Supplementary Material (A) - Testing the Circumplex Structure of the Soundscape Survey

Andrew Mitchell[a,*], Francesco Aletta[a]

[a] *University College London, Institute for Environmental Design and Engineering, Central House, 14 Upper Woburn Place, London, WC1H 0NN*

---

1. Testing the quasi-circumplex structure (Steps 1 and 2)

```
library(devtools)
library(ggplot2)
library(tidyverse)
library(dplyr)
library(readxl)
library(here)
library(knitr)
library(CircE)
library(RCurl)

source(here("utils/sem_funcs.R")) # Load our own functions
dir.create(here("outputs", Sys.Date()), showWarnings = FALSE) # Create a folder for the outputs
output_dir <- here("outputs", Sys.Date())

# Prep variables for the circumplex analysis
scales <- c("PAQ1", "PAQ2", "PAQ3", "PAQ4", "PAQ5", "PAQ6", "PAQ7", "PAQ8") # Names of the scales for
eq.angles <- c(0, 45, 90, 135, 180, 225, 270, 315) # Ideal angles for circumplex analysis

# Load in the SATP dataset from Zenodo
temp.file <- paste0(tempfile(), ".xlsx")
download.file(
    "https://zenodo.org/records/10159673/files/SATP%20Dataset%20v1.4.xlsx",
    temp.file,
    mode="wb")
satp <- read_excel(temp.file,
    na = c("", "N/A"),
    col_types = c(
        "text", "text", "text", # Lan, Rec, Part
        "numeric", # Age
        "text", # Gender
        "numeric", "numeric", "numeric", "numeric",
```

*Corresponding author
Email addresses: andrew.mitchell.18@ucl.ac.uk (Andrew Mitchell), f.aletta@ucl.ac.uk (Francesco Aletta)

```
        "numeric", "numeric", "numeric", "numeric", # PAQs
        "numeric", # loud
        "text", # Inst
        "numeric" # sequence
    )
  )
```

1.0.1. *Ipsatization*

```
# Ipsatize the data
# For each participant, we subtract the mean of their response to all scales
# across all recordings from their response to each scale for each recording
# This is done at the suggestion of JM Girard/R Circumplex

satp |>
    group_by(Participant) |>
    mutate(Mean = mean(c_across(scales), na.rm = TRUE)) -> parts_means

satp[scales] <- satp |>
    select(all_of(scales)) |>
    mutate(across(all_of(scales), ~ .x - parts_means$Mean))
```

1.1. *Step One: Tracey's Circular Order Model*

```
library(RTHORR)
source(here("utils/RTHORR_funcs.R")) # Load our own functions

# Run the RTHORR analysis
matrices <- list(na.omit(satp[scales]))
names(matrices) <- "SATP"
for (lang in unique(satp$Language)) {
    lang_data <- na.omit(satp[satp$Language == lang, ][scales])
    matrices[[lang]] <- lang_data
}
randall_df_output <- my_randall_from_df(matrices, names(matrices), ord="circular8")
knitr::kable(randall_df_output, digits=3, align = "c")
```

Table 1: Results of the Circular Order analysis of the SATP dataset.

mat	pred	met	tie	CI	p	description
1	288	283	0	0.965	0.000	SATP
2	288	272	0	0.889	0.000	arb
3	288	262	0	0.819	0.000	cmn
4	288	284	0	0.972	0.000	deu
5	288	276	0	0.917	0.000	ell
6	288	286	0	0.986	0.000	eng
7	288	278	0	0.931	0.000	fra
8	288	268	0	0.861	0.000	hrv
9	288	255	0	0.771	0.000	ind
10	288	275	0	0.910	0.000	ita
11	288	264	0	0.833	0.000	jpn

Table 1: Results of the Circular Order analysis of the SATP dataset.

mat	pred	met	tie	CI	p	description
12	288	262	0	0.819	0.000	kor
13	288	261	0	0.812	0.000	nld
14	288	254	0	0.764	0.000	por
15	288	283	0	0.965	0.000	spa
16	288	284	0	0.972	0.000	swe
17	288	261	0	0.812	0.000	tur
18	288	244	0	0.694	0.002	vie
19	288	241	0	0.674	0.002	zsm

```r
print("Pass: ")
```

```
[1] "Pass: "
```

```r
print(randall_df_output[randall_df_output$CI > 0.7, "description"])
```

```
 [1] "SATP" "arb"  "cmn"  "deu"  "ell"  "eng"  "fra"  "hrv"  "ind"  "ita"
[11] "jpn"  "kor"  "nld"  "por"  "spa"  "swe"  "tur"
```

```r
print("Fail: ")
```

```
[1] "Fail: "
```

```r
print(randall_df_output[randall_df_output$CI < 0.7, "description"])
```

```
[1] "vie" "zsm"
```

```r
pass <- randall_df_output[randall_df_output$CI > 0.7, "description"][-1]
satp <- satp[satp$Language %in% pass, ] # Filter to just the languages that pass
```

