library("car")

library("MASS")

library("psych")

library('dplyr')

library("plyr")

library("DescTools")

library("lmtest")

library("lmerTest")

library("sandwich")

library('lme4')

library('tidyverse')

library('data.table')

rm(list=ls())

#loading the data:

setwd("~/Dropbox/Research assistant/VESPHA research/data")

VRdata<-read.csv("VRdata.csv", header = T) #dataset with the physiological/eye tracking measures throughout the VR session

CRAVdata<-read.csv("craving\_measures.csv", header = T) #dataset with only the subjective craving measures, taken during the VR session

SURVdata<-read.csv("survey.csv", header= T) #dataset with the data from the survey taken before and after the VR session (i.e. demographic data, presence questionnaire, dependency on cigarettes/vapes, etc)

VRinitial<-read.csv('VRfirst\_measure.csv', header=T) #dataset with only the first measures of each environment for each participant. Used to check if people returned to baseline

#### descriptives ####

# no. of males and females; overall and split by nicotine status

summary(SURVdata$Gender)

table(SURVdata$Gender, by=SURVdata$condition)

table(SURVdata$Gender, by=SURVdata$specificCondition)

# age: mean, sd, range

c(mean(SURVdata$Age, na.rm=T), sd(SURVdata$Age), range(SURVdata$Age))

by(SURVdata$Age, SURVdata$condition, function(c){c(mean(c, na.rm=T), sd(c), range(c))})

by(SURVdata$Age, SURVdata$specificCondition, function(c){c(mean(c, na.rm=T), sd(c), range(c))})

# ethnicity

table(SURVdata$Ethnicity, SURVdata$condition)

table(SURVdata$Ethnicity, SURVdata$specificCondition)

#education

table(SURVdata$Education, SURVdata$condition)

table(SURVdata$Education, SURVdata$specificCondition)

#occupation

table(SURVdata$occupation, SURVdata$condition)

table(SURVdata$occupation, SURVdata$specificCondition)

#### randomisation ####

#age

Anova(glm(Age~as.factor(condition), data=SURVdata), type="II")

#mind you, with more groups with smaller numbers of ptcpts each, there's quite some chance of at least one accidentally being younger/older than the rest

Anova(glm(Age~as.factor(specificCondition), data=SURVdata), type="II")

#gender

chisq.test(table(SURVdata$Gender, SURVdata$condition))

#see remark above

chisq.test(table(SURVdata$Gender, SURVdata$SpecificCondition))

#SES

#education

chisq.test(table(SURVdata$Education, SURVdata$condition))

#see remark above

chisq.test(table(SURVdata$Education, SURVdata$specificCondition))

#occupation

chisq.test(table(SURVdata$occupation, SURVdata$condition))

#see remark above

chisq.test(table(SURVdata$occupation, SURVdata$specificCondition))

#baseline craving

Anova(glm(Baseline\_vaping\_urge\_1~Smoking\_status, data=SURVdata), type="II")

#### reliability ####

# cigarette dependence

psych::alpha(SURVdata[,44:53])

# ENDS dependence

psych::alpha(SURVdata[,34:43])

# presence

psych::alpha(SURVdata[,77:95])

#### manipulation check ####

#check if people return to baseline after nature environment for a number of physiological measures

summary(lmer(firstIBI~as.factor(time)+as.factor(scenario)+(1|ptcpt), data=VRinitial))

Anova(lmer(firstIBI~as.factor(time)+as.factor(scenario)+(1|ptcpt), data=VRinitial))

summary(lmer(firstSC~as.factor(time)+as.factor(scenario)+(1|ptcpt), data=VRinitial))

Anova(lmer(firstSC~as.factor(time)+as.factor(scenario)+(1|ptcpt), data=VRinitial))

summary(lmer(firstPupil~as.factor(time)+(1|ptcpt), data=VRinitial))

Anova(lmer(firstPupil~as.factor(time)+as.factor(scenario)+(1|ptcpt), data=VRinitial))

##### main tests #####

## Research question 1: do virtual cues induce craving?

## 1.1 and 1.4. Do physiological measures go up when people stare at cues, and is this moderated by vaping and smoking status?

