Player perceptions of gender balance

Introduction

The aim of this survey is to identify what people's perceptions are regarding the balance of dialogue between male and female characters in role-playing video games. This is part of a larger project that involves collecting a corpus of video game dialogue and making objective measures of gender bias. We've found that almost all games have more male dialogue than female dialogue, even when the main protagonists are female.

The purpose of the current study is to see whether the objective measures of gender bias align with gamer's subjective intuitions. This will help us frame the added value of objective corpus-based measures.

We have objective findings from 50 video games that demonstrate a clear bias against female dialogue. The main findings are:

- Games have an average of around 30% female dialogue.
- Very few games have more than 50% female dialogue.
- The proportion of female dialogue is increasing over time.

The first two statistics surprised us, suggesting that the severity of the bias isn't well known. But we want to be able to demonstrate this objectively.

We now want to compare this objective measure with people's subjective perceptions. The survey asks participants to guess what the main findings are. However, people may be aware of an imbalance without thinking it is necessarily a problem. So there is also a section that asks whether the participant thinks that male and female characters are under represented, and whether the representation needs to be improved.

We expect that people's intuitions may be influenced by their experience with role-playing games and identity (age category, gender). Therefore, we collected basic information about these aspects to control for these influences.

This study gained ethical approval from Cardiff University's ENCAP Research Ethics Committee (SREC reference: ENCAP/Roberts/17-01-2022).

Methodology

188 participants were recruited via online platforms including twitter, facebook, reddit (r/JRPG, r/rpg_gamers, r/FinalFantasy), Cardiff University yammer, Cardiff University Gaming Society Discord, and University of Glasgow Games and Gaming Lab. They followed a link that redirected them randomly to one of two surveys. The surveys were identical, except one asked questions framed in terms of the proportion of female dialogue, and one asked questions in terms of the proportion of male dialogue. Below, text in square brackets show how these two options varied.

Participants confirmed they were older than 18 years of age, then answered the following questions:

1. Age category
2. What is your gender? (free text)
3. How much experience do you have playing Role-Playing video games (RPGs)? (I've never played an RPG, I occasionally play RPGs, I play an RPG at least once a month, I play an RPG at least once a week)

4. What do you think? Let's say we counted all the words spoken by male and female characters (playable characters or NPCs) in a typical game. What percentage of words would be spoken by [female/male] characters? (number between 0 and 100)
5. What percentage of RPG games do you think have more [female/male] dialogue than [male/female] dialogue? (number between 0 and 100)
6. This question takes a bit more thought. Imagine we picked 10 popular RPG games. How many games do you think would have ...

- ... between 0% and 10% [female/male] dialogue?
- ... between 10% and 30% [female/male] dialogue?
- ... between 10% and 30% [female/male] dialogue?
- ... between 50% and 70% [female/male] dialogue?
- ... between 70% and 90% [female/male] dialogue?
- ... between 90% and 100% [female/male] dialogue?

5 point scales for strongly disagree, disagree, neither agree nor disagree, agree, strongly agree:

7. Do you disagree or agree with the following statements?

- [female/male] characters are under-represented in video games
- [male/female] characters are under-represented in video games
- The representation of [female/male] characters should be improved
- The representation of [male/female] characters should be improved

8. How do you think the proportion of [female/male] dialogue has changed over the last 20 years? (5 point scale: Increased a lot (more [female/male] dialogue now than before), Increased a bit, Stayed the same, Decreased a bit, Decreased a lot (less [female/male] dialogue now than before))

Responses for questions 4, 5 and 6 for the male-framed surveys were flipped to express the answers in terms of amount of female dialogue.

Below, we analyse the responses.

Analysis

Load survey data

```
library(ggplot2)
library(ggpubr)
library(cowplot)
library(viridis)
library(party)
library(rjson)
```

The survey data is in two files: one for each framing of the questions.

```
df = read.csv("surveyFemaleFraming.csv",stringsAsFactors = F)
dm = read.csv("surveyMaleFraming.csv", stringsAsFactors = F)

varNames = read.csv("varNames.csv",stringsAsFactors = F)
names(df) = varNames[match(names(df),varNames$SurveyQ),]$Label
names(dm) = varNames[match(names(dm),varNames$SurveyQ),]$Label
df = df[,names(dm)]

df$framing = "female"
dm$framing = "male"
```

```
d = rbind(df,dm)
```

