire surveys with the same setup. Compare to 'ASA_1.sav', we added one more question 'Bilingual_check' to 'ASA_1.1.sav' and 'ASA_1.2.sav' to double-check whether the participants were bilinguals in a sense that they were native Chinese and fluent English speaker. The same applies to the data files 'ASA_2.sav', 'ASA_2.1.sav' and 'ASA_2.2.sav'. As these files are similar data structure, we only report here the content of 'ASA_1.1.sav' as an example, and report the variables in all columns. The raw scores are 7-point scale, ranging from 1 (disagree), 4 (neither agree nor disagree) to 7 (agree). Before we could do any analysis, we first had to revise the score of items with an R indication. Some columns in Table 1 are normally hidden in the raw data files 'ASA_1.1.sav', i.e., DISTR0-DISTR6, USERL0-USERL6, COMME0-COMME6, PROLI0-PROLI6, STUDY0-STUDY6, SESSI0-SESSI6, RANDO0-RANDO6, and AGENT0-AGENT6. It is probably due to Qualtics automatically generating these variables to ensure the questionnaire survey flow, and they are normally not shown when viewing the content of the SPSS files. As these data are not relevant to the analysis in the current study, we can just ignore these columns.

```
d_ASA_1.1spss <- read.spss("ASA_1.1.sav", use.value.labels=TRUE, to.data.frame=TRUE)
d_ASA_1 <- read_sav("ASA_1.sav")
d_ASA_1.1 <- read_sav("ASA_1.1.sav")
d_ASA_1.2 <- read_sav("ASA_1.2.sav")
```

```
print_table_fields(d_ASA_1.1spss, "ASA_1.1.sav") # Print all column names in the file
```

Table 1: Fields and label from SPSS file ASA_1.1.sav

variable	label
StartDate	Start Date
EndDate	End Date
Status	Response Type
Progress	Progress
Duration___in...	Duration (in seconds)
Finished	Finished
RecordedDate	Recorded Date
ResponseId	Response ID
Distribution...	Distribution Channel
DISTR0	Distribution Channel
DISTR1	Distribution Channel
DISTR2	Distribution Channel
DISTR3	Distribution Channel
DISTR4	Distribution Channel
DISTR5	Distribution Channel
DISTR6	Distribution Channel
UserLanguage	User Language
USERL0	User Language
USERL1	User Language
USERL2	User Language
USERL3	User Language
USERL4	User Language
USERL5	User Language

variable	label
USERL6	User Language
CSNT1_1	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_2	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_3	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_4	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_5	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_6	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_7	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT2	You have answered “NO” to one (or more) of the consent questions. Please select the …
PROL_ID_C	Please enter your Prolific ID
PROL_0	Please enter your Prolific ID
PROL_1	Please enter your Prolific ID
PROL_2	Please enter your Prolific ID
PROL_3	Please enter your Prolific ID
PROL_4	Please enter your Prolific ID
PROL_5	Please enter your Prolific ID
PROL_6	Please enter your Prolific ID
Bilingual_ch…	Are you fluent in both Mandarin Chinese and English?
VID_CHECK_Q	Is the video above playing?
vid_check_AC	What is displayed in the video?
TIME_VID_INT…	Timing - First Click
TIME_VID_INT…	Timing - Last Click
TIME_VID_INT…	Timing - Page Submit
TIME_VID_INT…	Timing - Click Count
AttentionChe…	[Field-agentName] is singing, however, to show that you pay attention, select …
AttentionChe…	The user can see [Field-agentName], however, please as attention check select agree
AttentionChe…	[Field-agentName] speaks English, however this is an attention check so please …

variable	label
AttentionChe…	The user runs toward [Field-agentName], but to show you pay attention please select …
AttentionChe…	[Field-agentName] can walk or not, however, select the middle option labelled with …
AttentionChe…	You've seen [Field-agentName]. However, to show that you have watched the video …
AttentionChe…	[Field-agentName]'s body colour is black. Please ignore the previous statement and …
HLA1	[Field-agentName]'s appearance is human
HLA2	[Field-agentName] has the appearance of a human
HLA3	[Field-agentName] has a human-like outside
HLA4	[Field-agentName]'s appearance makes me think of a human
HLB1	A human would behave like [Field-agentName]
HLB2	[Field-agentName]'s manners is consistent with that of people
HLB3	[Field-agentName] behavior makes me think of human behavior
HLB4	[Field-agentName] behaves like a real person
HLB5	[Field-agentName] has a human-like manner
NA1	[Field-agentName] appears like something that could exist in nature
NA2	[Field-agentName] has a natural physique
NA3	[Field-agentName]'s resemblance has an organic origin
NA4	[Field-agentName] seems natural from the outward appearance
NA5	How [Field-agentName] is represented is realistic
NB1	[Field-agentName] is alive
NB2	[Field-agentName] acts naturally
NB3	[Field-agentName] reacts like a living organism
AAS1	[Field-agentName] 's appearance is appropriate
AAS2	[Field-agentName]'s physique is suitable for its role
AAS3	[Field-agentName]' appearance was suitable
AU1	[Field-agentName] is easy to use
AU2	Learning to work with [Field-agentName] is easy
AttentionChe…	[Field-agentName] talks to the user, but to show that you are not a bot, please …

variable	label
AttentionChe...	Make sure you select the middle option labelled with the number zero to show you ...
AttentionChe...	Please select disagree here to show you still pay attention
AttentionChe...	Please select agree to pass this attention check
AttentionChe...	To show you are really reading these questions, please the option labelled with the ...
AttentionChe...	To indicate you are still there please select agree
AttentionChe...	To show you are not sleeping, as attention check please select disagree
AttentionChe...	To show you are paying attention, please select the option labelled with the number ...
AU3	Learning how to communicate with [Field-agentName] is quick
PF1	[Field-agentName] does its task well
PF2	[Field-agentName] does not hinder the user
PF3	The user is capable of succeeding with [Field-agentName]
AL1	[Field-agentName]'s appearance is pleasing
AL2	I like [Field-agentName]
R_AL3	I dislike [Field-agentName]
AL4	[Field-agentName] is cooperative
AL5	I want to hang out with [Field-agentName]
AS1	[Field-agentName] can easily mix socially
AS2	It is easy to mingle with [Field-agentName]
AS3	[Field-agentName] interacts socially with the user
APP1	[Field-agentName] has a distinctive character
R_APP2	[Field-agentName] is characterless
APP3	[Field-agentName] is an individual
UAA1	The user will use [Field-agentName] again in future
UAA2	The user can see themselves using [Field-agentName] in the future
R_UAA3	The user oppose further interaction with [Field-agentName]
R_AE1	[Field-agentName] is boring
AE2	It is interesting to interact with [Field-agentName]

