**Data Science HW3 Total: 20 points Due: January 25th, 2019, NOON**

1. In your own words, specify the research question(s) (2 sentences or less) for Study 1 & 2 [4]

* Overall RQ: Did “physically experiencing” angular momentum lead to better understanding in students.
* RQ for study 1: Did participants in the “action group” improve in correct answers to the questions “Did Woody experience more or less force than the Template?” significantly more than participants in the “observation group?”
* RQ for study 2: Did participants in the “action group” show more learning as measured by more improvement in correct answers to the question “Is Woody or the Template experiencing more force?” compared to the action group, when participants were prompted to think about vector as well as magnitude factors?

1. Include and explain IVs, DVs for Study 1 & 2. [4]

* Study 1: The IV is group membership (action vs observation). The DVs are pre test scores and post test scores for both questions. Other DVs that weren’t used for analysis were the pre and post RTs.
* Study 2: The IV is group membership (action vs observation). The DVs are pre and post-test performance for both vector-dependent and magnitude-dependent trials. Other DVs that weren’t used for analysis were the pre and post RTs.

1. R section (please complete the following and **include your script and outputs** as in-line text below)
2. Reproduce Study 1 findings (ANOVAs, t-tests; include partial eta-squared values). Also *identify and explain* the ANOVAs used (e.g. factorial 2x5, one-way, etc.). [6; bonus 1 pt for plot]

* For the pretest findings, I ran a one-way ANOVA, type 1 because there is only one factor of interest (but I got the same answer when I ran a type 3)
* For the posttest findings, I ran an ANCOVA because the authors specifically describe that they controlled for pretest scores
* I also did t tests for each group between pre and post test scores as shown by the improvement scores

1. Reproduce Study 2 findings (ANOVAs, include partial eta- squared values). Also *identify and explain* the ANOVAs used. [6; bonus 1 pt for plot]

* For the pretest findings, I ran a one-way ANOVA, type 1 because there is only one factor of interest (but I got the same answer when I ran a type 3)
* For the general posttest findings, I ran an ANCOVA because the authors specifically describe that they controlled for pretest scores
* For the vector dependent and magnitude dependent trials, I ran ANCOVAs to control for accuracy at pretest for each type.

R script:

setwd("/Users/AbigailBergman/Desktop/Grad School/Winter Quarter 2019/Data Science/datascience\_repo/week\_03")

library(car) #run type III

library(Hmisc) #import SPSS file

library(heplots) # eta squared

library(tidyverse)

library(dplyr)

library(tidyr)

library(granovaGG)

```

```{r}

physics1 = spss.get("physics1.sav", use.value.labels = TRUE)

physics2 = spss.get("physics2.sav", use.value.labels = TRUE)

```

```{r}

#study1

action <- physics1 %>%

filter(group == "action")

observation <- physics1 %>%

filter(group == "observation")

#study1

#pretest - "one way ANOVA"

fit = aov(mag1 ~ group, data = physics1)

summary(fit)

#post test - ANCOVA to control for pretest

mod1 = lm(mag2 ~ group + mag1, data = physics1,

contrasts = list(group = contr.sum))

fit1 = Anova(mod1, type = 3)

fit1

#eta squared

etasq(mod1)

#t.test for action group change from pre to post

t.test(action$mag.improvement)

#t.test for obs group change from pre to post

t.test(observation$mag.improvement)

```

```{r}

physics1 %>%

group\_by(group)%>%

mutate(mag1=mean(mag1),

mag2=mean(mag2))%>%

gather("test","accuracy", 3:4)%>%

mutate(test=as.factor(test)) %>%

ggplot() +

geom\_point(aes(test, accuracy, shape = 3))+

geom\_line(aes(group = interaction(group),test, accuracy, color = group))+

geom\_hline(aes(yintercept = .5, color = "red"), linetype = "dashed")+

ylim(c(.2,.8))+

theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank(),

panel.background = element\_blank(), axis.line = element\_line(colour = "black"),

legend.position="none",

plot.title = element\_text(hjust=0.5))+

scale\_shape\_identity()+

labs(title = "Study 1")+

scale\_color\_manual(values=c("blue", "light blue", "red"))

```

```{r}

action1 <- physics1 %>%

filter(group == "action")

observation1 <- physics1 %>%

filter(group == "observation")

#study 2

#pretest - "one way ANOVA"

fit2 = aov(accuracy.pre ~ group, data = physics2)

summary(fit)

#posttest - ANCOVA

mod3 = lm(accuracy.post ~ group + accuracy.pre, data = physics2,

contrasts = list(group = contr.sum))

fit3 = Anova(mod3, type = 3)

fit3

#eta squared

etasq(mod3)

#vector dependent trials

mod4 = lm(accuracy.posttest.newVC ~ group + accuracy.pretest.newVC, data = physics2,

contrasts = list(group = contr.sum))

fit4 = Anova(mod4, type = 3)

fit4

etasq(mod4)

