Experience Familiarity Analysis

Alkis Antoniades

28 March, 2022

Contents

Introduction

This file can be used to reproduce results reported in Chapter 4 of the thesis report. Specifically, these are:

1. Means, 95% highest density intervals, ROPE, and posterior probability that our hypothesis holds for whether environment personalization has a positive effect on experience familiarity elicited, as presented by Section 4.2.1.

2. Means, 95% highest density intervals, and posterior probability that our hypotheses hold for the likelihood that each sense of presence subscale examined has a positive effect on elicited experience familiarity, as presented by Table 4.6 in Section 4.2.3.3.

Theoretical Concepts Descriptions

Highest Posterior Density Interval (HPDI)

Indicates which points of a distribution are most credible and summarizes the distribution by specifying an interval that spans most of it, in our case 95%. Every point inside this interval has higher credibility than points outside it.

Region of Practical Equivalence (ROPE)

Region corresponding to the null hypothesis that no effect can be posited to have been detected. Used in conjunction with HPDI to see whether they intersect, in which case we withhold judgment regarding our hypothesis, and if not, we reject the null hypothesis and examine the position of the HPDI relative to the ROPE to determine whether the effect detected is positive or negative.

Watanabe-Akaike Information Criterion (WAIC)

Used to estimate prediction error and therefore infer the relative quality of a statistical model for the given data. A lower score indicates the quality of the model is better.

Posterior Probability

Probability of an event occurring given the data provided. In the case of our models here, it is the probability that each of our hypotheses is likely to hold given the data used to construct them. extract.samples() method used collects given number of posterior samples from the specified model constructed.

Analysis

Required files: "Data/experience_familiarity_data.csv"

Output: Knitted PDF (or HTML) with the same name as the script.

Importing Packages

```r
library(foreign) # Open various data files
library(ggplot2)
library(pander) # For rendering output
library(rethinking) # map2stan
library(tidyr) # For wide to long format transformation of the data
```

Experience Familiarity Elicited by the VR System

Loading Data

```r
ExpFamdf = read.csv("Data/experience_familiarity_data.csv")
```