Flip the responses for the male-framed condition:

```
d[d$framing=="male",]$percentFemaleDialogue = 100-d[d$framing=="male",]$percentFemaleDialogue
d[d$framing=="male",]$percentGamesMoreFemale = 100-d[d$framing=="male",]$percentGamesMoreFemale

d[d$framing=="female",]$FemaleChange20y = as.character(
  factor(d[d$framing=="female",]$FemaleChange20y,
         levels = c("Decreased a lot (less female dialogue now than before)",
                    "Decreased a bit",
                    "Stayed the same",
                    "Increased a bit",
                    "Increased a lot (more female dialogue now than before)"),
         labels = c("Decreased a lot",
                    "Decreased a bit",
                    "Stayed the same",
                    "Increased a bit",
                    "Increased a lot")))

d[d$framing=="male",]$FemaleChange20y = as.character(
  factor(d[d$framing=="male",]$FemaleChange20y,
         levels = c("Decreased a lot (less male dialogue now than before)",
                    "Decreased a bit",
                    "Stayed the same",
                    "Increased a bit",
                    "Increased a lot (more male dialogue now than before)"),
         labels = c("Increased a lot",
                    "Increased a bit",
                    "Stayed the same",
                    "Decreased a bit",
                    "Decreased a lot")))

d$FemaleChange20y = factor(d$FemaleChange20y,
  levels = c("Decreased a lot",
                    "Decreased a bit",
                    "Stayed the same",
                    "Increased a bit",
                    "Increased a lot"))
```

Quickly check the variation with framing using a t-test and a permutation test (since the distributions are skewed). The guesses of the percentage of female dialogue do not vary by framing, but the guesses of the percentage of games with more than 50% female dialogue do vary, which we address below.

```
t.test(d$percentFemaleDialogue~d$framing)
```

```
##
##  Welch Two Sample t-test
##
## data:  d$percentFemaleDialogue by d$framing
## t = 0.58012, df = 184.05, p-value = 0.5625
## alternative hypothesis: true difference in means between group female and group male is not equal to
## 95 percent confidence interval:
##  -2.429237  4.452820
## sample estimates:
```

```
## mean in group female   mean in group male
##             31.36735             30.35556
```

```
trueDiff = diff(tapply(d$percentFemaleDialogue,d$framing,mean))
perm = function(){
  diff(tapply(sample(d$percentFemaleDialogue),d$framing,mean))
}
permDiff = replicate(1000,perm())
sum(permDiff>trueDiff)/1000
```

```
## [1] 0.719
```

```
t.test(d$percentGamesMoreFemale~d$framing)
```

```
##
##  Welch Two Sample t-test
##
## data:  d$percentGamesMoreFemale by d$framing
## t = -3.0485, df = 172.11, p-value = 0.002663
## alternative hypothesis: true difference in means between group female and group male is not equal to
## 95 percent confidence interval:
##  -17.130426  -3.665492
## sample estimates:
## mean in group female   mean in group male
##             19.10204             29.50000
```

```
trueDiff2 = diff(tapply(d$percentGamesMoreFemale,d$framing,mean))
perm2 = function(){
  diff(tapply(sample(d$percentGamesMoreFemale),d$framing,mean))
}
permDiff2 = replicate(1000,perm2())
sum(permDiff2>trueDiff2)/1000
```

```
## [1] 0.001
```

Flip the responses for question 6. We assume that the participants have an accurate, symmetrical understanding of the scale.

```
fcats = c("f10","f30","f50","f70","f90","f100")
d[d$framing=="male",fcats] = d[d$framing=="male",rev(fcats)]
```

Participants were given a free text box to describe their gender. We re-code categories that are obviously equivalent to 'male' and 'female' below (at least in terms of gender, rather than sex). There were only 4 participants who did not identify as either male or female. This is too few to analyses statistically. While their responses are analysed in general, they are not part of the analyses that focus on responses from male and female participants.

```
genderCoding = read.csv("genderCodes.csv",stringsAsFactors = F)
d$gender = tolower(d$genderString)
d[d$gender %in% genderCoding$string,]$gender = genderCoding[
      match(d[d$gender %in% genderCoding$string,]$gender,
      genderCoding$string),]$code
d$gender[d$gender==""] = NA
d$gender2 = d$gender
d[!is.na(d$gender2) & !(d$gender2 %in% c("female","male")),]$gender2 = "other"
```

Overall means:

```
percentFemaleDialogueGuessMean = mean(d$percentFemaleDialogue)
percentFemaleDialogueGuessSD = sd(d$percentFemaleDialogue)
percentGamesMore50FemaleMean = mean(d$percentGamesMoreFemale)
percentGamesMore50FemaleSD = sd(d$percentGamesMoreFemale)
```

The mean guess for the average percentage of female dialogue was 30.88% (sd = 12.02).