#magnitude dependent trials

mod5 = lm(accuracy.posttest.newNONVC ~ group + accuracy.pretest.newNONVC, data = physics2,

contrasts = list(group = contr.sum))

fit5 = Anova(mod5, type = 3)

fit5

```

```{r}

physics2 %>%

group\_by(group)%>%

mutate(accuracy.pre=mean(accuracy.pre),

accuracy.post=mean(accuracy.post))%>%

gather("test","accuracy", 3:4)%>%

mutate(test=as.factor(test)) %>%

ggplot() +

geom\_point(aes(test, accuracy, shape = 3))+

scale\_x\_discrete(limits = c("accuracy.pre", "accuracy.post"))+

geom\_hline(aes(yintercept = .32, color = "red"), linetype = "dashed")+

geom\_line(aes(group = interaction(group),test, accuracy, color = group))+

ylim(c(.2,.8))+

theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank(),

panel.background = element\_blank(), axis.line = element\_line(colour = "black"),

legend.position="none",

plot.title = element\_text(hjust=0.5)) +

scale\_shape\_identity()+

labs(title = "Study 2") +

scale\_color\_manual(values=c("blue", "light blue", "red"))

Rendered R:

hw\_03

Abby Bergman

1/23/2019

setwd("/Users/AbigailBergman/Desktop/Grad School/Winter Quarter 2019/Data Science/datascience\_repo/week\_03")  
library(car) #run type III

## Loading required package: carData

library(Hmisc) #import SPSS file

## Loading required package: lattice

## Loading required package: survival

## Loading required package: Formula

## Loading required package: ggplot2

##   
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:base':  
##   
## format.pval, units

library(heplots) # eta squared  
library(tidyverse)

## ── Attaching packages ─────────────────────────── tidyverse 1.2.1 ──

## ✔ tibble 1.4.2 ✔ purrr 0.2.5  
## ✔ tidyr 0.8.2 ✔ dplyr 0.7.7  
## ✔ readr 1.1.1 ✔ stringr 1.3.1  
## ✔ tibble 1.4.2 ✔ forcats 0.3.0

## ── Conflicts ────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ✖ dplyr::recode() masks car::recode()  
## ✖ purrr::some() masks car::some()  
## ✖ dplyr::src() masks Hmisc::src()  
## ✖ dplyr::summarize() masks Hmisc::summarize()

library(dplyr)  
library(tidyr)  
library(granovaGG)

physics1 = spss.get("physics1.sav", use.value.labels = TRUE)

## re-encoding from latin1

physics2 = spss.get("physics2.sav", use.value.labels = TRUE)

## re-encoding from latin1

#study1  
  
action <- physics1 %>%  
 filter(group == "action")  
  
observation <- physics1 %>%  
 filter(group == "observation")  
  
#study1  
#pretest - "one way ANOVA"  
fit = aov(mag1 ~ group, data = physics1)  
summary(fit)

## Df Sum Sq Mean Sq F value Pr(>F)  
## group 1 0.0006 0.00058 0.015 0.903  
## Residuals 42 1.6390 0.03902

#post test - ANCOVA to control for pretest  
mod1 = lm(mag2 ~ group + mag1, data = physics1,  
 contrasts = list(group = contr.sum))  
fit1 = Anova(mod1, type = 3)  
fit1

## Anova Table (Type III tests)  
##   
## Response: mag2  
## Sum Sq Df F value Pr(>F)   
## (Intercept) 0.32268 1 16.6546 0.000202 \*\*\*  
## group 0.10099 1 5.2125 0.027679 \*   
## mag1 0.61941 1 31.9699 1.341e-06 \*\*\*  
## Residuals 0.79436 41   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#eta squared  
etasq(mod1)

## Partial eta^2  
## group 0.1127946  
## mag1 0.4381241  
## Residuals NA

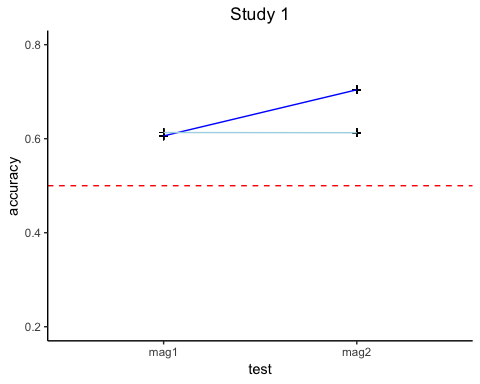
#t.test for action group change from pre to post  
t.test(action$mag.improvement)

##   
## One Sample t-test  
##   
## data: action$mag.improvement  
## t = 3.0668, df = 21, p-value = 0.005854  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## 0.03160329 0.16476035  
## sample estimates:  
## mean of x   
## 0.09818182

