

Female responses to urinary fractions

Holly

2023-05-30

```
#data
fractions_bias<-read.csv("urinary_fractions_behavioural_biases.csv", header=T)
fractions_data<-read.csv("urinary_fractions_behavioural_data.csv", header=T)
```

```
#packages
library(tidyverse)
library(ggplot2)
library(lme4)
library(car)
library(cowplot)
library(gridExtra)
library(ggpubr)
library(ggsignif)
library(emmeans)
```

```
#make subject and date factors
fractions_data$subject<-as.factor(fractions_data$subject)
fractions_data$date<-as.factor(fractions_data$date)

#check no females look weird
ggplot(data=fractions_data, aes(x=subject, y=time.on.tile, group=donor_sex, colour=donor_sex)) +
  geom_point()
#all females show some response
```

```

#LMM: time spent sniffing LMW vs fresh HWM

#separate out LMW and fresh HWM trials
fractions_noaged<-fractions_data[fractions_data$fraction != "aged",]

#run linear mixed model to compare female sniffing responses to male compare to female fractions
sniff_noaged.lm<-lmer(lgsniff_tile~(fraction+donor_sex+contact) + (1|subject/date), data=fractions_noaged)

#check model
par(mfrow=c(2,2),mai=c(0.5,0.5,1.25,0.5))
plot(resid(sniff_noaged.lm)~fitted(sniff_noaged.lm),
     main='Residual plot', ylab='Residuals', xlab='Fitted values')
abline(h=0, lty=2)
title("Residual plots for time spent investigating fresh HWM vs LMW", line=-2, outer=T, cex.main=2)
hist(resid(sniff_noaged.lm),main='Histogram of residuals', xlab = 'Residuals')
qqnorm(resid(sniff_noaged.lm))
qqline(resid(sniff_noaged.lm))

#check if any 2-way interactions are significant
sniff_noaged2.lm<-lmer(lgsniff_tile~(fraction+contact*donor_sex) + (1|subject/date), data=fractions_noaged)
Anova(sniff_noaged2.lm)
sniff_noaged3.lm<-lmer(lgsniff_tile~(fraction*contact+donor_sex) + (1|subject/date), data=fractions_noaged)
Anova(sniff_noaged3.lm)
sniff_noaged4.lm<-lmer(lgsniff_tile~(fraction*donor_sex+contact) + (1|subject/date), data=fractions_noaged)
Anova(sniff_noaged4.lm)

anova(sniff_noaged.lm, sniff_noaged2.lm, sniff_noaged3.lm, sniff_noaged4.lm)
#simplest model best

#check 3-way interaction
sniff_noaged5.lm<-lmer(lgsniff_tile~(fraction*contact*donor_sex) + (1|subject/date), data=fractions_noaged)
Anova(sniff_noaged5.lm)

anova(sniff_noaged.lm, sniff_noaged5.lm)
#simplest model still best
#look at model outputs
Anova(sniff_noaged.lm)
summary(sniff_noaged.lm)