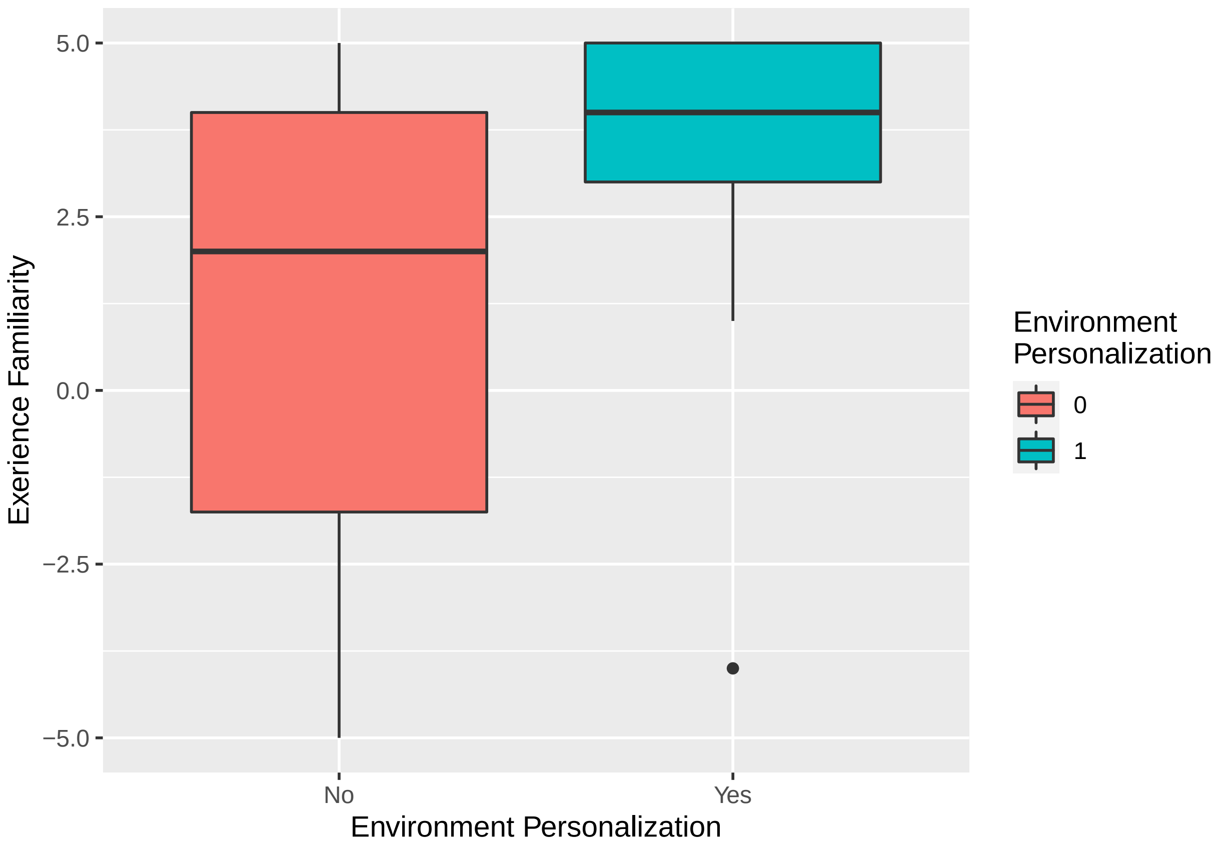
Adjusting Columns

```r
ExpFamdf$subjectF <- factor(ExpFamdf$subject, levels = c(0:19), labels = c(0:19))
ExpFamdf$subjectN <- as.numeric(ExpFamdf$subject)
```

```
ExpFamdf$env_persN <- as.numeric(ExpFamdf$env_pers)
ExpFamdf$exp_famN <- as.numeric(ExpFamdf$experience_familiarity)
ExpFamdf$env_pers <- factor(ExpFamdf$env_pers)
```

The Boxplot below compares the difference in experience familiarity elicited in personalized and non-personalized virtual environments. Presented as Figure 11 of Section 4.2.1 in the thesis report.

```
expfam_comparison <- ggplot(ExpFamdf, aes(x=factor(env_persN, labels=c("No", "Yes")),
                                          y=exp_famN, fill=env_pers)) +
  geom_boxplot(show.legend = TRUE) +
  labs(
    x = "Environment Personalization",
    y = "Exerience Familiarity", fill = "Environment\nPersonalization"
  )
print(expfam_comparison)
```



Model with a general mean and a random intercept for each participant(subject).

```
set.seed(42)

h1_m0 <- ulam(
  alist(
    exp_famN ~ dstudent(v, mu, sigma),
    v ~ gamma(2,0.1),
    mu <- aGeneral + a[subjectF] * sigmaSubjectF,
```

```r
    aGeneral ~ dnorm(0, 10),
    a[subjectF] ~ dnorm(0, 1),
    sigma ~ dexp(1),
    sigmaSubjectF ~ dexp(1)
  ), data = ExpFamdf, iter = 10000, chains = 4, cores = 4, log_lik=TRUE)
```

```r
precis(h1_m0, prob=.95)
```

```
## 19 vector or matrix parameters hidden. Use depth=2 to show them.

##                    mean        sd      2.5%     97.5%     n_eff     Rhat4
## v             1.3761084 0.2131798 1.0181639 1.8506698 11102.261 1.0001335
## aGeneral      3.6395242 0.1884873 3.2539818 3.9968311  7222.758 1.0004035
## sigma         1.0063045 0.1461010 0.7451425 1.3150946  8942.364 0.9999886
## sigmaSubjectF 0.5675542 0.1733963 0.2747212 0.9424313  5210.019 1.0002635
```