1.2. Structural Equation Modelling using Browne's Stochastic Circumplex Model

1.2.0.1. Inter-rater Reliability.

Note that we do not report a measure of Inter-rater Reliability (IRR) for the original survey data as a test of the validity of the survey instrument. Although this has been used previously in soundscape research (see Erfanian et al., 2021, and Tarlao et al. (2020)), we feel that IRR imposes assumptions which do not necessarily hold in soundscape research. Primarily, IRR tests make the assumption that for a reliable response, raters should have high agreement about the subject. To rely on an IRR for a soundscape survey instrument would impose an external requirement that all listeners agree on the emotional affect evoked by a particular sound. In the context of the SATP dataset, where all participants were exposed to the same recording, an IRR metric such as Kendall's coefficient of concordance W, could be applied for the ratings given for each recording, with a high W theoretically indicating good instrument reliability. However, what it would actually indicate is a high degree of agreement regarding the emotional affect of that recording. As noted in Mitchell et al. (2022), this assumption should not necessarily be applied to soundscape assessments, given that respondents would be expected to have differing perceptual responses to the same sound and, in fact, this variability in response should be one of the outcomes to be investigated by researchers. A low IRR would therefore not necessarily indicate a poor measurement instrument, but instead could indicate a sound for which there is valid disagreement about the perception. A sound which for some groups is considered pleasant and for others it is annoying would result in a low IRR, completely independent of the validity of the measurement instrument.

1.2.1. *Run `CircE` Analysis*

The bulk of the code for this process has been pulled out into a separate `sem_funcs.R` file, which is loaded at the beginning of the analysis. This file contains the functions used to run the circumplex analysis and compile the results into a single table.

`step_one_test(data, model_type, scales = c("PAQ1", "PAQ2", "PAQ3", "PAQ4", "PAQ5", "PAQ6", "PAQ7", "PAQ8"), m = 3) )` is the function used to run the circumplex analysis for a single model for a single language. It takes the data for that language, the model_type (one of `Circumplex`, `Equal comm.`, `Equal ang.`, or `Unconstrained`), the names of the scales, and the number of betas for the fourier series correlation function (we're using `m=3` by default). It then runs the analysis and returns a list of the results, including a list of the desired results (`res_list`) and the model object (`res_model`).

`run_all_models(data, datasource, language, m)` is the function used to run the circumplex analysis for all four models for a single language. It takes the data for that language, the name of the data source (e.g. `SATP`), the language code, and the number of betas for the fourier series correlation function (m). It then runs the analysis for each of the four models and returns a list of the results, including a list of the four results and a table combining the results from all four models.

First, we run the circumplex analysis for the English data. This is done separately from the other languages to set up the results data table.

```
# Run the models for English
satp_eng <- satp[satp$Language == "eng", ]
circe_satp_eng <- run_all_models(satp_eng, "SATP", "eng", m = 3)
```

Then, we run the circumplex analysis for each of the other languages. This is done in a loop, with each language being run separately. The results for each language are then added to the results table.

Within each loop, we check for any errors in execution and append these to a list of errors to inspect later.

```
languages <- unique(satp$Language) # Get a list of all the languages

full_table <- circe_satp_eng$res_table # Start with the English results
for (lang in languages) {
    if (lang == "eng") {
        next # Skip English, we've already done it
    }
    print("======================================")
    print(lang)
    print("======================================")
    lang_data <- satp[satp$Language == lang, ] # Filter to just the language we want

    pass_on_error <- FALSE
    errors <- list()
    tryCatch(
      lang_res <- run_all_models(lang_data, "SATP", lang, m = 3),
      error = function(e) {
        pass_on_error <<- TRUE
        errors[lang] <<- e
      }
    )
    if (pass_on_error) {
      next
    }
    # lang_res <- run_all_models(lang_data, "SATP", lang, m = 3) # Run the models
```

```
    full_table <- rbind(full_table, lang_res$res_table) # Add the results to the table
}
```

```
# Catching any errors

for (name in names(errors)) {
  print("==========  Error in:  =============")
  print(errors[name])
}
```

1.2.2. SEM Analysis Results

Below is the table of results for the circumplex analysis of the soundscape survey translations. The table includes the results for each of the four models for each language. The results are presented in the order of the models, with the unconstrained model first, followed by the equal spacing model, the equal communality model, and the circumplex model. The results for each model include the χ^2 test, CFI, GFI, SRMR, RMSEA, MCSC, and GDIFF. These results are saved to a CSV file for later use.