variable	label
AE3	The user enjoys interacting with [Field-agentName]
R_AE4	[Field-agentName] is unpleasant to deal with
HLA1_CH	[Field-agentName] 的外观和人类一样
HLA2_CH	[Field-agentName] 具有人类的外观
HLA3_CH	[Field-agentName] 具有和人类相似的外观
HLA4_CH	[Field-agentName] 的外观让我联想到人类
HLB1_CH	人类会做出和 [Field-agentName] 一样的行为
HLB2_CH	[Field-agentName] 的举止与人一致
HLB3_CH	[Field-agentName] 的行为让我联想到人类
HLB4_CH	[Field-agentName] 的行为像一个真人
HLB5_CH	[Field-agentName] 具有类人的举止
NA1_CH	[Field-agentName] 看着像自然界中可能存在的物体
NA2_CH	[Field-agentName] 具有自然的体形
NA3_CH	[Field-agentName] 的外观具有自然的起源
NA4_CH	[Field-agentName] 从外观来看是自然的
NA5_CH	[Field-agentName] 的外观具有现实意义
NB1_CH	[Field-agentName] 是活着的
NB2_CH	[Field-agentName] 行动自然
NB3_CH	[Field-agentName] 可以像生物一样做出反应
AAS1_CH	[Field-agentName] 的外观是合适的
AAS2_CH	[Field-agentName] 的体格适合其角色
AAS3_CH	[Field-agentName] 的外观适宜
AU1_CH	[Field-agentName] 易于使用
AU2_CH	学习使用 [Field-agentName] 很容易
AU3_CH	学会和 [Field-agentName] 交流是快捷的
PF1_CH	[Field-agentName] 很好地完成了它的任务
PF2_CH	[Field-agentName] 不妨碍用户

variable	label
PF3_CH	用户能够成功使用 [Field-agentName]
AL1_CH	[Field-agentName] 的外观令人满意
AL2_CH	我喜欢 [Field-agentName]
R_AL3_CH	我不喜欢 [Field-agentName]
AL4_CH	[Field-agentName] 的表现很配合
AL5_CH	我愿意与 [Field-agentName] 玩耍
AS1_CH	[Field-agentName] 可以很容易参与社交
AS2_CH	很容易与 [Field-agentName] 打成一片
AS3_CH	[Field-agentName] 与用户进行社交互动
APP1_CH	[Field-agentName] 具有独特的性格
R_APP2_CH	[Field-agentName] 是没有个性的
APP3_CH	[Field-agentName] 是一个个体
UAA1_CH	用户将在未来再次使用 [Field-agentName]
UAA2_CH	用户可以看到自己将来会使用 [Field-agentName]
R_UAA3_CH	用户反对与 [Field-agentName] 进一步互动
R_AE1_CH	[Field-agentName] 是令人厌倦的
AE2_CH	和 [Field-agentName] 互动是有趣的
AE3_CH	用户喜欢和 [Field-agentName] 互动
R_AE4_CH	[Field-agentName] 不好相处
COMMENT	If you have any comments you can leave them here:
COMME0	If you have any comments you can leave them here:
COMME1	If you have any comments you can leave them here:
COMME2	If you have any comments you can leave them here:
COMME3	If you have any comments you can leave them here:
COMME4	If you have any comments you can leave them here:
COMME5	If you have any comments you can leave them here:
COMME6	If you have any comments you can leave them here:

variable	label
PROLIFIC_PID	PROLIFIC_PID
PROLI0	PROLIFIC_PID
PROLI1	PROLIFIC_PID
PROLI2	PROLIFIC_PID
PROLI3	PROLIFIC_PID
PROLI4	PROLIFIC_PID
PROLI5	PROLIFIC_PID
PROLI6	PROLIFIC_PID
STUDY_ID	STUDY_ID
STUDY0	STUDY_ID
STUDY1	STUDY_ID
STUDY2	STUDY_ID
STUDY3	STUDY_ID
STUDY4	STUDY_ID
STUDY5	STUDY_ID
STUDY6	STUDY_ID
SESSION_ID	SESSION_ID
SESSI0	SESSION_ID
SESSI1	SESSION_ID
SESSI2	SESSION_ID
SESSI3	SESSION_ID
SESSI4	SESSION_ID
SESSI5	SESSION_ID
SESSI6	SESSION_ID
randomflag	randomflag
RANDO0	randomflag
RANDO1	randomflag

variable	label
RANDO2	randomflag
RANDO3	randomflag
RANDO4	randomflag
RANDO5	randomflag
RANDO6	randomflag
agentName	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT0	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT1	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT2	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT3	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT4	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT5	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT6	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...

2.2 File ASA_2.sav, ASA_2.1.sav and ASA_2.2.sav

In similar manner we just described for first 44 questions, we processed the three raw data files consisting of the remaining 46 English ASA item scores and corresponding Chinese translation scores.

```
d_ASA_2.1spss <- read.spss("ASA_2.1.sav", use.value.labels=TRUE, to.data.frame=TRUE)
d_ASA_2 <- read_sav("ASA_2.sav")
d_ASA_2.1 <- read_sav("ASA_2.1.sav")
d_ASA_2.2 <- read_sav("ASA_2.2.sav")
```

```
print_table_fields(d_ASA_2.1spss, "ASA_2.1.sav") # Print all column names in the file
```

Table 2: Fields and label from SPSS file ASA_2.1.sav

variable	label
StartDate	Start Date
EndDate	End Date
Status	Response Type
Progress	Progress
Duration___in...	Duration (in seconds)
Finished	Finished
RecordedDate	Recorded Date
ResponseId	Response ID
Distribution...	Distribution Channel
DISTR0	Distribution Channel
DISTR1	Distribution Channel
DISTR2	Distribution Channel
DISTR3	Distribution Channel
DISTR4	Distribution Channel
DISTR5	Distribution Channel
DISTR6	Distribution Channel
UserLanguage	User Language
USERL0	User Language
USERL1	User Language
USERL2	User Language
USERL3	User Language
USERL4	User Language
USERL5	User Language

variable	label
USERL6	User Language
CSNT1_1	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_2	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_3	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_4	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_5	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_6	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT1_7	Consent Form for “Validation Studies of a Translated Questionnaire for Evaluating …
CSNT2	You have answered “NO” to one (or more) of the consent questions. Please select the …
PROL_ID_C	Please enter your Prolific ID
PROL_0	Please enter your Prolific ID
PROL_1	Please enter your Prolific ID
PROL_2	Please enter your Prolific ID
PROL_3	Please enter your Prolific ID
PROL_4	Please enter your Prolific ID
PROL_5	Please enter your Prolific ID
PROL_6	Please enter your Prolific ID
Bilingual_ch…	Are you fluent in both Mandarin Chinese and English?
VID_CHECK_Q	Is the video above playing?
vid_check_AC	What is displayed in the video?
TIME_VID_INT…	Timing - First Click
TIME_VID_INT…	Timing - Last Click
TIME_VID_INT…	Timing - Page Submit
TIME_VID_INT…	Timing - Click Count
AttentionChe…	[Field-agentName] is singing, however, to show that you pay attention, select …
AttentionChe…	The user can see [Field-agentName], however, please as attention check select agree
AttentionChe…	[Field-agentName] speaks English, however this is an attention check so please …