Model with a general mean, a random intercept for each participant(subject), and a personalization fixed effect.

```r
set.seed(42)

h1_m1 <- ulam(
  alist(
    exp_famN ~ dstudent(v, mu, sigma),
    v ~ gamma(2,0.1),
    mu <- aGeneral + a[subjectF] * sigmaSubjectF + b * env_persN,
    aGeneral ~ dnorm(0, 10),
    a[subjectF] ~ dnorm(0, 1),
    b ~ dnorm(0, 10),
    sigma ~ dexp(1),
    sigmaSubjectF ~ dexp(1)
  ), data = ExpFamdf, iter = 10000, chains = 4, cores = 4, log_lik=TRUE)
```

```r
precis(h1_m1, prob=.95)
```

The results here are reported in Section 4.2.1 of the thesis report.

```
## 19 vector or matrix parameters hidden. Use depth=2 to show them.

##                    mean        sd      2.5%     97.5%     n_eff     Rhat4
## v             1.5364250 0.2332270 1.1428765 2.0489236 10734.532 0.9998544
## aGeneral      2.6807669 0.2487515 2.1774689 3.1588921  8172.948 0.9999508
## b             1.3632939 0.2299759 0.9264253 1.8280839 10749.957 0.9998606
## sigma         1.0161049 0.1313532 0.7777731 1.2928415  9436.123 0.9998690
## sigmaSubjectF 0.5178861 0.1730551 0.2208934 0.8971087  4820.026 1.0000849
```

Watanabe-Akaike Information Criterion (WAIC).

```r
compare(h1_m0, h1_m1, func=WAIC)
```

```
##            WAIC       SE    dWAIC      dSE    pWAIC       weight
## h1_m1  975.3259 40.79099  0.00000       NA 16.21213 1.00000e+00
## h1_m0 1016.7059 43.21230 41.38007 14.19468 17.22968 1.03379e-09
```

Highest Density Intervals from the Model Posterior (HPDI).

```
set.seed(42)

h1_post <-extract.samples(h1_m1, n=10000)
```

```
round(HPDI(h1_post$b, prob = 0.95),2)
```

```
## |0.95 0.95|
##  0.92  1.82
```

Computing the Region of Practical Equivalence (ROPE). Range in which values have a negligible magnitude.

```
h1_ropeUpper <- c(h1_post$sigma * 0.1)
h1_ropeUpper <- sum(h1_ropeUpper) / length(h1_ropeUpper)
h1_ropeLower = h1_ropeUpper * -1

round(h1_ropeLower,2)
```

```
## [1] -0.1
```

```
round(h1_ropeUpper,2)
```

```
## [1] 0.1
```

Posterior probability that the hypothesis "Environment Personalization has a positive effect on the elicited experience familiarity" holds.

```
h1_post_prob <- length(h1_post[which(h1_post$b>0)])/length(h1_post$b)
cat("Calculated posterior probability: ",  h1_post_prob)
```

```
## Calculated posterior probability:  1
```

Computing the Highest Density Intervals from the Posterior (HPDI).

```
hpdi_lower = HPDI(h1_post$b, prob = 0.95)[1]
hpdi_higher = HPDI(h1_post$b, prob = 0.95)[2]
```

Comparing HPDI and ROPE to examine whether they intersect each other. If they do not, we can reject the null hypothesis that "there is no relation between environment personalization and experience familiarity", and if HPDI is higher, we can conclude that environment personalization has some positive effect on experience familiarity.

The following Plot is presented as Figure 12 of Section 4.2.1 in the thesis report.