The mean guess for the percentage of games with more than 50% female dialogue was 24.08% (sd = 23.68).

```
cat(nrow(d),file="../../results/latexStats/Survey_N.tex")
cat(round(percentFemaleDialogueGuessMean),
    file="../../results/latexStats/Survey_percentFemaleDialogueGuessMean.tex")
cat(round(percentFemaleDialogueGuessSD,2),
    file="../../results/latexStats/Survey_percentFemaleDialogueGuessSD.tex")
cat(round(percentGamesMore50FemaleMean),
    file="../../results/latexStats/Survey_percentGamesMore50FemaleMean.tex")
cat(round(percentGamesMore50FemaleSD,2),
    file="../../results/latexStats/Survey_percentGamesMore50FemaleSD.tex")
```

Load game data

Obtain the true values from the corpus

```
folders = list.dirs("../../data", recursive = T)
folders = folders[sapply(folders,function(X){
  "stats_by_character.csv" %in% list.files(X)
})]


trueFemaleDialogue = c()
for(folder in folders){
  alternativeMeasure = FALSE
  js = fromJSON(file = paste0(folder,"/meta.json"))
  if(!is.null(js$alternativeMeasure)){
    alternativeMeasure = js$alternativeMeasure
  }
  if(!alternativeMeasure){
    stats = read.csv(paste0(folder,"/stats.csv"),stringsAsFactors = F)
    fprop = stats[stats$group=="female",]$words/
      (stats[stats$group=="female",]$words +
         stats[stats$group=="male",]$words)
    trueFemaleDialogue = c(trueFemaleDialogue,fprop)
  }
}
trueFemaleDialogueDist = trueFemaleDialogue[!is.na(trueFemaleDialogue)]
truePercentFemaleDialogue = mean(trueFemaleDialogueDist)
truePercentGamesMoreThan50 = sum(trueFemaleDialogueDist>=0.5)/length(trueFemaleDialogueDist)
```

Statistical test: is the mean guess significantly different from the true value?

```
mainTestPercentFemaleDialouge = t.test(d$percentFemaleDialogue,
                                       mu=truePercentFemaleDialogue*100)
mpx = "p < 0.0001"
if(mainTestPercentFemaleDialouge$p.value>0.0001){
  mpx = paste0("p = ",round(mainTestPercentFemaleDialouge$p.value,4))
}
```

```
cat(paste0("t = ",round(mainTestPercentFemaleDialouge$statistic,2),", ",mpx),
    file = "../../results/latexStats/Survey_percentFemaleDialogue_Stat.tex")
mainTestPercentFemaleDialouge
```

```
##
##  One Sample t-test
##
## data:  d$percentFemaleDialogue
## t = -0.68151, df = 187, p-value = 0.4964
## alternative hypothesis: true mean is not equal to 31.48047
## 95 percent confidence interval:
##  29.15347 32.61249
## sample estimates:
## mean of x
##  30.88298
```

```
mainTestPercentGamesMoreThan50 = t.test(d$percentGamesMoreFemale,
                                        mu=truePercentGamesMoreThan50*100)
mpx = "p < 0.0001"
if(mainTestPercentGamesMoreThan50$p.value>0.0001){
  mpx = paste0("p = ",round(mainTestPercentGamesMoreThan50$p.value,4))
}
cat(paste0("t = ",round(mainTestPercentGamesMoreThan50$statistic,2),", ",mpx),
    file = "../../results/latexStats/Survey_percentMoreThan50_Stat.tex")
mainTestPercentGamesMoreThan50
```