```

```
#LMM: time spent nearby LMW vs fresh HWM
```

```
#run linear mixed model to compare time females spent nearby male compare to female fractions
```

```
time_noaged.lm<-lmer(lgtime_on_tile~(fraction+contact+donor_sex) + (1|subject/date),  
                    data=fractions_noaged)
```

```
par(mfrow=c(2,2),mai=c(0.5,0.5,1.25,0.5))
```

```
plot(resid(time_noaged.lm)~fitted(time_noaged.lm),  
     main='Residual plot', ylab='Residuals', xlab='Fitted values')
```

```
abline(h=0, lty=2)
```

```
title("Residual plots for time nearby fresh HWM vs LMW", line=-2, outer=T, cex.main=2)
```

```
hist(resid(time_noaged.lm), main='Histogram of residuals', xlab = 'Residuals')
```

```
qqnorm(resid(time_noaged.lm))
```

```
qqline(resid(time_noaged.lm))
```

```
#look for 2-way interactions
```

```
time_noaged2.lm<-lmer(lgtime_on_tile~(fraction*contact+donor_sex) + (1|subject/date),  
                    data=fractions_noaged)
```

```
Anova(time_noaged2.lm)
```

```
time_noaged3.lm<-lmer(lgtime_on_tile~(fraction+contact*donor_sex) + (1|subject/date),  
                    data=fractions_noaged)
```

```
Anova(time_noaged3.lm)
```

```
time_noaged4.lm<-lmer(lgtime_on_tile~(fraction*donor_sex+contact) + (1|subject/date),  
                    data=fractions_noaged)
```

```
Anova(time_noaged4.lm)
```

```
anova(time_noaged.lm, time_noaged2.lm, time_noaged3.lm, time_noaged4.lm)
```

```
#simplest model best and non of 2 way interactions significant.
```

```
#check 3-way interaction
```

```
time_noaged5.lm<-lmer(lgtime_on_tile~(fraction*contact*donor_sex) +  
                    (1|subject/date) , data=fractions_noaged)
```

```
Anova(time_noaged5.lm)
```

```
anova(time_noaged2.lm, time_noaged5.lm)
```

```
#simplest model is better
```

```
#look at model output
```

```
Anova(time_noaged.lm)
```

```
summary(time_noaged.lm)
```

```

#LMM: LMW vs fresh HMW sniffing bias to male

#separate out just LMW and fresh HMW tests
bias_noaged<-fractions_bias[fractions_bias$fraction!="aged hmw",]

#run linear mixed model to compare sniffing bias to male fraction for LMW and fresh H
MW and
# with & without contact

bias_noaged_sniff.lmm<-lmer(sqrtsniff_bias~fraction +
                             contact + (1|subject), data=bias_noaged)

#check model
par(mfrow=c(2,2))
plot(resid(bias_noaged_sniff.lmm)~fitted(bias_noaged_sniff.lmm))
abline(h=0, lty=2)
hist(resid(bias_noaged_sniff.lmm))
qqnorm(resid(bias_noaged_sniff.lmm))
qqline(resid(bias_noaged_sniff.lmm))
summary(bias_noaged_sniff.lmm)

#check for interaction
bias_noaged_sniff2.lmm<-lmer(lgsniff_bias~fraction*contact + (1|subject), data=bias_n
oaged)
Anova(bias_noaged_sniff2.lmm)
anova(bias_noaged_sniff.lmm, bias_noaged_sniff2.lmm)

#simplest model is best
Anova(bias_noaged_sniff.lmm)
summary(bias_noaged_sniff.lmm)