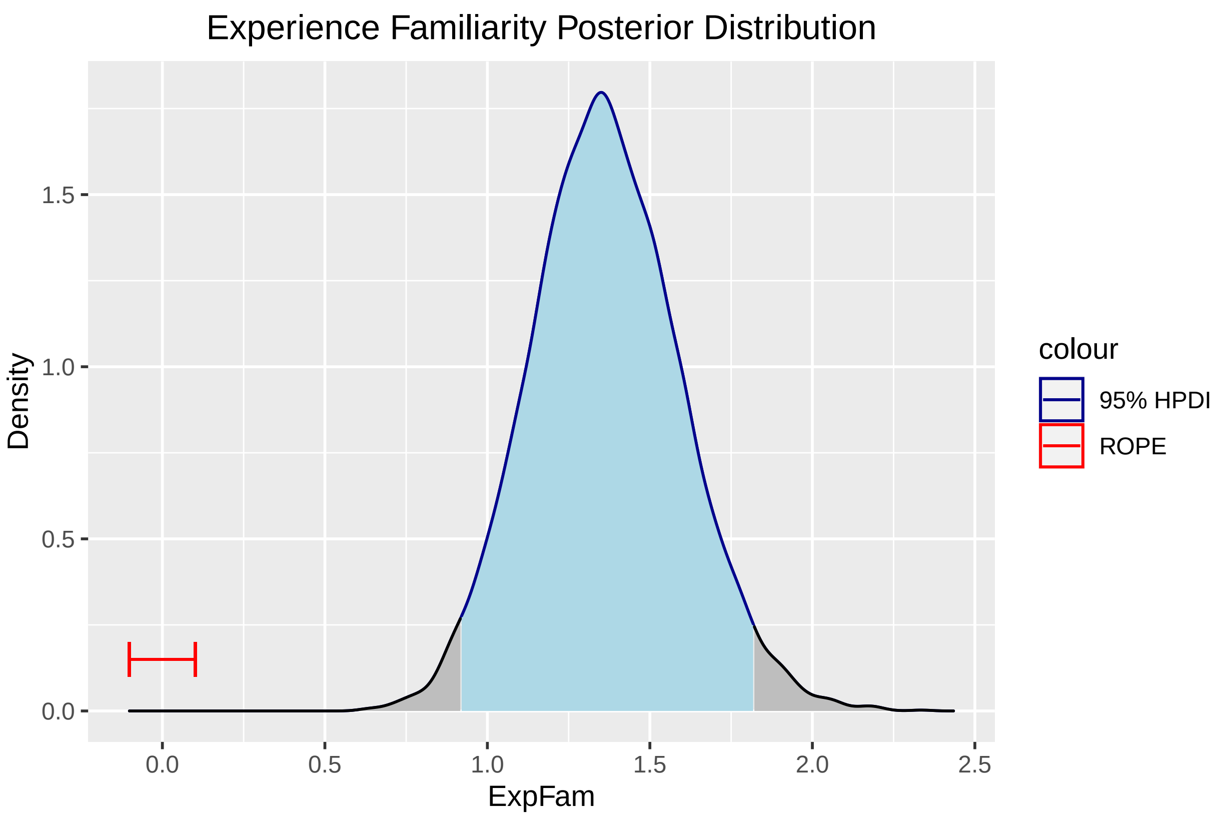
```
h1_post_df = data.frame(h1_post)
dens_df <- data.frame(x=as.numeric(density(h1_post$b)$x), y=as.numeric(density(h1_post$b)$y))

color_correcting_df <- data.frame(x= h1_ropeLower, y = min(dens_df$y))

dens_df <- rbind(dens_df, color_correcting_df)
```

```
dens_plot <- ggplot(h1_post_df, aes(b, color = "95% HPDI")) +
         geom_density() +
  geom_area(data = subset(dens_df, x >= hpdi_lower & x <= hpdi_higher),
                        aes(x=x,y=y), fill = "lightblue") +
  geom_area(data = subset(dens_df, x < hpdi_lower),
                        aes(x=x,y=y), fill = "grey", color = "black") +
  geom_area(data = subset(dens_df, x > hpdi_higher),
                        aes(x=x,y=y), fill = "grey", color = "black") +
         geom_segment(aes(x = h1_ropeLower, y = 0.1, xend = h1_ropeLower, yend = 0.2, color = "ROPE")) +
         geom_segment(aes(x = h1_ropeUpper, y = 0.1, xend = h1_ropeUpper, yend = 0.2, color = "ROPE")) +
         geom_segment(aes(x = h1_ropeLower, y = 0.15, xend = h1_ropeUpper, yend = 0.15, color = "ROPE")
         labs(
         x = "ExpFam", y = "Density",
         title = "Experience Familiarity Posterior Distribution",
         caption = "Blue area represents 95% highest posterior density region") +
  theme(plot.title = element_text(hjust = 0.5), plot.caption =  element_text(hjust = 0.5))

dens_plot + scale_color_manual(values = c("darkblue", "red")) + guides(color=guide_legend(override.aes=
```



Blue area represents 95% highest posterior density region

Experience Familiarity and Sense of Presence

The results here are reported in Table 4.6 of Section 4.2.3.2 in the thesis report.