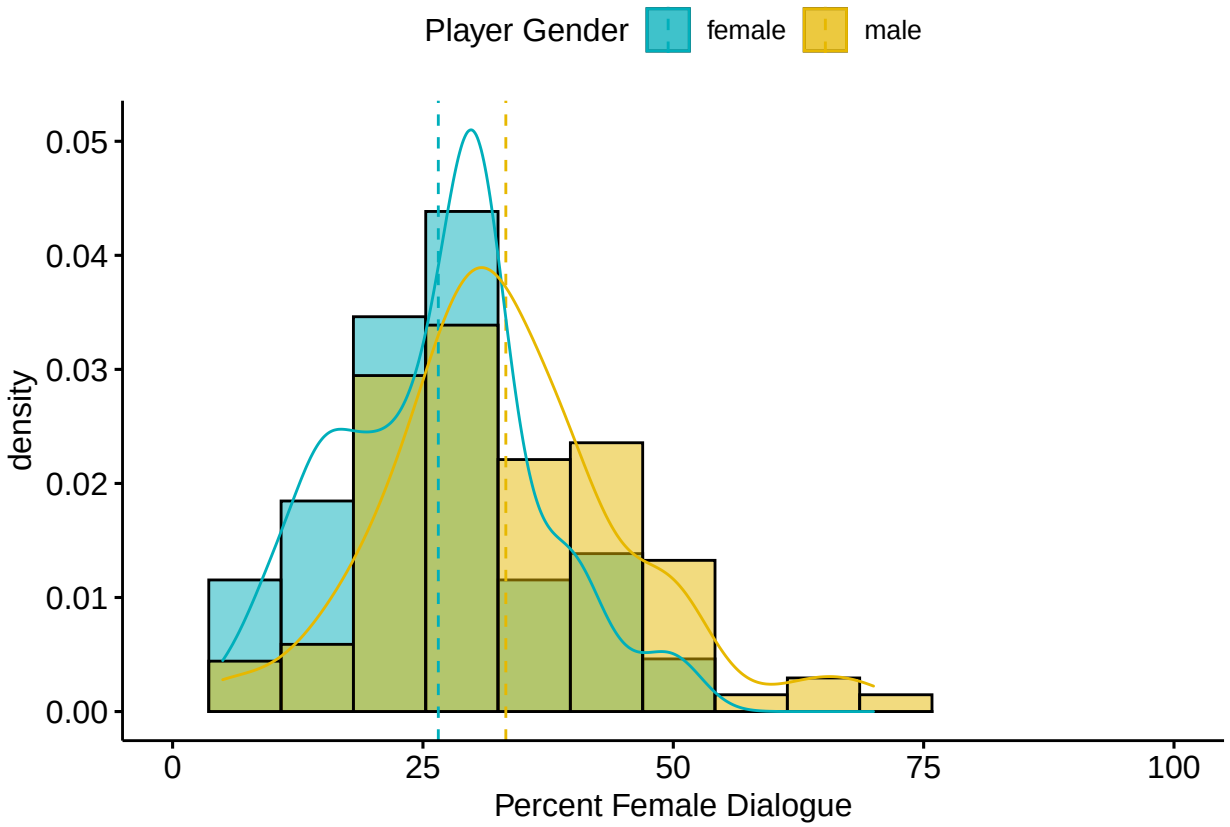
```
##
##  One Sample t-test
##
## data:  d$percentGamesMoreFemale
## t = 10.468, df = 187, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 6
## 95 percent confidence interval:
##  20.67269 27.48689
## sample estimates:
## mean of x
##  24.07979
```

Variation by gender

Below we visualise the data in various ways. In particular, we are interested in whether the estimates vary by the gender of the participant.

Distribution of guesses for the percentage of female dialogue:

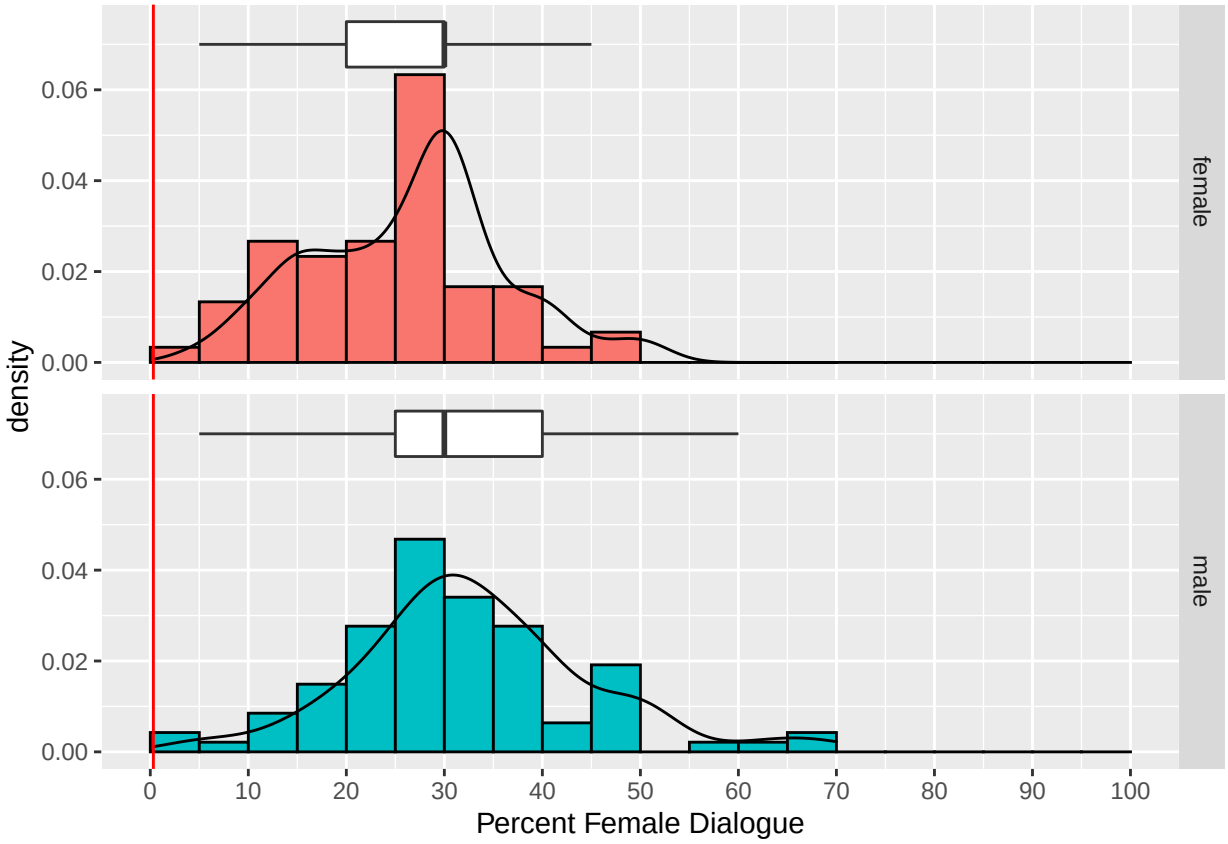
```
gghistogram(
  d[d$gender %in% c("female","male"),],
  bins = 10,
  x = "percentFemaleDialogue",
  y = "..density..",
  add = "mean", #rug = TRUE,
  fill = "gender", palette = c("#00AFBB", "#E7B800"),
  add_density = TRUE,
  xlim=c(0,100),
  legend.title="Player Gender") +
  xlab("Percent Female Dialogue")
```




```
ggplot(d[d$gender %in% c("female","male"),],
       aes(x=percentFemaleDialogue)) +
  geom_histogram(aes(y = ..density.., fill=gender),breaks=seq(0,100,5), color='black')+
  geom_density() +
  xlim(0,100) +
  facet_grid(rows=vars(gender)) +
  scale_x_continuous(breaks = seq(0,100,10)) +
  geom_boxplot(width=0.01,position=position_nudge(y=0.07),outlier.alpha = 0) +
  xlab("Percent Female Dialogue") +
```

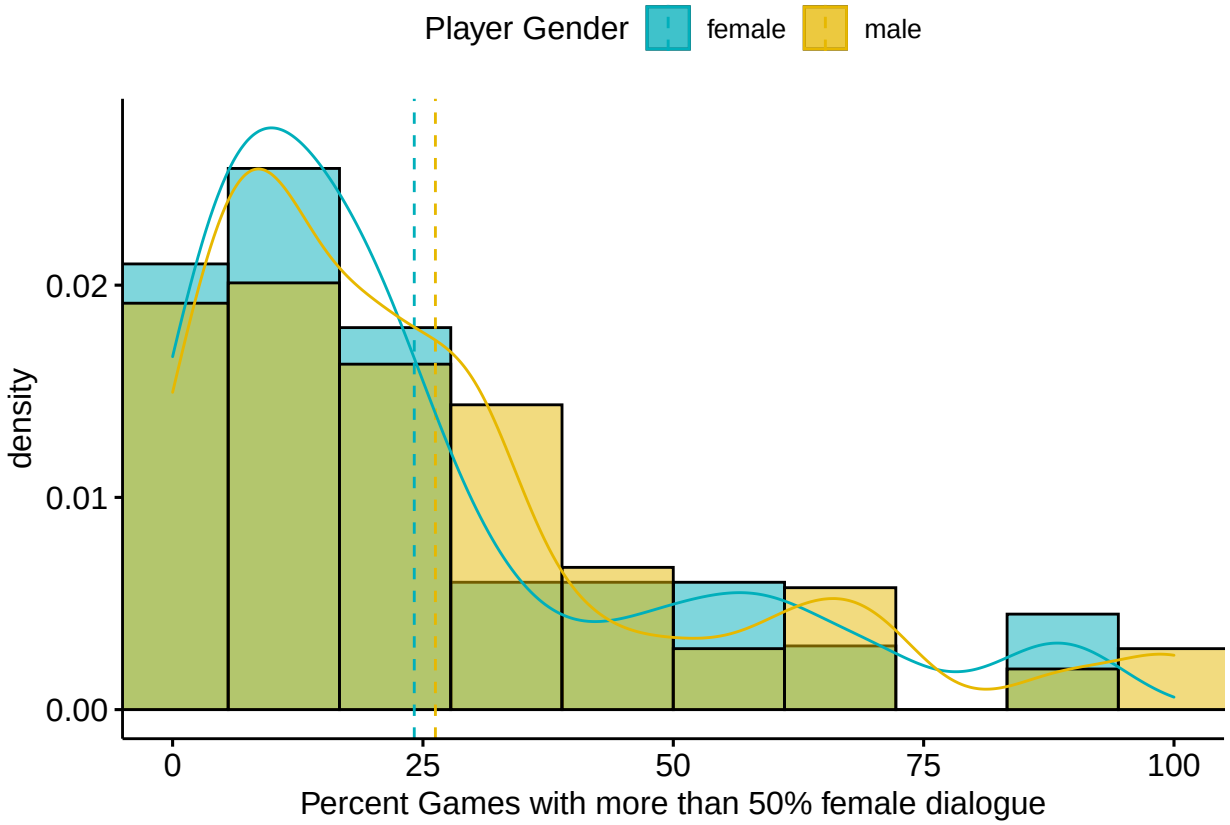
```
  geom_vline(xintercept = truePercentFemaleDialogue,color="red") +
  theme(legend.position = "none")
```

```
## Scale for 'x' is already present. Adding another scale for 'x', which will
## replace the existing scale.
```