#t.test for obs group change from pre to post  
t.test(observation$mag.improvement)

##   
## One Sample t-test  
##   
## data: observation$mag.improvement  
## t = -0.013008, df = 21, p-value = 0.9897  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## -0.07312168 0.07221259  
## sample estimates:  
## mean of x   
## -0.0004545455

physics1 %>%  
 group\_by(group)%>%  
 mutate(mag1=mean(mag1),  
 mag2=mean(mag2))%>%  
 gather("test","accuracy", 3:4)%>%  
 mutate(test=as.factor(test)) %>%  
 ggplot() +  
 geom\_point(aes(test, accuracy, shape = 3))+  
 geom\_line(aes(group = interaction(group),test, accuracy, color = group))+  
 geom\_hline(aes(yintercept = .5, color = "red"), linetype = "dashed")+  
 ylim(c(.2,.8))+  
 theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank(),  
panel.background = element\_blank(), axis.line = element\_line(colour = "black"),   
legend.position="none",   
plot.title = element\_text(hjust=0.5))+  
 scale\_shape\_identity()+  
 labs(title = "Study 1")+  
 scale\_color\_manual(values=c("blue", "light blue", "red"))



action1 <- physics1 %>%  
 filter(group == "action")  
  
observation1 <- physics1 %>%  
 filter(group == "observation")  
  
#study 2  
#pretest - "one way ANOVA"  
fit2 = aov(accuracy.pre ~ group, data = physics2)  
summary(fit)

## Df Sum Sq Mean Sq F value Pr(>F)  
## group 1 0.0006 0.00058 0.015 0.903  
## Residuals 42 1.6390 0.03902

#posttest - ANCOVA  
mod3 = lm(accuracy.post ~ group + accuracy.pre, data = physics2,  
 contrasts = list(group = contr.sum))  
fit3 = Anova(mod3, type = 3)  
fit3

## Anova Table (Type III tests)  
##   
## Response: accuracy.post  
## Sum Sq Df F value Pr(>F)   
## (Intercept) 0.24921 1 13.5620 0.0008201 \*\*\*  
## group 0.09357 1 5.0923 0.0307783 \*   
## accuracy.pre 0.06223 1 3.3865 0.0747412 .   
## Residuals 0.60640 33   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#eta squared  
etasq(mod3)

## Partial eta^2  
## group 0.13368253  
## accuracy.pre 0.09307002  
## Residuals NA

#vector dependent trials  
mod4 = lm(accuracy.posttest.newVC ~ group + accuracy.pretest.newVC, data = physics2,  
 contrasts = list(group = contr.sum))  
fit4 = Anova(mod4, type = 3)  
fit4

## Anova Table (Type III tests)  
##   
## Response: accuracy.posttest.newVC  
## Sum Sq Df F value Pr(>F)   
## (Intercept) 0.45684 1 13.0269 0.001005 \*\*  
## group 0.19868 1 5.6653 0.023236 \*   
## accuracy.pretest.newVC 0.03265 1 0.9310 0.341622   
## Residuals 1.15727 33   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

etasq(mod4)

## Partial eta^2  
## group 0.14652180  
## accuracy.pretest.newVC 0.02743797  
## Residuals NA

#magnitude dependent trials  
mod5 = lm(accuracy.posttest.newNONVC ~ group + accuracy.pretest.newNONVC, data = physics2,  
 contrasts = list(group = contr.sum))  
fit5 = Anova(mod5, type = 3)  
fit5

## Anova Table (Type III tests)  
##   
## Response: accuracy.posttest.newNONVC  
## Sum Sq Df F value Pr(>F)   
## (Intercept) 1.05437 1 53.2239 2.261e-08 \*\*\*  
## group 0.00694 1 0.3502 0.5580   
## accuracy.pretest.newNONVC 0.03350 1 1.6909 0.2025   
## Residuals 0.65373 33   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

physics2 %>%  
 group\_by(group)%>%  
 mutate(accuracy.pre=mean(accuracy.pre),  
 accuracy.post=mean(accuracy.post))%>%  
 gather("test","accuracy", 3:4)%>%  
 mutate(test=as.factor(test)) %>%  
 ggplot() +  
 geom\_point(aes(test, accuracy, shape = 3))+  
 scale\_x\_discrete(limits = c("accuracy.pre", "accuracy.post"))+  
 geom\_hline(aes(yintercept = .32, color = "red"), linetype = "dashed")+  
 geom\_line(aes(group = interaction(group),test, accuracy, color = group))+  
 ylim(c(.2,.8))+  
 theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank(),  
panel.background = element\_blank(), axis.line = element\_line(colour = "black"),  
legend.position="none",  
plot.title = element\_text(hjust=0.5)) +  
 scale\_shape\_identity()+  
 labs(title = "Study 2") +  
 scale\_color\_manual(values=c("blue", "light blue", "red"))