variable	label
AttentionChe…	The user runs toward [Field-agentName], but to show you pay attention please select …
AttentionChe…	[Field-agentName] can walk or not, however, select the middle option labelled with …
AttentionChe…	You've seen [Field-agentName]. However, to show that you have watched the video …
AttentionChe…	[Field-agentName]'s body colour is black. Please ignore the previous statement and …
UE1	The user was concentrated during the interaction with [Field-agentName]
UE2	The interaction captured the user's attention
UE3	The user was alert during the interaction with [Field-agentName]
UT1	[Field-agentName] always gives good advice
UT2	[Field-agentName] acts truthfully
UT3	The user can rely on [Field-agentName]
UAL1	[Field-agentName] and the user have a strategic alliance
UAL2	Collaborating with [Field-agentName] is like a joint venture
UAL3	[Field-agentName] joins the user for mutual benefit
UAL4	[Field-agentName] can collaborate in a productive way
UAL5	[Field-agentName] and the user are in sync with each other
UAL6	[Field-agentName] understands the user
AA1	[Field-agentName] remains focused on the user throughout the interaction
AA2	[Field-agentName] is attentive
AA3	The user receives [Field-agentName]'s full attention throughout the interaction
R_AC1	[Field-agentName]'s behavior does not make sense
R_AC2	[Field-agentName]'s behavior is irrational
R_AC3	[Field-agentName] is inconsistent
R_AC4	[Field-agentName] appears confused
AI1	[Field-agentName] acts intentionally
AI2	[Field-agentName] knows what it is doing
R_AI3	[Field-agentName] has no clue of what it is doing
AI4	[Field-agentName] can make its own decision

variable	label
AttentionChe...	[Field-agentName] talks to the user, but to show that you are not a bot, please ...
AttentionChe...	Make sure you select the middle option labelled with the number zero to show you ...
AttentionChe...	Please select disagree here to show you still pay attention
AttentionChe...	Please select agree to pass this attention check
AttentionChe...	To show you are really reading these questions, please the option labelled with the ...
AttentionChe...	To indicate you are still there please select agree
AttentionChe...	To show you are not sleeping, as attention check please select disagree
AttentionChe...	To show you are paying attention, please select the option labelled with the number ...
AT1	The user sees the interaction with [Field-agentName] as something positive
AT2	The user views the interaction as something favorable
R_AT3	The user thinks negatively of the interaction with [Field-agentName]
SP1	[Field-agentName] has a social presence
SP2	[Field-agentName] is a social entity
SP3	The user has the same social presence as [Field-agentName]
IIS1	The user's friends would recommend them to use [Field-agentName]
IIS2	Others would encourage the user to use [Field-agentName]
IIS3	[Field-agentName] makes the user look good
IIS4	People would look favorably at the user because of their interaction with ...
AEI1	[Field-agentName] is emotional
AEI2	[Field-agentName] experiences emotions
R_AEI3	[Field-agentName] is emotionless
AEI4	[Field-agentName] can express its feelings
R_AEI5	[Field-agentName] cannot experience emotions
UEP1	[Field-agentName]'s attitudes influences how the user feels]
UEP2	The user is influenced by [Field-agentName]'s moods
UEP3	The emotions the user feels during the interaction are caused by [Field-agentName]
UEP4	The user's Interaction with [Field-agentName] gives them an emotional sensation

variable	label
UAI1	The user's emotions influence the mood of the interaction
UAI2	[Field-agentName] reciprocates the user's actions
UAI3	[Field-agentName]'s and the user's behaviors are in direct response to each other's …
UAI4	[Field-agentName]'s and the user's emotions change to what they do to each other
UE1_CH	用户在与 [Field-agentName] 互动中是注意力集中的
UE2_CH	交互吸引了用户的注意力
UE3_CH	用户在与 [Field-agentName] 交互期间保持警觉
UT1_CH	[Field-agentName] 总是提供好的建议
UT2_CH	[Field-agentName] 如实行事
UT3_CH	[Field-agentName] 是可靠的
UAL1_CH	用户与 [Field-agentName] 有战略联盟
UAL2_CH	与 [Field-agentName] 合作就像合资
UAL3_CH	[Field-agentName] 加入用户互惠互利
UAL4_CH	用户与 [Field-agentName] 可以高效协作
UAL5_CH	[Field-agentName] 和用户相互同步
UAL6_CH	[Field-agentName] 理解用户
AA1_CH	[Field-agentName] 在整个交互过程中始终关注用户
AA2_CH	[Field-agentName] 是专注的
AA3_CH	在互动中用户获得了 [Field-agentName] 的全部关注
R_AC1_CH	[Field-agentName] 的行为讲不通
R_AC2_CH	[Field-agentName] 的行为是不合逻辑的
R_AC3_CH	[Field-agentName] 的行为不连贯
R_AC4_CH	[Field-agentName] 的行为显得很混乱
AI1_CH	[Field-agentName] 的行动是有意图的
AI2_CH	[Field-agentName] 知道它在做什么
R_AI3_CH	[Field-agentName] 不知道它在做什么
AI4_CH	[Field-agentName] 可以自己做决定

variable	label
AT1_CH	用户认为与 [Field-agentName] 的交互是积极的
AT2_CH	用户认为交互是有利的
R_AT3_CH	用户对与 [Field-agentName] 的交互持负面看法
SP1_CH	[Field-agentName] 具有社会临场感
SP2_CH	[Field-agentName] 是一个社会实体
SP3_CH	用户具有与 [Field-agentName] 相同的社会临场感
IIS1_CH	用户的朋友会推荐他们使用 [Field-agentName]
IIS2_CH	其他人会鼓励用户使用 [Field-agentName]
IIS3_CH	[Field-agentName] 使用户看起来很好
IIS4_CH	与 [Field-agentName] 的交互让人们对该用户有好印象
AEI1_CH	[Field-agentName] 是情绪化的
AEI2_CH	[Field-agentName] 可以体验情绪
R_AEI3_CH	[Field-agentName] 是没有情绪的
AEI4_CH	[Field-agentName] 可以表达它的情感
R_AEI5_CH	[Field-agentName] 无法体验情绪
UEP1_CH	[Field-agentName] 的态度影响用户的感受
UEP2_CH	用户受 [Field-agentName] 的情绪影响
UEP3_CH	用户在交互过程中感受到的情绪是由 [Field-agentName] 引起的
UEP4_CH	用户与 [Field-agentName] 的互动给他们一种情感上的感觉
UAI1_CH	用户的情绪会影响交互的氛围
UAI2_CH	[Field-agentName] 会回应用户的行为
UAI3_CH	[Field-agentName] 和用户的行为是对彼此行为的直接反应
UAI4_CH	[Field-agentName] 和用户的情绪根据他们对彼此的行为而改变
COMMENT	If you have any comments you can leave them here:
COMME0	If you have any comments you can leave them here:
COMME1	If you have any comments you can leave them here:
COMME2	If you have any comments you can leave them here:

variable	label
COMME3	If you have any comments you can leave them here:
COMME4	If you have any comments you can leave them here:
COMME5	If you have any comments you can leave them here:
COMME6	If you have any comments you can leave them here:
PROLIFIC_PID	PROLIFIC_PID
PROLI0	PROLIFIC_PID
PROLI1	PROLIFIC_PID
PROLI2	PROLIFIC_PID
PROLI3	PROLIFIC_PID
PROLI4	PROLIFIC_PID
PROLI5	PROLIFIC_PID
PROLI6	PROLIFIC_PID
STUDY_ID	STUDY_ID
STUDY0	STUDY_ID
STUDY1	STUDY_ID
STUDY2	STUDY_ID
STUDY3	STUDY_ID
STUDY4	STUDY_ID
STUDY5	STUDY_ID
STUDY6	STUDY_ID
SESSION_ID	SESSION_ID
SESSI0	SESSION_ID
SESSI1	SESSION_ID
SESSI2	SESSION_ID
SESSI3	SESSION_ID
SESSI4	SESSION_ID
SESSI5	SESSION_ID

variable	label
SESSI6	SESSION_ID
randomflag	randomflag
RANDO0	randomflag
RANDO1	randomflag
RANDO2	randomflag
RANDO3	randomflag
RANDO4	randomflag
RANDO5	randomflag
RANDO6	randomflag
agentName	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT0	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT1	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT2	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT3	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT4	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT5	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...
AGENT6	iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI, HAL 9000, SIM SENSEI, CHAPPIE, AIBO, ...

2.3 File result_all.csv

The input file 'result_all.csv' is data we obtained in a previous study. It includes information of start date, end date, record date, agent name, check score and evaluation scores of 131 ASA items. Participants that failed more than two attention control questions have already been removed from this file. However, the file contained scores of 131 ASA questionnaire items. They include the 90 official ASAQ items and 41 old items that were removed from the final version of the ASAQ questionnaire. To clean up the data set we removed these 41 old items. To make it clear, in this file, all scores have already been converted into the range [-3,3] (-3 disagree, 0 neither agree nor disagree, +3 agree), and reverse-scoring questionnaire items have been reversed.

```
data_mix <- read.csv2("result_all.csv", header = TRUE, sep =",")
```

Table 3: Fields and label from file result_all.csv

variable	label
STARTDATE	Survey start date
ENDDATE	Survey end date
RECORDDATE	Record date
GROUP	Agent name
CheckScore	Attention check score
C01D01Q3	[The agent]'s appearance is human
C01D01Q10	[The agent] has the appearance of a human
C01D01Q13	[The agent] has a human-like outside
C01D01Q15	[The agent]'s appearance makes me think of a human
C01D01Q16	[The agent] has a human likeness
C01D02Q7	A human would behave like [the agent]
C01D02Q10	[The agent]'s manners is consistent with that of people
C01D02Q0	[The agent] behavior makes me think of human behavior
C01D02Q1	[The agent] behaves like a real person
C01D02Q9	[The agent] has a human-like manner
C01D03Q7	[The agent] appears like something that could exist in nature
C01D03Q9	[The agent] has a natural physique
C01D03Q10	[The agent]'s resemblance has an organic origin
C01D03Q11	[The agent] seems natural from the outward appearance

variable	label
C01D03Q12	How [the agent] is represented is realistic
C01D04Q0	[The agent] is alive
C01D04Q13	[The agent] acts naturally
C01D04Q15	[The agent] reacts like a living organism
C01D04Q8	[The agent]'s behavior is like an animal
C01D04Q9	There are animals that behave similar to [the agent]
C01D05Q2	[The agent] 's appearance is appropriate
C01D05Q9	[The agent]'s physique is suitable for its role
C01D05Q11	[The agent]'s appearance is functional
C01D05Q12	[The agent]'s physique is fit for purpose
C01D05Q0	[The agent]' appearance was suitable
C02D00Q1	[The agent] is easy to use
C02D00Q2	Learning to work with [the agent] is easy
C02D00Q8	Learning how to communicate with [the agent] is quick
R_C02D00Q11	[R] A manual is needed to communicate with [the agent]
R_C02D00Q15	[R] [The agent] is difficult to use
C03D01Q1	The amount time it took to complete the task is acceptable
C03D01Q7	[The agent] does its task well
C03D01Q5	[The agent] enhances the user's effectiveness on the task
C03D01Q6	[The agent] allows the user to get the user's task done more quickly
C03D01Q12	[The agent] does not hinder the user
C03D02Q0	The user did well
C03D02Q5	The user is able to carry out the task well
C03D02Q10	The user is capable of succeeding with [the agent]
R_C03D02Q13	[R] The user failed the task
R_C03D02Q1	[R] The user did badly
C04D00Q1	[The agent]'s appearance is pleasing
C04D00Q10	I like [the agent]
R_C04D00Q11	[R] I dislike [the agent]
C04D00Q4	[The agent] is cooperative
C04D00Q12	I want to hang out with [the agent]

variable	label
R_C05D00Q3	[R] Interacting with [the agent] is awkward
C05D00Q6	[The agent] can easily mix socially
C05D00Q7	It is easy to mingle with [the agent]
C05D00Q1	[The agent] interacts socially with the user
C05D00Q18	[The agent] is empathic
C06D01Q1	[The agent] has a distinctive character
R_C06D01Q6	[R] [The agent] is characterless
R_C06D01Q13	[R] [The agent] lacks character
C06D01Q8	[The agent] has a coherent identity
C06D01Q11	[The agent] is an individual
C07D00Q0	The user will use [the agent] again in future
C07D00Q11	The user can see themselves using [the agent] in the future
C07D00Q13	[The agent] is acceptable
R_C07D00Q15	[R] The user opposes further interaction with [the agent]
C07D00Q14	The user is positive about future interactions with [the agent]
R_C08D00Q3	[R] [The agent] is boring
C08D00Q4	It is interesting to interact with [the agent]
R_C08D00Q10	[R] [The agent] is terrible to cope with
C08D00Q0	The user enjoys interacting with [the agent]
R_C08D00Q9	[R] [The agent] is unpleasant to deal with
C09D00Q4	The user was concentrated during the interaction with [the agent]
C09D00Q5	The user forgot about their surroundings when interacting with [the agent]
C09D00Q10	The interaction captured the user's attention
C09D00Q3	Interacting with [the agent] was absorbing
C09D00Q1	The user was alert during the interaction with [the agent]
R_C09D00Q9	[R] The user was looking for distractions during the interaction with [the agent]
C10D00Q1	The user is confident that the information provided by [the agent] is trustworthy
C10D00Q6	[The agent] always gives good advice
C10D00Q10	[The agent] is an expert in the topic discussed
C10D00Q16	[The agent] acts truthfully