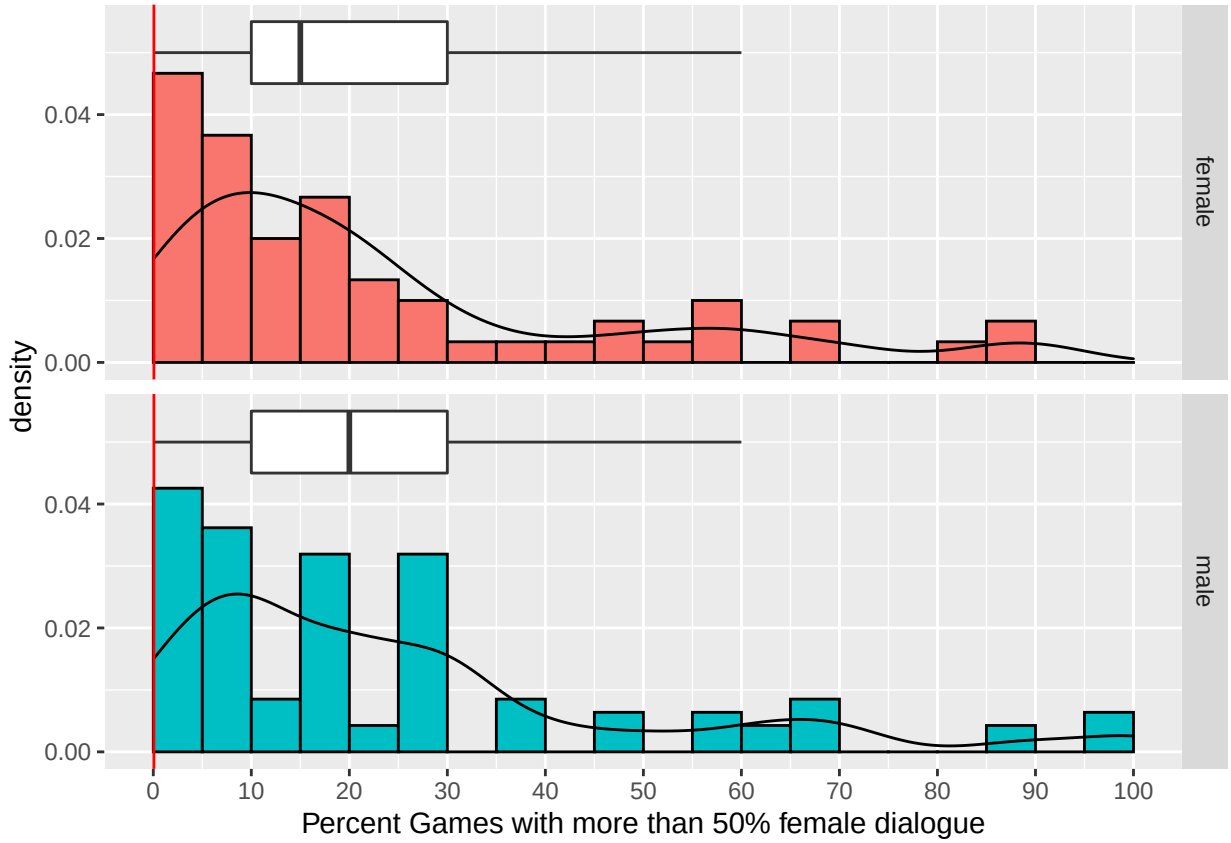
Distribution of guesses for the percentage of games with more than 50% female dialogue:

```
gghistogram(
  d[d$gender %in% c("female","male"),],
  bins = 10,
  x = "percentGamesMoreFemale",
  y = "..density..",
  add = "mean", #rug = TRUE,
  fill = "gender", palette = c("#00AFBB", "#E7B800"),
  add_density = TRUE,
  xlim=c(0,100),
  legend.title="Player Gender") +
  xlab("Percent Games with more than 50% female dialogue")
```




```
ggplot(d[d$gender %in% c("female","male"),],
       aes(x=percentGamesMoreFemale)) +
  geom_histogram(aes(y = ..density.., fill=gender),breaks=seq(0,100,5), color='black')+
  geom_density() +
  xlim(0,100) +
  facet_grid(rows=vars(gender)) +
  scale_x_continuous(breaks = seq(0,100,10)) +
  geom_boxplot(width=0.01,position=position_nudge(y=0.05),outlier.alpha = 0) +
  xlab("Percent Games with more than 50% female dialogue") +
  geom_vline(xintercept = truePercentGamesMoreThan50,color="red") +
  theme(legend.position = "none")
```

```
## Scale for 'x' is already present. Adding another scale for 'x', which will
## replace the existing scale.
```