Importantly, this table also reports the derived angles for each scale for the unconstrained and Equal comm. models. These angles will be carried over and used in the next stage of the analysis, where we will validate the survey instrument by correlating the survey responses with the acoustic indices using the Structural Summary Method (SSM).

```
write.csv(full_table, here(output_dir, "sem-fit-ipsatized.csv"))
if (is_html) {
    kable(full_table, digits = 3, align = "c") %>%
        kableExtra::kable_styling(bootstrap_options = c("striped", "hover", "condensed", "responsive",
} else {
    kableExtra::kbl(
      full_table[, ! colnames(full_table) %in% c('Dataset', 'Model Type', 'RMSEA.L', 'RMSEA.U', 'GDIFF
      format = "latex",
      row.names = FALSE,
      booktabs = T,
      digits = 3,
      align = "c",
      longtable = TRUE,
      linesep = c("", "", "", "\\addlinespace")
    ) |>
    kableExtra::kable_classic_2()
}
```

Language	n	m	ChiSq	df	p	CFI	GFI	AGFI	SRMR	MCSC	RMSEA
eng	864	3	75.86	10	0	0.99	0.98	0.93	0.04	-0.94	0.09
eng	864	3	370.06	17	0	0.93	0.91	0.8	0.05	-0.9	0.16
eng	864	3	534.25	17	0	0.9	0.87	0.72	0.1	-0.92	0.19
eng	864	3	830.65	24	0	0.85	0.81	0.72	0.11	-0.92	0.2
arb	809	3	44.04	10	0	0.99	0.99	0.96	0.02	-0.82	0.06
arb	809	3	119.26	17	0	0.97	0.97	0.93	0.04	-0.85	0.09
arb	809	3	527.62	17	0	0.86	0.86	0.71	0.18	-0.85	0.19
arb	809	3	649.44	24	0	0.82	0.84	0.76	0.18	-0.83	0.18
cmn	1832	3	172.22	10	0	0.98	0.98	0.92	0.02	-0.99	0.09
cmn	1832	3	366.93	17	0	0.96	0.95	0.9	0.04	-0.99	0.11