#step 1: create the unconditional means models (which wil tell us how much of the variance can be accounted for by within-person or within-scenario variance)

pupil\_m0<-lmer(pupilSize~1+(1|ptcpt), data=VRdata) #null model for pupil dilation...

IBI\_m0<-lmer(IBI~1+(1|ptcpt), data=VRdata) #...for inter beat interval...

SC\_m0<-lmer(SC~1+(1|ptcpt), data=VRdata) #...for skin conductance

#step 2: calculate intra-class correlation (ICC) for the different models

# ICC: how much scores within an individual correlate with one another

ICCpupil\_m0<-as.data.frame(VarCorr(pupil\_m0))[1,4]/(as.data.frame(VarCorr(pupil\_m0))[1,4]+as.data.frame(VarCorr(pupil\_m0))[2,4])

ICCibi\_m0<-as.data.frame(VarCorr(IBI\_m0))[1,4]/(as.data.frame(VarCorr(IBI\_m0))[1,4]+as.data.frame(VarCorr(IBI\_m0))[2,4])

ICCsc\_m0<-as.data.frame(VarCorr(SC\_m0))[1,4]/(as.data.frame(VarCorr(SC\_m0))[1,4]+as.data.frame(VarCorr(SC\_m0))[2,4])

#step 3: create the the full models, where physiological response is predicted by scenario, smoking & vaping status, and cue attention.

# Random slope for participant; random effect for scenario (basically: accounting for the fact that all individulas will have a different starting point, and a different response to the scenario, even when taking into account smoking status)

# Remaining questions: do we nest scenario under ptcpt or how do we account for repeated measures there

# Also, do we want to add a random effect for attentionCue (i.e. cue type?)

pupil\_m1<-lmer(pupilSize ~ 1 + scenario + cue + specificCondition + (1+scenario|ptcpt), data = VRdata)

IBI\_m1<-lmer(IBI ~ 1 + scenario + cue + specificCondition + (1+scenario|ptcpt), data = VRdata)

SC\_m1<-lmer(SC ~ 1 + scenario + cue + specificCondition + (1+scenario|ptcpt), data = VRdata)

#step 4: to test for model significance, you can run an anova comparing it to the null model. This will give you the overall significance

anova(pupil\_m0, pupil\_m1)

anova(IBI\_m0, IBI\_m1)

anova(SC\_m0, SC\_m1)

#step 5: to test for relevance, calculate the model explained variance (R2)

Rsq<-function(m0,m1){

SSt<-sum((m0@resp$mu-m0@resp$y)^2)

SSr<-sum((m1@resp$mu-m1@resp$y)^2)

SSm<-SSt-SSr

return(SSm/SSt)

}

Rsq(pupil\_m0, pupil\_m1)

Rsq(IBI\_m0, IBI\_m1)

Rsq(SC\_m0, SC\_m1)

#step6: inspect the model coefficients to see which was significant

summary(pupil\_m1)

summary(IBI\_m1)

summary(SC\_m1)

## 1.2 and 1.5. Is there a relationship between number of cues paid attention to in a scenario and increase in subjective craving, and is this moderated by smoking and vaping status?

## 1.3 and 1.5. Is there a relationship between time spent looking at cues in a scenario and increase in subjective craving?

#prep: create a little dataframe with the data we need

CUEdata<-data.frame('firstCrav'= unlist(c(by(CRAVdata$craving, CRAVdata$id, function(s){s[1]}), use.names = F)),

'lastCrav' = unlist(c(by(CRAVdata$craving, CRAVdata$id, function(s){s[length(s)]}), use.names = F)),

'difCrav' = unlist(c(by(CRAVdata$craving, CRAVdata$id, function(s){s[length(s)]-s[1]}), use.names = F)),

'ptcpt' = unlist(c(by(CRAVdata$ptcpt, CRAVdata$id, function(s){s[1]}), use.names = F)),

'no\_cue'= as.vector(by(VRdata$attentionCue, VRdata$filename, function(s){length(unique(s))-1})),

'time\_cue'=as.vector(by(VRdata$cue, VRdata$filename, function(s){sum(s)/4})),

'smoking\_status' = levels(CRAVdata$smoking\_status)[unlist(by(CRAVdata$smoking\_status, CRAVdata$id, function(s){s[1]}), use.names = F)],