variable	label
C10D00Q18	The user can rely on [the agent]
C11D01Q6	[The agent] appreciates the user's efforts
C11D01Q9	[The agent] and the user have a strategic alliance
C11D01Q10	Collaborating with [the agent] is like a joint venture
C11D01Q11	[The agent] joins the user for mutual benefit
C11D01Q14	[The agent] can collaborate in a productive way
C11D02Q1	The user can understand [the agent]
C11D02Q3	[The agent] cares about the user
C11D02Q4	[The agent] and the user are in sync with each other
C11D02Q7	[The agent] understands the user
C11D02Q8	[The agent] and the user are close
C12D00Q0	[The agent] remains focused on the user throughout the interaction
C12D00Q13	[The agent] is attentive
C12D00Q1	The user receives [the agent]'s full attention throughout the interaction
C12D00Q8	[The agent] notices if the user leaves
R_C12D00Q6	[R] [The agent] does not notice what happens around it
R_C13D00Q3	[R] [The agent]'s behavior does not make sense
C13D00Q7	[The agent] is logical
R_C13D00Q11	[R] [The agent]'s behavior is irrational
R_C13D00Q6	[R] [The agent] is inconsistent
R_C13D00Q15	[R] [The agent] appears confused
C14D00Q0	[The agent] acts intentionally
C14D00Q10	[The agent] knows what it is doing
R_C14D00Q13	[R] [The agent] has no clue of what it is doing
C14D00Q15	I can understand [the agent]'s reasoning
C14D00Q16	[The agent] can make its own decision
C15D00Q10	The user sees the interaction with [the agent] as something positive
C15D00Q11	The user views the interaction as something favorable

variable	label
C15D00Q8	Interacting with [the agent] is rewarding
C15D00Q12	Engaging with [the agent] is a good thing
R_C15D00Q15	[R] The user thinks negatively of the interaction with [the agent]
C16D00Q11	During the interaction the user perceives [the agent] as a social entity
C16D00Q12	[The agent] has a social presence
C16D00Q13	[The agent] is a social entity
C16D00Q16	The user behaves as if [the agent] is a social entity
C16D00Q17	The user has the same social presence as [the agent]
C17D00Q2	The user's friends would recommend them to use [the agent]
C17D00Q3	Others would encourage the user to use [the agent]
C17D00Q8	The user feels] cool when others see them interacting with [the agent]
C17D00Q9	[The agent] makes the user look good
C17D00Q10	People would look favorably at the user because of their interaction with [the agent]
C18D01Q2	[The agent] is emotional
C18D01Q3	[The agent] experiences emotions
R_C18D01Q5	[R] [The agent] is emotionless
C18D01Q8	[The agent] can express its feelings
R_C18D01Q14	[R] [The agent] cannot experience emotions
C18D03Q0	[The agent]'s attitudes influences how the user feels
C18D03Q1	The user is influenced by [the agent]'s moods
C18D03Q2	The emotions the user feels during the interaction are caused by [the agent]
C18D03Q9	The user's interaction with [the agent] is an emotional experience
C18D03Q12	The user's Interaction with [the agent] gives them an emotional sensation
C19D00Q3	[The agent]'s emotions change to what the user does

variable	label
C19D00Q7	The user's emotions influence the mood of the interaction
C19D00Q14	[The agent] reciprocates the user's actions
C19D00Q26	[The agent]'s and the user's behaviors are in direct response to each other's behavior
C19D00Q28	[The agent]'s and the user's emotions change to what they do to each other

3 Transformation results as input data files for further analysis

3.1 File data01.sav

We removed scores of attention control questions and other irrelevant data, e.g.,'StartDate' and 'Duration', retaining ratings of English items and corresponding Chinese translations, also with 'AgentID' and 'agentName'. Some of participants didn't complete the questionnaire survey, resulting in the null values in the raw data files 'ASA_1.sav', 'ASA_1.1.sav' and 'ASA_1.2.sav'. We omitted data rows of these participants that contain null values.

```
d_ASA_1 <- data.frame(select(d_ASA_1, AttentionCheck_1:R_AE4_CH, agentName))
d_ASA_1.1 <- data.frame(select(d_ASA_1.1, AttentionCheck_1:R_AE4_CH, agentName))
d_ASA_1.2 <- data.frame(select(d_ASA_1.2, AttentionCheck_1:R_AE4_CH, agentName))
# Select attention control scores, English scores and Chinese scores, and 'agentName'
d_ASA_1 <- na.omit(d_ASA_1)
d_ASA_1.1 <- na.omit(d_ASA_1.1)
d_ASA_1.2 <- na.omit(d_ASA_1.2)
# Omit data rows containing NA
dd1 <- rbind(d_ASA_1, d_ASA_1.1, d_ASA_1.2)
# Combine scores of 121 bilingual participants
i <- grep("AttentionCheck",colnames(dd1))
# Find the column number of attention control questions
Atten <- c(1,7,7,1,4,7,1,7,4,1,7,6,7,1,5)
# Correct answers for the 15 attention control questions
x <- NULL # Row number of participant who failed the attention check
for (j in (1:nrow(dd1))){
# Find participants who failed attention check in 'dd1'
  count <- 0
  # The number of incorrectly answered attention control questions of each participant
  for (k in 1:15){
```

```
      if (as.numeric(dd1[[i[k]]][j])!=Atten[k]) # Check whether each participant's
      # attention control question answers are consistent with the correct answers
        count <- count+1
  }
      if (count>2)
      # Row number of the participant who failed more than two
      # attention control questions were added to 'x'
      x <- append(x,j)
      # Participants who failed more than two attention control questions
}
m <- length(x) # The number of participants who failed attention check
if (m!=0)
  dd1 <- dd1[-x,] # Participants who failed attention check were excluded
```