density
0.04
0.02
0.00
female
male
0
10
20
30
40
50
60
70
80
90
100
Percent Games with more than 50% female dialogue

Variation by Age

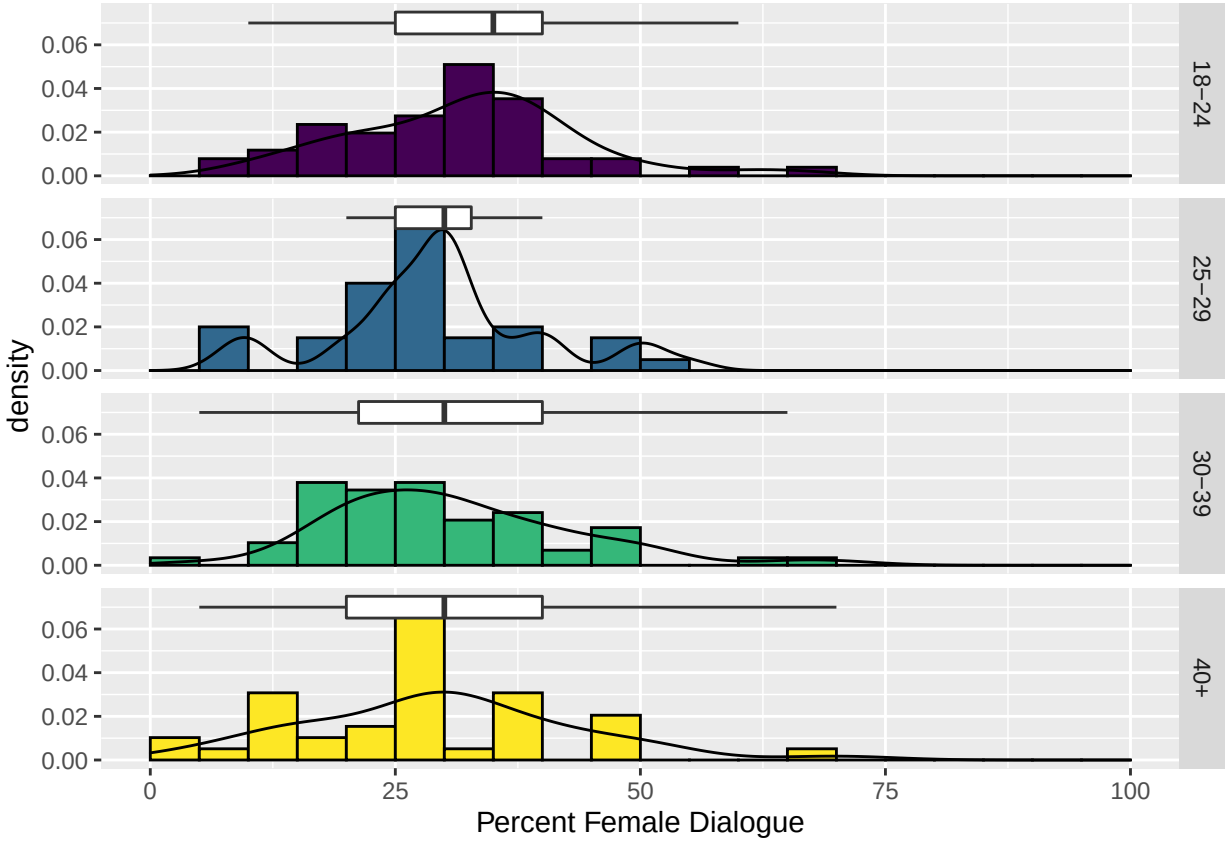
Below we plot the distribution of guesses by the age of the respondent. One hypothesis is that older players may be more likely to have played earlier games compared to younger players (or less likely to have played newer games). Given that earlier games have less female dialogue, guesses of older players may be lower than younger players. Having said this, accessing games has become much easier in the last 10 years, so we may not expect a strong result.