Models for the General Presence Subscale

```
set.seed(42)

h2_m0 <- ulam(
  alist(
    exp_famN ~ dstudent(v, mu, sigma),
    v ~ gamma(2,0.1),
    mu <- aGeneral + a[subjectF] * sigmaSubjectF + c * sop_pres,
    aGeneral ~ dnorm(0, 10),
    a[subjectF] ~ dnorm(0, 1),
    c ~ dnorm(0,10),
    sigma ~ dexp(1),
    sigmaSubjectF ~ dexp(1)
  ), data = ExpFamdf, iter = 10000, chains = 4, cores = 4, log_lik=TRUE)
```

```
precis(h2_m0, prob=.95)
```

```
## 19 vector or matrix parameters hidden. Use depth=2 to show them.
```

```
##                     mean        sd       2.5%     97.5%     n_eff    Rhat4
## v             1.38326863 0.2126535  1.0165552 1.8485722 10786.287 1.000106
## aGeneral      3.23025619 0.6305162  1.9228670 4.4273959  6488.852 1.000837
## c             0.09645427 0.1428405 -0.1791190 0.3856597  6414.692 1.000907
## sigma         1.01224139 0.1462328  0.7451488 1.3148741  9033.922 1.000072
## sigmaSubjectF 0.58698231 0.1928130  0.2769441 0.9973876  3349.799 1.001582
```

```
set.seed(42)

h2_m0_post <-extract.samples(h2_m0, n=10000)
length(h2_m0_post[which(h2_m0_post$c>0)])/length(h2_m0_post$c)
```

Posterior probability that the hypothesis "This subscale has a positive effect on elicited experience familiarity" holds.

```
## [1] 0.7648
```

```
round(HPDI(h2_m0_post$c, prob = 0.95), 2)
```

```
## |0.95 0.95|
## -0.18  0.38
```

Models for the Spatial Presence Subscale

```
set.seed(42)

h3_m0 <- ulam(
  alist(
    exp_famN ~ dstudent(v, mu, sigma),
    v ~ gamma(2,0.1),
    mu <- aGeneral + a[subjectF] * sigmaSubjectF + c * sop_sp,
    aGeneral ~ dnorm(0, 10),
    a[subjectF] ~ dnorm(0, 1),
    c ~ dnorm(0,10),
    sigma ~ dexp(1),
```

```r
    sigmaSubjectF ~ dexp(1)
  ), data = ExpFamdf, iter = 10000, chains = 4, cores = 4, log_lik=TRUE)
```

```r
precis(h3_m0, prob=.95)
```

```
## 19 vector or matrix parameters hidden. Use depth=2 to show them.

##                     mean        sd       2.5%     97.5%     n_eff     Rhat4
## v             1.37637871 0.2112087  1.0175624 1.8438703 10459.741 0.9999066
## aGeneral      3.54542616 0.5885681  2.3366954 4.6839803  5113.802 1.0000174
## c             0.02492242 0.1532025 -0.2771121 0.3367621  5530.275 0.9998701
## sigma         1.00657532 0.1441438  0.7445660 1.3083943  8762.823 0.9998592
## sigmaSubjectF 0.60538701 0.2082446  0.2905311 1.0507749  2551.451 1.0002498
```

```r
set.seed(42)

h3_m0_post <-extract.samples(h3_m0, n=10000)
length(h3_m0_post[which(h3_m0_post$c>0)])/length(h3_m0_post$c)
```