```

```
#LMM: LMW vs fresh HMW time bias to male
```

```
#run linear mixed model to compare time bias to male fraction for LMW and fresh HMW and
```

```
# with & without contact
```

```
bias_noaged_time.lmm<-lmer(time_bias~fraction + contact + (1|subject), data=bias_noaged)
```

```
#check for interaction
```

```
bias_noaged_time2.lmm<-lmer(time_bias~fraction*contact + (1|subject), data=bias_noaged)
```

```
anova(bias_noaged_time.lmm, bias_noaged_time2.lmm)
```

```
#model including interaction is best fit
```

```
par(mfrow=c(2,2), mai=c(0.5,0.5,1.25,0.5))
```

```
plot(resid(bias_noaged_time2.lmm)~fitted(bias_noaged_time2.lmm),  
     main='', ylab='Residuals', xlab='Fitted values')
```

```
title('Residual plot', outer=F,line=1)
```

```
abline(h=0, lty=2)
```

```
title("Residual plots for bias in time nearby male \n over female HMW vs LMW urine fractions",  
      line=-3.5, outer=T, cex.main=2)
```

```
hist(resid(bias_noaged_time2.lmm),xlab = 'Residuals', main = '')
```

```
title('Histogram of Residuals', outer=F,line=1)
```

```
qqnorm(resid(bias_noaged_time2.lmm))
```

```
qqline(resid(bias_noaged_time2.lmm))
```

```
#model output
```

```
Anova(bias_noaged_time2.lmm)
```

```
summary(bias_noaged_time2.lmm)
```

```
#use contrast to compare fraction and contact
```

```
noaged_time.emm<-emmeans(bias_noaged_time2.lmm, ~ fraction|contact)
```

```
contrast(noaged_time.emm, interaction = T, simple = list("fraction","contact"))
```

```

#LMM: time spent sniffing fresh vs aged HWM

#subset dataset to just include HWM and aged HWM
hwm<-fractions_data[fractions_data$fraction!="lmw",]

#run linear mixed model to compare time females spent sniffing male compare to female
HWM fractions
sniff_hmw.lm<-lmer(lgsniff_tile~(fraction+donor_sex+contact) + (1|subject/date), data
=hmw)

#check model
par(mfrow=c(2,2))
plot(resid(sniff_hmw.lm)~fitted(sniff_hmw.lm))
abline(h=0, lty=2)
hist(resid(sniff_hmw.lm))
qqnorm(resid(sniff_hmw.lm))
qqline(resid(sniff_hmw.lm))

#check if any 2-way interactions are significant
sniff_hmw2.lm<-lmer(lgsniff_tile~(fraction+contact*donor_sex) + (1|subject/date), dat
a=hmw)
Anova(sniff_hmw2.lm)
sniff_hmw3.lm<-lmer(lgsniff_tile~(fraction*contact+donor_sex) + (1|subject/date), dat
a=hmw)
Anova(sniff_hmw3.lm)
sniff_hmw4.lm<-lmer(lgsniff_tile~(fraction*donor_sex+contact) + (1|subject/date), dat
a=hmw)
Anova(sniff_hmw4.lm)

anova(sniff_hmw.lm, sniff_hmw2.lm, sniff_hmw3.lm, sniff_hmw4.lm)
#simplest model best

#check 3-way interaction
sniff_hmw5.lm<-lmer(lgsniff_tile~(fraction*contact*donor_sex) + (1|subject/date), dat
a=hmw)
Anova(sniff_hmw5.lm)

anova(sniff_hmw.lm, sniff_hmw5.lm)
#simplest model still best

#model outputs
Anova(sniff_hmw.lm)
summary(sniff_hmw.lm)

```

#LMM: time spent nearby fresh vs aged HWM

#run linear mixed model to compare time females spent nearby male compare to female HWM fractions

```
time_hmw.lm<-lmer(lgtime_on_tile~(fraction+contact+donor_sex) + (1|subject/date),  
                  data=hmw)
```

#check model

```
par(mfrow=c(2,2))  
plot(resid(time_hmw.lm)~fitted(time_hmw.lm))  
abline(h=0, lty=2)  
hist(resid(time_hmw.lm))  
qqnorm(resid(time_hmw.lm))  
qqline(resid(time_hmw.lm))  
summary(time_hmw.lm)  
Anova(time_hmw.lm)
```

#look for 2-way interactions

```
time_hmw2.lm<-lmer(lgtime_on_tile~(fraction*contact+donor_sex) + (1|subject/date),  
                  data=hmw)
```

```
Anova(time_hmw2.lm)
```

```
time_hmw3.lm<-lmer(lgtime_on_tile~(fraction+contact*donor_sex) + (1|subject/date),  
                  data=hmw)
```

```
Anova(time_hmw3.lm)
```

```
time_hmw4.lm<-lmer(lgtime_on_tile~(fraction*donor_sex+contact) + (1|subject/date),  
                  data=hmw)
```

```
Anova(time_hmw4.lm)
```

```
anova(time_hmw.lm, time_hmw2.lm, time_hmw3.lm, time_hmw4.lm)
```

#simplest model best and non of 2 way interactions significant.