All participants' evaluation data were included as none of the participants failed more than 2 out of the 15 attention control questions. The number of participants failing more than 2 attention control questions is: 0. Here, we converted scores in 'd_ASA_1', 'd_ASA_1.1' and 'd_ASA_1.2' from the range [1,7] into the range [-3,3], and reversed scores of questionnaire items and Chinese translations with R indication. Thus, we obtained the output data file 'data01.sav' as one of the three input data files for further analysis.

```
d1 <- as.data.frame(select(dd1, HLA1:AU2, AU3:R_AE4_CH))
# Select scores of 44 English items and corresponding Chinese translations
for (j in 1:ncol(d1)){
# Go step by step to 44 items and corresponding translations
  d1[[j]][] <- d1[[j]][]-4 # Transform scores from the range [1,7] into [-3,3]
}
for (i in grep("R",colnames(d1))){
# Find column number of reversing-scoring items and translations
  d1[[i]][] <- d1[[i]][]*(-1)
  # Reverse scores of reverse-scoring items and translations
}
dd1 <- as.data.frame(cbind(d1, select(dd1,agentName)))
# Add the 'agentName' column data
dd1$AgentID <- 0 # Add a column 'AgentID' to facilitate analysis for comparison
# between different cultural backgrounds in the main markdown file
# No specific meaning for 14 numbers, just a code for each agent
dd1$AgentID[dd1$agentName=='iCAT']<-1
dd1$AgentID[dd1$agentName=='DEEPBLUE']<-2
dd1$AgentID[dd1$agentName=='AMY']<-3
```

```
dd1$AgentID[dd1$agentName=='FURBY']<-4
dd1$AgentID[dd1$agentName=='POPPIE']<-5
dd1$AgentID[dd1$agentName=='SIRI']<-6
dd1$AgentID[dd1$agentName=='HAL 9000']<-7
dd1$AgentID[dd1$agentName=='SIM SENSEI']<-8
dd1$AgentID[dd1$agentName=='CHAPPIE']<-9
dd1$AgentID[dd1$agentName=='AIBO']<-10
dd1$AgentID[dd1$agentName=='SARAH']<-11
dd1$AgentID[dd1$agentName=='NAO']<-12
dd1$AgentID[dd1$agentName=='MARCUS']<-13
dd1$AgentID[dd1$agentName=='DOG']<-14
# Label AgentID for 14 ASAs
attr(dd1$agentName,"label")<-c("iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI,
HAL 9000, SIM SENSEI, CHAPPIE, AIBO, SARAH, NAO, MARCUS, DOG")
# Add label to 'AgentID'
attr(dd1$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 'AgentID'
```

3.2 File data02.sav

Following the same procedure as explained in subsection 'File data01.sav', we transformed the evaluation data of the remaining 46 ASA items and corresponding Chinese translations.

```
d_ASA_2 <- data.frame(select(d_ASA_2, AttentionCheck_1:UAI4_CH, agentName))
d_ASA_2.1 <- data.frame(select(d_ASA_2.1, AttentionCheck_1:UAI4_CH, agentName))
d_ASA_2.2 <- data.frame(select(d_ASA_2.2, AttentionCheck_1:UAI4_CH, agentName))
# Select attention control scores, English scores and Chinese scores, and 'agentName'
d_ASA_2 <- na.omit(d_ASA_2)
d_ASA_2.1 <- na.omit(d_ASA_2.1)
d_ASA_2.2 <- na.omit(d_ASA_2.2)
# Omit data rows containing NA
dd2 <- rbind(d_ASA_2, d_ASA_2.1, d_ASA_2.2)
# Combine scores of 121 bilingual participants
i <- grep("AttentionCheck",colnames(dd2))
# Find the column number of attention control questions
Atten <- c(1,7,7,1,4,7,1,7,4,1,7,6,7,1,5)
# Correct answers for the 15 attention control questions
y <- NULL # Row number of participant who failed the attention check
for (j in (1:nrow(dd2))){
```

```
# Find participants who failed attention checks in 'dd2'
  count <- 0
  # The number of incorrectly answered attention control questions of each participant
  for (k in 1:15){
      if (as.numeric(dd2[[i[k]]][j])!=Atten[k]) # Check whether each participant's
      # attention control question answers are consistent with the correct answers
        count <- count+1
  }
      if (count>2)
      # Row number of the participant who failed more than two
      # attention control questions were added to 'x'
      y <- append(y,j)
      # Participants who failed more than two attention control questions
}
m <- length(y) # The number of participants who failed attention check
if (m!=0)
  dd2 <- dd2[-y,] # Participants who failed attention check were excluded
```

The number of participants failing more than 2 attention control questions is: 0. All 121 participants' data satisfy the inclusion requirements. Similarly, the output this time resulted in the second input data files we used for the analysis we reported in the paper.

```
d2 <- as.data.frame(select(dd2, UE1:AI4, AT1:UAI4_CH))
# Select scores of 46 English items and corresponding Chinese translations
for (j in 1:ncol(d2)){
# Go step by step to 46 items and corresponding translations
  d2[[j]][] <- d2[[j]][]-4 # Transform scores from the range [1,7] into [-3,3]
}
for (i in grep("R",colnames(d2))){
# Find column number of reversing-scoring items and translations
  d2[[i]][] <- d2[[i]][]*(-1)
  # Reverse scores of reverse-scoring items and translations
}
dd2 <- as.data.frame(cbind(d2, select(dd2,agentName)))
# Add the 'agentName' column data
dd2$AgentID <- 0  # Add a column 'AgentID' to facilitate analysis for comparison
# between different cultural backgrounds in the main markdown file.
# No specific meaning for 14 numbers,just a code for each agent
dd2$AgentID[dd2$agentName=='iCAT']<-1
dd2$AgentID[dd2$agentName=='DEEPBLUE']<-2
```

```
dd2$AgentID[dd2$agentName=='AMY']<-3
dd2$AgentID[dd2$agentName=='FURBY']<-4
dd2$AgentID[dd2$agentName=='POPPIE']<-5
dd2$AgentID[dd2$agentName=='SIRI']<-6
dd2$AgentID[dd2$agentName=='HAL 9000']<-7
dd2$AgentID[dd2$agentName=='SIM SENSEI']<-8
dd2$AgentID[dd2$agentName=='CHAPPIE']<-9
dd2$AgentID[dd2$agentName=='AIBO']<-10
dd2$AgentID[dd2$agentName=='SARAH']<-11
dd2$AgentID[dd2$agentName=='NAO']<-12
dd2$AgentID[dd2$agentName=='MARCUS']<-13
dd2$AgentID[dd2$agentName=='DOG']<-14
# Label AgentID for 14 ASAs
attr(dd2$agentName,"label")<-c("iCAT, DEEPBLUE, AMY, FURBY, POPPIE, SIRI,
HAL 9000, SIM SENSEI, CHAPPIE, AIBO, SARAH, NAO, MARCUS, DOG")
# Add label to 'AgentID'
attr(dd2$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 'AgentID'
```