Posterior probability that the hypothesis "This subscale has a positive effect on elicited experience familiarity" holds.

```
## [1] 0.5731
```

```r
round(HPDI(h3_m0_post$c, prob = 0.95), 2)
```

```
## |0.95 0.95|
## -0.30  0.31
```

Models for the Involvement Subscale

```r
set.seed(42)

h4_m0 <- ulam(
  alist(
    exp_famN ~ dstudent(v, mu, sigma),
    v ~ gamma(2,0.1),
    mu <- aGeneral + a[subjectF] * sigmaSubjectF + c * sop_inv,
    aGeneral ~ dnorm(0, 10),
    a[subjectF] ~ dnorm(0, 1),
    c ~ dnorm(0,10),
    sigma ~ dexp(1),
    sigmaSubjectF ~ dexp(1)
  ), data = ExpFamdf, iter = 10000, chains = 4, cores = 4, log_lik=TRUE)
```

```r
precis(h4_m0, prob=.95)
```

```
## 19 vector or matrix parameters hidden. Use depth=2 to show them.

##                    mean        sd       2.5%     97.5%     n_eff    Rhat4
## v             1.3802303 0.2087484  1.0235906 1.8388417 11291.579 1.000513
## aGeneral      2.9995304 0.5060054  1.9664112 3.9715184  5267.895 1.000259
## c             0.1859508 0.1351043 -0.0740286 0.4546175  5243.739 1.000248
## sigma         1.0097175 0.1439184  0.7489007 1.3128322  9456.944 1.000391
## sigmaSubjectF 0.5461040 0.1857699  0.2453599 0.9460226  4381.988 1.001535
```

```r
set.seed(42)

h4_m0_post <-extract.samples(h4_m0, n=10000)
length(h4_m0_post[which(h4_m0_post$c>0)])/length(h4_m0_post$c)
```

Posterior probability that the hypothesis "This subscale has a positive effect on elicited experience familiarity" holds.

```
## [1] 0.9171
```

```r
round(HPDI(h4_m0_post$c, prob = 0.95), 2)
```

```
## |0.95 0.95|
## -0.08  0.45
```

Models for the Experienced Realism Subscale

```r
set.seed(42)

h5_m0 <- ulam(
  alist(
    exp_famN ~ dstudent(v, mu, sigma),
    v ~ gamma(2,0.1),
    mu <- aGeneral + a[subjectF] * sigmaSubjectF + c * sop_real,
    aGeneral ~ dnorm(0, 10),
    a[subjectF] ~ dnorm(0, 1),
    c ~ dnorm(0,10),
    sigma ~ dexp(1),
    sigmaSubjectF ~ dexp(1)
  ), data = ExpFamdf, iter = 10000, chains = 4, cores = 4, log_lik=TRUE)
```

```r
precis(h5_m0, prob=.95)
```

```
## 19 vector or matrix parameters hidden. Use depth=2 to show them.
```

```
##                     mean        sd       2.5%     97.5%     n_eff     Rhat4
## v             1.37869694 0.2118843  1.0128301 1.8419834 10050.277 0.9999444
## aGeneral      3.52764744 0.4997325  2.5133198 4.4890804  5782.530 1.0002727
## c             0.03692867 0.1659547 -0.2947397 0.3698802  5992.147 1.0002005
## sigma         1.00849861 0.1451480  0.7424846 1.3113179  8722.949 0.9999379
## sigmaSubjectF 0.60262964 0.1929153  0.2991171 1.0336599  4251.651 1.0009895
```

```r
set.seed(42)

h5_m0_post <-extract.samples(h5_m0, n=10000)
length(h5_m0_post[which(h5_m0_post$c>0)])/length(h5_m0_post$c)
```

Posterior probability that the hypothesis "This subscale has a positive effect on elicited experience familiarity" holds.

```
## [1] 0.5863
```

```r
round(HPDI(h5_m0_post$c, prob = 0.95), 2)
```

```
## |0.95 0.95|
```

```
## -0.30  0.36
```