#try 3-way interaction

```
time_hmw5.lm<-lmer(lgtime_on_tile~(fraction*contact*donor_sex) +  
                  (1|subject/date) , data=hmw)
```

```
Anova(time_hmw5.lm)
```

```
anova(time_hmw.lm, time_hmw5.lm)
```

#simplest model is better

#look at model output

```
Anova(time_hmw.lm)
```

```
summary(time_hmw.lm)
```

```

#LMM: fresh vs aged HMW sniffing bias to male

#separate out just fresh & aged HMW tests
bias_hmw<-fractions_bias[fractions_bias$fraction!="lmw",]

#run linear mixed model to compare sniffing bias to male fraction for fresh & aged HM
W and
# with & without contact

bias_hmw_sniff.lmm<-lmer(lgsniff_bias~fraction + contact + (1|subject), data=bias_hm
w)

par(mfrow=c(2,2))
plot(resid(bias_hmw_sniff.lmm)~fitted(bias_hmw_sniff.lmm))
abline(h=0, lty=2)
hist(resid(bias_hmw_sniff.lmm))
qqnorm(resid(bias_hmw_sniff.lmm))
qqline(resid(bias_hmw_sniff.lmm))

#check for interaction
bias_hmw_sniff2.lmm<-lmer(lgsniff_bias~fraction*contact + (1|subject), data=bias_hmw)
Anova(bias_hmw_sniff2.lmm)
anova(bias_hmw_sniff.lmm,bias_hmw_sniff2.lmm)
#simplest model best

#model outputs
summary(bias_hmw_sniff.lmm)
Anova(bias_hmw_sniff.lmm)

```



```

#LMM: fresh vs aged HMW time bias to male

#run linear mixed model to compare time bias to male fraction for fresh & aged HMW and
# with & without contact
bias_hmw_time.lmm<-lmer(lgtime_bias~ contact + fraction + (1|subject), data=bias_hmw)

par(mfrow=c(2,2),mai=c(0.5,0.5,1.25,0.5))
plot(resid(bias_hmw_time.lmm)~fitted(bias_hmw_time.lmm),
     main='', ylab='Residuals', xlab='Fitted values')
title('Residual plot', outer=F,line=1)
abline(h=0, lty=2)
title("Residual plots for bias in time nearby male \n over female fresh vs aged HMW
fractions", line=-3.5, outer=T, cex.main=2)
hist(resid(bias_hmw_time.lmm),xlab = 'Residuals', main = '')
title('Histogram of Residuals', outer=F,line=1)
qqnorm(resid(bias_hmw_time.lmm))
qqline(resid(bias_hmw_time.lmm))


#check for interaction
bias_hmw_time2.lmm<-lmer(time_bias~fraction*contact + (1|subject), data=bias_hmw)
anova(bias_hmw_time.lmm, bias_hmw_time2.lmm)
Anova(bias_hmw_time2.lmm)
#simplest model best


#model outputs
summary(bias_hmw_time.lmm)
Anova(bias_hmw_time.lmm)

```

#effect sizes

```
se <- function(x) sqrt(var(x)/length(x))
```

#effect of contact on time sniffing/ nearby male & female fractions

```
averages_contact<-fractions_data %>% group_by(contact) %>%
```

```
  summarise(mean_time=mean(time.on.tile), se_time=se(time.on.tile), mean_sniff=mean(t  
ime.sniffing.tile), se_sniff=se(time.sniffing.tile))
```

#effect of donor sex and fraction on time sniffing/ nearby male & female fractions

```
averages_contact_sex<-fractions_data %>% group_by(fraction, donor_sex) %>%
```

```
  summarise(mean_time=mean(time.on.tile), se_time=se(time.on.tile), mean_sniff=mean(t  
ime.sniffing.tile), se_sniff=se(time.sniffing.tile))
```

#effect of fraction on time sniffing/ nearby male & female fractions

```
averages_fraction<-fractions_data %>% group_by(fraction) %>%
```

```
  summarise(mean_time=mean(time.on.tile), se_time=se(time.on.tile), mean_sniff=mean(t  
ime.sniffing.tile), se_sniff=se(time.sniffing.tile))
```

#effect of fraction on bias in time sniffing/ nearby male fractions

```
averages_fraction_bias<-fractions_bias %>% group_by(fraction) %>%
```

```
  summarise(mean_time=mean(time_bias),se_time=se(time_bias), mean_sniff=mean(sniff_bi  
as), se_sniff=se(sniff_bias))
```