3.3 File data_culture.sav

To formulate the input data file for the subsection 'Comparison of Human-ASA Interaction between Different Cultural Backgrounds' in the paper, we transformed the three data files into the file 'data_culture.sav'. As this was only a comparison on construct/dimension level based on data collected with the English version of the questionnaire items, we first calculated 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 English version of ASAQ. This we did both for data frame 'dd1' and 'dd2'.

```
d1_bilingual <-dd1
HLA_EN<-data.frame(select(d1_bilingual,HLA1:HLA4))
# Select English item scores of HLA1, HLA2, HLA3 and HLA4 for Dimension 'HLA'
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:AU2,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))
# Select English item scores of each construct/dimension for Construct 1-8
d1_bilingual$HLA_EN<-rowMeans(HLA_EN)
# Calculate English item score mean of HLA1, HLA2, HLA3 and HLA4 as construct score
d1_bilingual$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)
# Calculate English mean of each construct/dimension per participant
bilingualdata1<-melt(select(d1_bilingual,agentName:AE_EN),id.vars=c("agentName","AgentID"),
                     variable.name="Construct",value.name="Rating")
# Transform data into the long format
bilingualdata1$Culture <- 0
# Add a column 'Culture', '0' is participant group with bilingual cultural background
bilingualdata1$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
# Label ConstructID for English constructs/dimensions
```

Then we did the same manipulation to data frame 'dd2'.

```
d2_bilingual<-dd2
UE_EN<-data.frame(select(d2_bilingual,UE1:UE3))
# Select English item scores of UE1, UE2 and UE3 for Construct 'UE'
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))
# Select English scores of the remaining 12 constructs/dimensions
d2_bilingual$UE_EN<-rowMeans(UE_EN)
# Calculate English item score mean of UE1, UE2 and UE3 as construct score
d2_bilingual$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)
# Calculate English mean of each construct/dimension per participant
bilingualdata2<-melt(select(d2_bilingual,agentName:UAI_EN),id.vars=c("agentName","AgentID"),
                     variable.name="Construct",value.name="Rating")
# Transform data into the long format
bilingualdata2$Culture <- 0
# Add a column 'Culture', '0' is participant group with bilingual cultural background
bilingualdata2$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
# Label ConstructID for English constructs/dimensions
```