'vaping\_status' = levels(CRAVdata$vaping\_status)[unlist(by(CRAVdata$vaping\_status, CRAVdata$id, function(s){s[1]}), use.names = F)]

)

#step 1: null model

subjCrav\_m0<-lmer(difCrav~1+(1|ptcpt), data=CUEdata)

#ICC

ICCsubjCrav\_m0<-as.data.frame(VarCorr(subjCrav\_m0))[1,4]/(as.data.frame(VarCorr(subjCrav\_m0))[1,4]+as.data.frame(VarCorr(subjCrav\_m0))[2,4])

#step 3: alternative model

subjCrav\_time<- lmer(difCrav~ 1 + time\_cue + smoking\_status \* vaping\_status + (1+scenario|ptcpt), data= CUEdata)

subjCrav\_number<- lmer(difCrav~ 1 + no\_cue + smoking\_status \* vaping\_status +(1+scenario|ptcpt), data= CUEdata)

#step 4: comparing

anova(subjCrav\_m0, subjCrav\_number)

anova(subjCrav\_m0, subjCrav\_time)

#step 5: explained variance

Rsq(subjCrav\_m0, subjCrav\_number)

Rsq(subjCrav\_m0, subjCrav\_time)

#step 6: significance individual factors

summary(subjCrav\_number)

summary(subjCrav\_time)

##1.4. Is subjective craving higher at the end of a scenario than at the start?

#step 1: null model

increaseCrav\_m0<-lmer(lastCrav~1+(1|ptcpt), data=CUEdata)

#step 2: ICC

ICCincreaseCrav\_m0<-as.data.frame(VarCorr(increaseCrav\_m0))[1,4]/(as.data.frame(VarCorr(increaseCrav\_m0))[1,4]+as.data.frame(VarCorr(increaseCrav\_m0))[2,4])

#step 3: alternative model

increaseCrav\_m1<- lmer(lastCrav~ 1 + firstCrav + (1|ptcpt), data= CUEdata)

#step 4:

anova(increaseCrav\_m0, increaseCrav\_m1)

#step 5: explained variance

Rsq(increaseCrav\_m0, increaseCrav\_m1)

## Research question 2: Reliability/validity of the measures

# 2.1 Do physiological measures correlate with subjective craving?

#step 1: null model (note that those are different from the models constructed to answer 1.1 - they use a different (thinned out) dataset)

pupil\_m0<-lmer(pupil~1+(1|ptcpt), data=CRAVdata) #null model for pupil dilation...

IBI\_m0<-lmer(IBI~1+(1|ptcpt), data=CRAVdata) #...for inter beat interval...

SC\_m0<-lmer(SC~1+(1|ptcpt), data=CRAVdata) #...for skin conductance

#step 2: ICC

ICCpupil\_m0<-as.data.frame(VarCorr(pupil\_m0))[1,4]/(as.data.frame(VarCorr(pupil\_m0))[1,4]+as.data.frame(VarCorr(pupil\_m0))[2,4])

ICCibi\_m0<-as.data.frame(VarCorr(IBI\_m0))[1,4]/(as.data.frame(VarCorr(IBI\_m0))[1,4]+as.data.frame(VarCorr(IBI\_m0))[2,4])

ICCsc\_m0<-as.data.frame(VarCorr(SC\_m0))[1,4]/(as.data.frame(VarCorr(SC\_m0))[1,4]+as.data.frame(VarCorr(SC\_m0))[2,4])

#step 3: full models

pupil\_m1<-lmer(pupil ~ 1 + craving + smoking\_status \* vaping\_status + (1+scenario|ptcpt), data = CRAVdata)

IBI\_m1<-lmer(IBI ~ 1 + craving + smoking\_status \* vaping\_status + (1+scenario|ptcpt), data = CRAVdata)

SC\_m1<-lmer(SC ~ 1 + craving + smoking\_status \* vaping\_status + (1+scenario|ptcpt), data = CRAVdata)

#step 4

anova(pupil\_m0, pupil\_m1)

anova(IBI\_m0, IBI\_m1)

anova(SC\_m0, SC\_m1)

#step 5: to test for relevance, calculate the model explained variance (R2)