#effect of contact on bias in time sniffing/ nearby male fractions

```
averages_bias_contat<-fractions_bias %>% group_by(fraction, contact) %>%
```

```
  summarise(mean_time=mean(time_bias),se_time=se(time_bias), mean_sniff=mean(sniff_bi  
as), se_sniff=se(sniff_bias))
```

#Fig 3a and b

#urinary fractions graphs

```
fractions_bias$fraction<-factor(fractions_bias$fraction, levels= c("lmw", "hmw", "aged hmw"))
```

#time on tile

```
fraction_time.gg<-ggplot(data=fractions_bias, aes(x=fraction, y=time_bias, fill=contact)) +  
  geom_boxplot(aes(fill=contact), outlier.shape = NA) + theme_bw() +  
  ylab("Time nearby not sniffing \n (male - female urine bias, s)") +  
  geom_point(position=position_jitterdodge(),aes(shape=contact),  
    stroke=1.5,colour="black", size=3) +  
  theme(axis.title.x = element_blank(),  
    axis.text.x = element_text(size=24, face= "bold", colour = "black"),  
    axis.ticks.x = element_blank(),  
    axis.title.y = element_text(size=24, face= "bold", colour = "black"),  
    axis.text.y = element_text(size=20),  
    panel.border = element_blank(), panel.grid.minor=element_blank(),  
    panel.grid.major = element_blank(),  
    axis.line = element_line(colour = "black", linewidth = 0.5),  
    axis.line.x = element_blank(),  
    legend.position = "none",  
    plot.margin=unit(c(1,1,1,1),"cm")) +  
  geom_hline(yintercept=0, linetype="dashed", color="black", size=1) +  
  scale_fill_manual(values=c("#43a2ca","#a8ddb5"), name="", breaks=c("y", "n"),  
    labels=c("No contact", "Contact")) +  
  scale_shape_manual(values=c(15,17), breaks=c("y", "n"),  
    labels=c("No contact", "Contact"), name="") +  
  scale_x_discrete(breaks=c("lmw", "hmw", "aged hmw"),  
    labels=c("LMW", "HMW", "Aged HMW"))+  
  scale_y_continuous(expand = expansion(mult = c(0.05, .1)))
```

#add significance

```
fraction_time2.gg<-fraction_time.gg + geom_signif(  
  y_position = c(160, 150, 230), xmin = c(0.8, 1.8, 2.8), xmax = c(1.2, 2.2,3.2),  
  annotation = c("?", "NS", "NS"), tip_length = 0, size=1, textsize = 10)
```