cmn	1832	3	1542.18	17	0	0.83	0.83	0.64	0.29	-0.96	0.22
cmn	1832	3	1716.87	24	0	0.81	0.81	0.72	0.28	-0.96	0.2
deu	810	3	23.37	10	0.0094	1	1	0.99	0.01	-1	0.04
deu	810	3	316.72	17	0	0.94	0.92	0.82	0.06	-1	0.15
deu	810	3	403.26	17	0	0.93	0.89	0.77	0.13	-0.98	0.17
deu	810	3	766.61	24	0	0.86	0.81	0.72	0.14	-0.97	0.2
ell	810	3	71.48	10	0	0.98	0.98	0.93	0.03	-1	0.09
ell	810	3	246.73	17	0	0.93	0.93	0.86	0.08	-1	0.13
ell	810	3	445.38	17	0	0.87	0.88	0.75	0.13	-0.95	0.18
ell	810	3	595.98	24	0	0.82	0.85	0.77	0.14	-0.93	0.17
fra	891	3	41.46	10	0	0.99	0.99	0.97	0.02	-0.95	0.06
fra	891	3	357.31	17	0	0.92	0.91	0.82	0.1	-0.96	0.15
fra	891	3	267.55	17	0	0.94	0.93	0.86	0.12	-0.94	0.13
fra	891	3	625.06	24	0	0.86	0.86	0.78	0.14	-0.92	0.17
hrv	864	3	58.92	10	0	0.99	0.99	0.95	0.01	-0.91	0.07
hrv	864	3	290.7	17	0	0.95	0.93	0.84	0.06	-0.9	0.14
hrv	864	3	1394.2	17	0	0.74	0.72	0.4	0.2	-0.88	0.31
hrv	864	3	1688.57	24	0	0.69	0.68	0.51	0.21	-0.87	0.28
ind	891	3	39.96	10	0	0.99	0.99	0.97	0.02	-0.96	0.06
ind	891	3	315.19	17	0	0.93	0.92	0.84	0.08	-0.9	0.14
ind	891	3	732.27	17	0	0.84	0.83	0.65	0.21	-0.93	0.22
ind	891	3	1177.14	24	0	0.74	0.76	0.63	0.26	-0.92	0.23
ita	810	3	58.38	10	0	0.99	0.98	0.95	0.03	-0.96	0.08
ita	810	3	251.28	17	0	0.94	0.93	0.86	0.07	-0.95	0.13
ita	810	3	660.37	17	0	0.85	0.84	0.65	0.16	-0.9	0.22
ita	810	3	933.15	24	0	0.78	0.78	0.67	0.17	-0.9	0.22
jpn	917	3	26.8	10	0.0028	1	1	0.99	0.01	-0.94	0.04
jpn	917	3	440.26	17	0	0.89	0.9	0.78	0.09	-0.99	0.16
jpn	917	3	760.84	17	0	0.81	0.83	0.64	0.18	-0.9	0.22
jpn	917	3	1200.25	24	0	0.7	0.76	0.64	0.2	-0.91	0.23
kor	810	3	42.95	10	0	0.99	0.99	0.96	0.01	-1	0.06
kor	810	3	220.34	17	0	0.95	0.94	0.88	0.08	-0.99	0.12
kor	810	3	605.83	17	0	0.86	0.85	0.67	0.22	-1	0.21
kor	810	3	763.88	24	0	0.82	0.81	0.72	0.23	-1	0.2
nld	864	3	31.44	10	5e-04	1	0.99	0.98	0.01	-0.94	0.05
nld	864	3	225.41	17	0	0.97	0.94	0.88	0.06	-0.92	0.12
nld	864	3	1069.73	17	0	0.83	0.77	0.5	0.25	-0.9	0.27
nld	864	3	1505.27	24	0	0.77	0.7	0.55	0.25	-0.88	0.27
por	1890	3	98.28	10	0	0.99	0.99	0.96	0.02	-0.9	0.07
por	1890	3	703.27	17	0	0.92	0.92	0.82	0.09	-0.87	0.15
por	1890	3	1279.36	17	0	0.86	0.86	0.7	0.22	-0.84	0.2
por	1890	3	2116.47	24	0	0.77	0.78	0.68	0.23	-0.83	0.22
spa	1647	3	117.44	10	0	0.99	0.98	0.94	0.04	-0.98	0.08
spa	1647	3	671.28	17	0	0.92	0.91	0.81	0.06	-0.97	0.15
spa	1647	3	489.92	17	0	0.94	0.93	0.86	0.11	-0.98	0.13
spa	1647	3	1049.99	24	0	0.88	0.86	0.8	0.13	-0.99	0.16
swe	945	3	63.54	10	0	0.99	0.99	0.95	0.02	-0.94	0.07

swe	945	3	326.09	17	0	0.94	0.92	0.84	0.05	-0.95	0.14
swe	945	3	683.64	17	0	0.88	0.85	0.68	0.13	-0.92	0.2
swe	945	3	893	24	0	0.84	0.81	0.72	0.14	-0.93	0.2
tur	918	3	107.19	10	0	0.98	0.97	0.91	0.03	-0.9	0.1
tur	918	3	358.89	17	0	0.93	0.92	0.82	0.08	-0.92	0.15
tur	918	3	744.57	17	0	0.84	0.84	0.65	0.23	-0.87	0.22
tur	918	3	1057.39	24	0	0.78	0.78	0.67	0.23	-0.83	0.22

References

Erfanian, M., Mitchell, A., Aletta, F., Kang, J., 2021. Psychological well-being and demographic factors can mediate soundscape pleasantness and eventfulness: A large sample study. Journal of Environmental Psychology 77, 101660. doi:10.1016/j.jenvp.2021.101660.

Mitchell, A., Aletta, F., Kang, J., 2022. How to analyse and represent quantitative soundscape data. JASA Express Letters 2, 037201. doi:10.1121/10.0009794, arXiv:https://doi.org/10.1121/10.0009794.

Tarlao, C., Steffens, J., Guastavino, C., 2020. Investigating contextual influences on urban soundscape evaluations with structural equation modeling. Building and Environment 188. URL: http://www.sciencedirect.com/science/article/pii/S036013232030857X, doi:10.1016/j.buildenv.2020.107490.