Rsq(pupil\_m0, pupil\_m1)

Rsq(IBI\_m0, IBI\_m1)

Rsq(SC\_m0, SC\_m1)

#step6: inspect the model coefficients to see which was significant

summary(pupil\_m1)

summary(IBI\_m1)

summary(SC\_m1)

# 2.2. Does vaping dependence (PSECDI) correlate with physiological measures and subjective craving?

# first, create a vector that repeats the PSECDI scores

PSECDI<-c()

for(i in 1:nrow(SURVdata)) {

PSECDI<-c(PSECDI, rep(SURVdata$PSECI[i], length(which(CRAVdata$ptcpt==SURVdata$Ptctpt\_no[i]))))

}

#make sure that dependence is labeled as 0 for the participants who didn't fill out the PSECDI scale

PSECDI[which(is.na(PSECDI))]<-0

# null models have already been created, and ICCs calculated. These models predict physiological + subjective measures from PESCDI

pupil\_m2<-lmer(pupil ~ 1 + PSECDI + (1+scenario|ptcpt), data = CRAVdata)

IBI\_m2<-lmer(IBI ~ 1 + PSECDI + (1+scenario|ptcpt), data = CRAVdata)

SC\_m2<-lmer(SC ~ 1 + PSECDI + (1+scenario|ptcpt), data = CRAVdata)

subjCrave\_m2<-lmer(craving ~ 1+ PSECDI + (1+scenario|ptcpt), data = CRAVdata)

#step 4

anova(pupil\_m0, pupil\_m2)

anova(IBI\_m0, IBI\_m2)

anova(SC\_m0, SC\_m2)

anova(subjCrav\_m0, subjCrave\_m2)

#step 5: to test for relevance, calculate the model explained variance (R2)

Rsq(pupil\_m0, pupil\_m2)

Rsq(IBI\_m0, IBI\_m2)

Rsq(SC\_m0, SC\_m2)

Rsq(subjCrav\_m0, subjCrave\_m2)

#step6: inspect the model coefficients to see which was significant

summary(pupil\_m2)

summary(IBI\_m2)

summary(SC\_m2)

summary(subjCrave\_m2)

# 2.3. Does cigarette dependence (PSCDI) correlate with physiological measures and subjective craving?

PSCDI<-c()

for(i in 1:nrow(SURVdata)) {

PSCDI<-c(PSCDI, rep(SURVdata$PSCI[i], length(which(CRAVdata$ptcpt==SURVdata$Ptctpt\_no[i]))))

}

#make sure that dependence is labeled as 0 for the participants who didn't fill out the PSCDI scale

PSCDI[which(is.na(PSCDI))]<-0

pupil\_m3<-lmer(pupil ~ 1 + PSCDI + (1+scenario|ptcpt), data = CRAVdata)

IBI\_m3-lmer(IBI ~ 1 + PSCDI + (1+scenario|ptcpt), data = CRAVdata)

SC\_m3<-lmer(SC ~ 1 + PSCDI + (1+scenario|ptcpt), data = CRAVdata)

subjCrave\_m3<-lmer(craving ~ 1+ PSCDI + (1+scenario|ptcpt), data = CRAVdata)

#step 4

anova(pupil\_m0, pupil\_m3)

anova(IBI\_m0, IBI\_m3)

anova(SC\_m0, SC\_m3)

anova(subjCrav\_m0, subjCrave\_m3)

#step 5: to test for relevance, calculate the model explained variance (R2)

Rsq(pupil\_m0, pupil\_m3)

Rsq(IBI\_m0, IBI\_m3)

Rsq(SC\_m0, SC\_m3)

Rsq(subjCrav\_m0, subjCrave\_m3)

#step6: inspect the model coefficients to see which was significant

summary(pupil\_m3)

summary(IBI\_m3)

summary(SC\_m3)

summary(subjCrave\_m3)

####### Further tests

#test difference between baseline craving and post-VR craving

t.test(SURVdata$Baseline\_vaping\_urge\_1, SURVdata$craving\_post2\_1, paired=T)

#test difference between baseline and post-VR cybersickness

t.test(SURVdata$CSQ\_b, SURVdata$CSQ\_p, paired = T)