#time sniffing stimulus

```
fraction_sniff.gg<-ggplot(data=fractions_bias, aes(x=fraction, y=sniff_bias)) +  
  geom_boxplot(aes(fill=contact), outlier.shape = NA) + theme_bw() +  
  ylab("Time sniffing \n (male - female urine bias, s)") +  
  geom_point(position=position_jitterdodge(),aes(shape=contact),  
    stroke=1.5, colour="black", size=3) +  
  theme(axis.title.x = element_blank(),  
    axis.text.x = element_blank(),  
    axis.ticks.x = element_blank(),  
    axis.title.y = element_text(size=24, face= "bold", colour = "black"),  
    axis.text.y = element_text(size=20),  
    panel.border = element_blank(), panel.grid.minor=element_blank(),  
    panel.grid.major = element_blank(),axis.line = element_line(colour = "black",  
linewidth = 0.5),  
    axis.line.x = element_blank(),  
    plot.margin=unit(c(1,1,1,1),"cm")) +
```

```

geom_hline(yintercept=0, linetype="dashed", color="black", size=1) +
scale_fill_manual(values=c("#43a2ca", "#a8ddb5"), name="", breaks=c("y", "n"),
                  labels=c("No contact", "Contact")) +
scale_shape_manual(values=c(15,17), breaks=c("y", "n"),
                   labels=c("No contact", "Contact"), name="") +
theme(legend.text = element_text(size=22, face= "bold", colour = "black"),
      legend.key.size = unit(2,'cm')) +
scale_x_discrete(breaks=c("lmw", "hmw", "aged hmw"),
                 labels=c("LMW", "HMW", "Aged HMW"))

#add significance
sniff_sig <- data.frame(signif = c("*", "NS"),
                      y_position = c(60,80),
                      group = c(1,2),
                      start = c("lmw", "hmw"),
                      end = c("hmw", "aged hmw"))

fraction_sniff2.gg<-fraction_sniff.gg + geom_signif(aes(
  xmin= start,
  xmax= end,
  y_position = y_position,
  annotations = signif),
  data = sniff_sig , manual = TRUE, inherit.aes = FALSE, size=1, textsize = 10, tip
_length = 0)

#combine both behavioral graphs into one figure
#get legend
fraction_legend<-get_legend(fraction_sniff.gg)
#remove legend from sniffing graph
fraction_sniff3.gg<-fraction_sniff2.gg + theme(legend.position = "none")

#plot on grid
fraction_grid<-plot_grid(fraction_sniff3.gg, fraction_time2.gg, align="vh", ncol=1)
fraction_grid

#add legend to grid
grid.arrange(arrangeGrob(fraction_grid, fraction_legend, ncol=2, widths= c(1,0.2)))

#save plot
png("/Users/user/Library/CloudStorage/OneDrive-Personal/PhD/Chapter 3 BV female respo
nse/Scientific Reports Paper/images/Fractions behaviour reviewer.png", width = 30, hei
ght = 30, units = 'cm', res = 300)
# Make plot
fraction_grid
dev.off()

#save legend (manually added to Fig 3)
png("Fractions legend.png", width = 20, height = 15, units = 'cm', res = 300)
as_ggplot(fraction_legend)
dev.off()

```

```

#fig 3c
protein<-read.csv("urinary_fractions_protein_data.csv")

#make sex and fraction factors
protein$fraction<-factor(protein$fraction, levels=c('intact', 'lmw1','lmw2','hmw'))
protein$donor.sex<-factor(protein$donor.sex, levels=c('m','f'))

#make plot
protein.gg<-ggplot(data=protein, aes(x=fraction, y=av.protein)) +
  geom_boxplot(aes(fill=donor.sex), outlier.shape = NA) +
  geom_point(position=position_jitterdodge(),aes(shape=donor.sex), stroke=1.5, colour
="black", size=3) +
  theme_minimal() + ylab("Urinary protein (µg / mL)") +
  theme(axis.title.x = element_blank(),
        axis.text.x = element_text(size=28, face= "bold", colour="black"),
        axis.title.y = element_text(face="bold", size=28),
        axis.text.y = element_text(size=24),
        panel.grid.minor=element_blank(),
        panel.grid.major = element_blank(),axis.line = element_line(colour = "black",
linewidth = 0.5),
        axis.line.x = element_blank()) +
  scale_x_discrete(breaks=c("intact", "lmw1", "lmw2", "hmw"),
                  labels=c("Intact", "LMW", "LMW2", "HMW")) +
  scale_shape_manual(name=NULL, breaks=c('m', 'f'), labels=c("Male","Female"),
                    values=c(16,18)) +
  scale_fill_manual(name=NULL, breaks=c('m', 'f'), labels=c("Male","Female"),
                   values=c("#e34a33","#fdbb84")) +
  theme(legend.text = element_text(size=28, face="bold"),
        legend.key.size = unit(2,'cm'))

protein.gg2<-protein.gg + theme(legend.position = 'none')
protien_legend<-get_legend(protein.gg)
#save plot
png("Fractions protein.png", width = 28, height = 20, units = 'cm', res = 300)
# Make plot
protein.gg2
dev.off()

#save legend
png("Fractions protein legend.png", width = 20, height = 15, units = 'cm', res = 300)
as_ggplot(protien_legend)
dev.off()

```