Next, we calculated scores for the 24 constructs/dimensions in the file 'result_all.csv'. The two studies shared the same set of human-ASA interaction video clips and participants rated these agents on the same English ASA questionnaire items. The output of the transformation result 'data_culture.sav' in this subsection was used as input data file for further analysis of how two cultural backgrounds influence human-ASA interaction. The output file includes five variables: 'agentName', 'AgentID', 'Rating', 'Culture', and 'ConstructID'. Moreover, the description of all columns in the 'data_culture.sav' can be found in the document 'Questionnaire translation and validation paper analysis.pdf'.

```
colnames(data_mix)[5]<-"agentName"
# Change the column name 'GROUP' to 'agentName'
d_mix <- na.omit(select(data_mix,-(RID:RECORDDATE),-CheckScore))
# Omit irrelevant data, i.e.,'STARTDATE','ENDDATE','RECORDDATE','CheckScore',
d_mix<-select(d_mix,agentName,C01D01Q3,C01D01Q10,C01D01Q13,C01D01Q15,C01D02Q7,
              C01D02Q10,C01D02Q0,C01D02Q1,C01D02Q9,C01D03Q7,C01D03Q9,C01D03Q10,
              C01D03Q11,C01D03Q12,C01D04Q0,C01D04Q13,C01D04Q15,C01D05Q2,C01D05Q9,
              C01D05Q0,C02D00Q1,C02D00Q2,C02D00Q8,C03D01Q7,C03D01Q12,C03D02Q10,
              C04D00Q1,C04D00Q10,R_C04D00Q11,C04D00Q4,C04D00Q12,C05D00Q6,C05D00Q7,
              C05D00Q1,C06D01Q1,R_C06D01Q6,C06D01Q11,C07D00Q0,C07D00Q11,R_C07D00Q15,
              R_C08D00Q3,C08D00Q4,C08D00Q0,R_C08D00Q9,C09D00Q4,C09D00Q10,C09D00Q1,
              C10D00Q6,C10D00Q16,C10D00Q18,C11D01Q9,C11D01Q10,C11D01Q11,C11D01Q14,
              C11D02Q4,C11D02Q7,C12D00Q0,C12D00Q13,C12D00Q1,R_C13D00Q3,R_C13D00Q11,
              R_C13D00Q6,R_C13D00Q15,C14D00Q0,C14D00Q10,R_C14D00Q13,C14D00Q16,
              C15D00Q10,C15D00Q11,R_C15D00Q15,C16D00Q12,C16D00Q13,C16D00Q17,C17D00Q2,
              C17D00Q3,C17D00Q9,C17D00Q10,C18D01Q2,C18D01Q3,R_C18D01Q5,C18D01Q8,
              R_C18D01Q14,C18D03Q0,C18D03Q1,C18D03Q2,C18D03Q12,C19D00Q7,C19D00Q14,
              C19D00Q26,C19D00Q28)
# Select 90 items rated in the current study from Construct 1-19 and 'agentName'
names(d_mix)<-c('agentName','HLA1','HLA2','HLA3','HLA4','HLB1','HLB2',
                'HLB3','HLB4','HLB5','NA1','NA2','NA3','NA4','NA5','NB1',
                'NB2','NB3','AAS1','AAS2','AAS3','AU1','AU2',
```

```
                'AU3','PF1','PF2','PF3','AL1','AL2','R_AL3','AL4',
                'AL5','AS1','AS2','AS3','APP1','R_APP2','APP3','UAA1',
                'UAA2','R_UAA3','R_AE1','AE2','AE3','R_AE4','UE1','UE2',
                'UE3','UT1','UT2','UT3','UAL1','UAL2','UAL3','UAL4',
                'UAL5','UAL6','AA1','AA2','AA3','R_AC1','R_AC2','R_AC3',
                'R_AC4','AI1','AI2','R_AI3','AI4','AT1','AT2','R_AT3',
                'SP1','SP2','SP3','IIS1','IIS2','IIS3','IIS4','AEI1',
                'AEI2','R_AEI3','AEI4','R_AEI5','UEP1','UEP2','UEP3','UEP4',
                'UAI1','UAI2','UAI3','UAI4')
# Rename the column names to match data frame 'dd1' and 'dd2'
d_mix$Culture<-1 # Add a column 'Culture'
d_mix$AgentID <-0 # 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
d_mix$AgentID[d_mix$agentName=='iCAT']<-1
d_mix$AgentID[d_mix$agentName=='DEEPBLUE']<-2
d_mix$AgentID[d_mix$agentName=='AMY']<-3
d_mix$AgentID[d_mix$agentName=='FURBY']<-4
d_mix$AgentID[d_mix$agentName=='POPPIE']<-5
d_mix$AgentID[d_mix$agentName=='SIRI']<-6
d_mix$AgentID[d_mix$agentName=='HAL 9000']<-7
d_mix$AgentID[d_mix$agentName=='SIM SENSEI']<-8
d_mix$AgentID[d_mix$agentName=='CHAPPIE']<-9
d_mix$AgentID[d_mix$agentName=='AIBO']<-10
d_mix$AgentID[d_mix$agentName=='SARAH']<-11
d_mix$AgentID[d_mix$agentName=='NAO']<-12
d_mix$AgentID[d_mix$agentName=='MARCUS']<-13
d_mix$AgentID[d_mix$agentName=='DOG']<-14
# Label 'AgentID' for 14 ASAs to match 'bilingualdata1' and 'bilingualdata2'
HLA_MIX<-select(d_mix,HLA1:HLA4)
#Select item scores of HLA1, HLA2, HLA3 and HLA4 for Dimension 'HLA'
HLB_MIX<-select(d_mix,HLB1:HLB5)
NA_MIX<-select(d_mix,NA1:NA5)
NB_MIX<-select(d_mix,NB1:NB3)
AAS_MIX<-select(d_mix,AAS1:AAS3)
AU_MIX<-select(d_mix,AU1:AU3)
PF_MIX<-select(d_mix,PF1:PF3)
AL_MIX<-select(d_mix,AL1:AL5)
AS_MIX<-select(d_mix,AS1:AS3)
```

```
APP_MIX<-select(d_mix,APP1:APP3)
UAA_MIX<-select(d_mix,UAA1:R_UAA3)
AE_MIX<-select(d_mix,R_AE1:R_AE4)
UE_MIX<-select(d_mix,UE1:UE3)
UT_MIX<-select(d_mix,UT1:UT3)
UAL_MIX<-select(d_mix,UAL1:UAL6)
AA_MIX<-select(d_mix,AA1:AA3)
AC_MIX<-select(d_mix,R_AC1:R_AC4)
AI_MIX<-select(d_mix,AI1:AI4)
AT_MIX<-select(d_mix,AT1:R_AT3)
SP_MIX<-select(d_mix,SP1:SP3)
IIS_MIX<-select(d_mix,IIS1:IIS4)
AEI_MIX<-select(d_mix,AEI1:R_AEI5)
UEP_MIX<-select(d_mix,UEP1:UEP4)
UAI_MIX<-select(d_mix,UAI1:UAI4)
# Select item scores of 24 constructs/dimensions
d_mix$HLA_MIX<-rowMeans(HLA_MIX)
d_mix$HLB_MIX<-rowMeans(HLB_MIX)
d_mix$NA_MIX<-rowMeans(NA_MIX)
d_mix$NB_MIX<-rowMeans(NB_MIX)
d_mix$AAS_MIX<-rowMeans(AAS_MIX)
d_mix$AU_MIX<-rowMeans(AU_MIX)
d_mix$PF_MIX<-rowMeans(PF_MIX)
d_mix$AL_MIX<-rowMeans(AL_MIX)
d_mix$AS_MIX<-rowMeans(AS_MIX)
d_mix$APP_MIX<-rowMeans(APP_MIX)
d_mix$UAA_MIX<-rowMeans(UAA_MIX)
d_mix$AE_MIX<-rowMeans(AE_MIX)
d_mix$UE_MIX<-rowMeans(UE_MIX)
d_mix$UT_MIX<-rowMeans(UT_MIX)
d_mix$UAL_MIX<-rowMeans(UAL_MIX)
d_mix$AA_MIX<-rowMeans(AA_MIX)
d_mix$AC_MIX<-rowMeans(AC_MIX)
d_mix$AI_MIX<-rowMeans(AI_MIX)
d_mix$AT_MIX<-rowMeans(AT_MIX)
d_mix$SP_MIX<-rowMeans(SP_MIX)
d_mix$IIS_MIX<-rowMeans(IIS_MIX)
d_mix$AEI_MIX<-rowMeans(AEI_MIX)
d_mix$UEP_MIX<-rowMeans(UEP_MIX)
```

```r
d_mix$UAI_MIX<-rowMeans(UAI_MIX)
# Calculate mean of each construct/dimension per participant
mixdata<-melt(select(d_mix,agentName,Culture:UAI_MIX),id.vars=c("agentName","AgentID","Culture"),
              variable.name="Construct",value.name="Rating")
# Transform data into the long format
mixdata$ConstructID <-0
# Add a column 'ConstructID'
mixdata$ConstructID[mixdata$Construct=="HLA_MIX"]<-1
mixdata$ConstructID[mixdata$Construct=="HLB_MIX"]<-2
mixdata$ConstructID[mixdata$Construct=="NA_MIX"]<-3
mixdata$ConstructID[mixdata$Construct=="NB_MIX"]<-4
mixdata$ConstructID[mixdata$Construct=="AAS_MIX"]<-5
mixdata$ConstructID[mixdata$Construct=="AU_MIX"]<-6
mixdata$ConstructID[mixdata$Construct=="PF_MIX"]<-7
mixdata$ConstructID[mixdata$Construct=="AL_MIX"]<-8
mixdata$ConstructID[mixdata$Construct=="AS_MIX"]<-9
mixdata$ConstructID[mixdata$Construct=="APP_MIX"]<-10
mixdata$ConstructID[mixdata$Construct=="UAA_MIX"]<-11
mixdata$ConstructID[mixdata$Construct=="AE_MIX"]<-12
mixdata$ConstructID[mixdata$Construct=="UE_MIX"]<-13
mixdata$ConstructID[mixdata$Construct=="UT_MIX"]<-14
mixdata$ConstructID[mixdata$Construct=="UAL_MIX"]<-15
mixdata$ConstructID[mixdata$Construct=="AA_MIX"]<-16
mixdata$ConstructID[mixdata$Construct=="AC_MIX"]<-17
mixdata$ConstructID[mixdata$Construct=="AI_MIX"]<-18
mixdata$ConstructID[mixdata$Construct=="AT_MIX"]<-19
mixdata$ConstructID[mixdata$Construct=="SP_MIX"]<-20
mixdata$ConstructID[mixdata$Construct=="IIS_MIX"]<-21
mixdata$ConstructID[mixdata$Construct=="AEI_MIX"]<-22
mixdata$ConstructID[mixdata$Construct=="UEP_MIX"]<-23
mixdata$ConstructID[mixdata$Construct=="UAI_MIX"]<-24
# Label ConstructID for 24 constructs/dimensions
data_culture <- rbind(bilingualdata1, bilingualdata2, mixdata)
# Combine data 'bilingualdata1', 'bilingualdata2' and 'mixdata'
data_culture <- select(data_culture, -Construct)
# Omit data of column 'Construct', as they are not useful for further analysis
```

Reference

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 *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*, 1–8.