There were relatively few participants over 40, so these are collapsed into a "+40" category.

The red dotted line shows the actual proportion in the corpus.

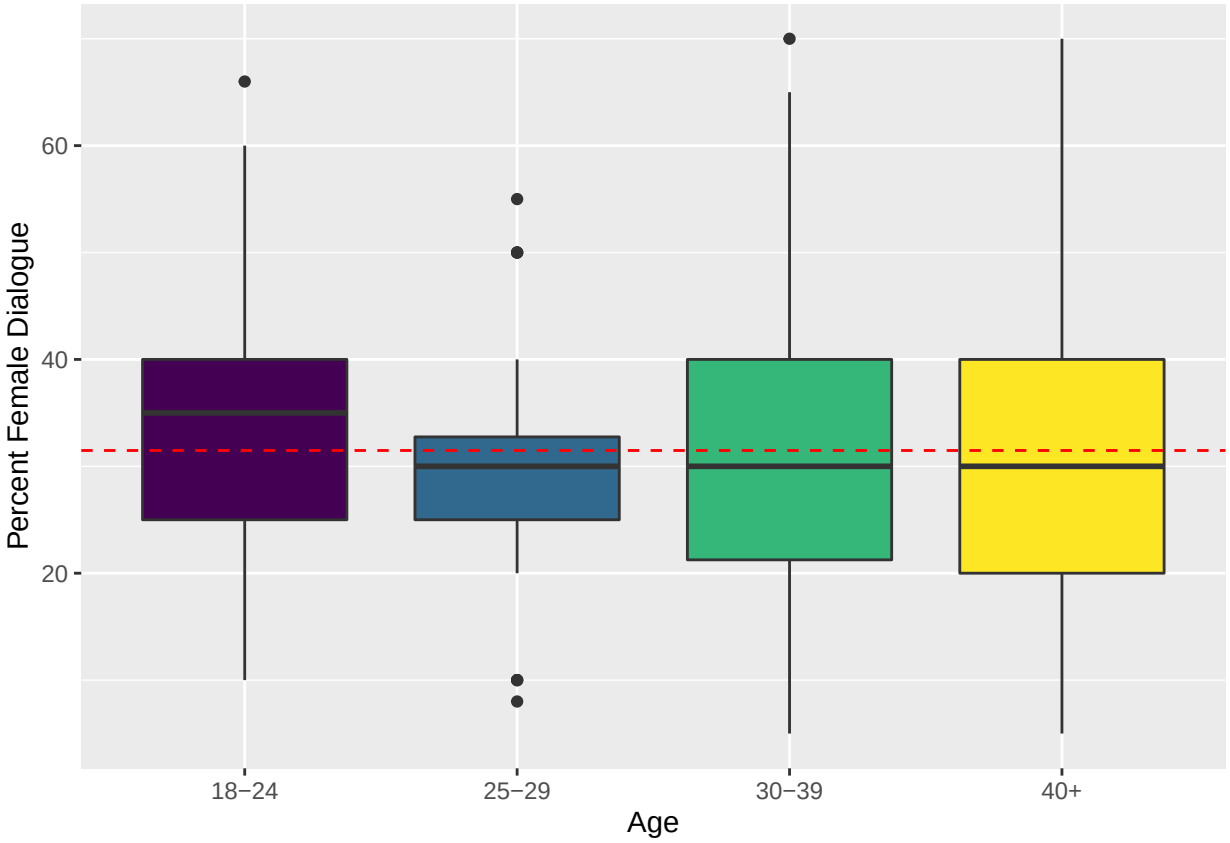
```
d$age2 = d$age
d$age2[d$age2 %in% c("40-49","50-59","60-69","70+")] = "40+"
d$age2 = factor(d$age2,ordered = T)

ggplot(d, aes(x=percentFemaleDialogue)) +
  geom_histogram(aes(y = ..density.., fill=age2),breaks=seq(0,100,5), color='black')+
  geom_density() +
  xlim(0,100) +
  facet_grid(rows=vars(age2)) +
  geom_boxplot(width=0.01,position=position_nudge(y=0.07),outlier.alpha = 0) +
  xlab("Percent Female Dialogue") +
  theme(legend.position = "none")
```



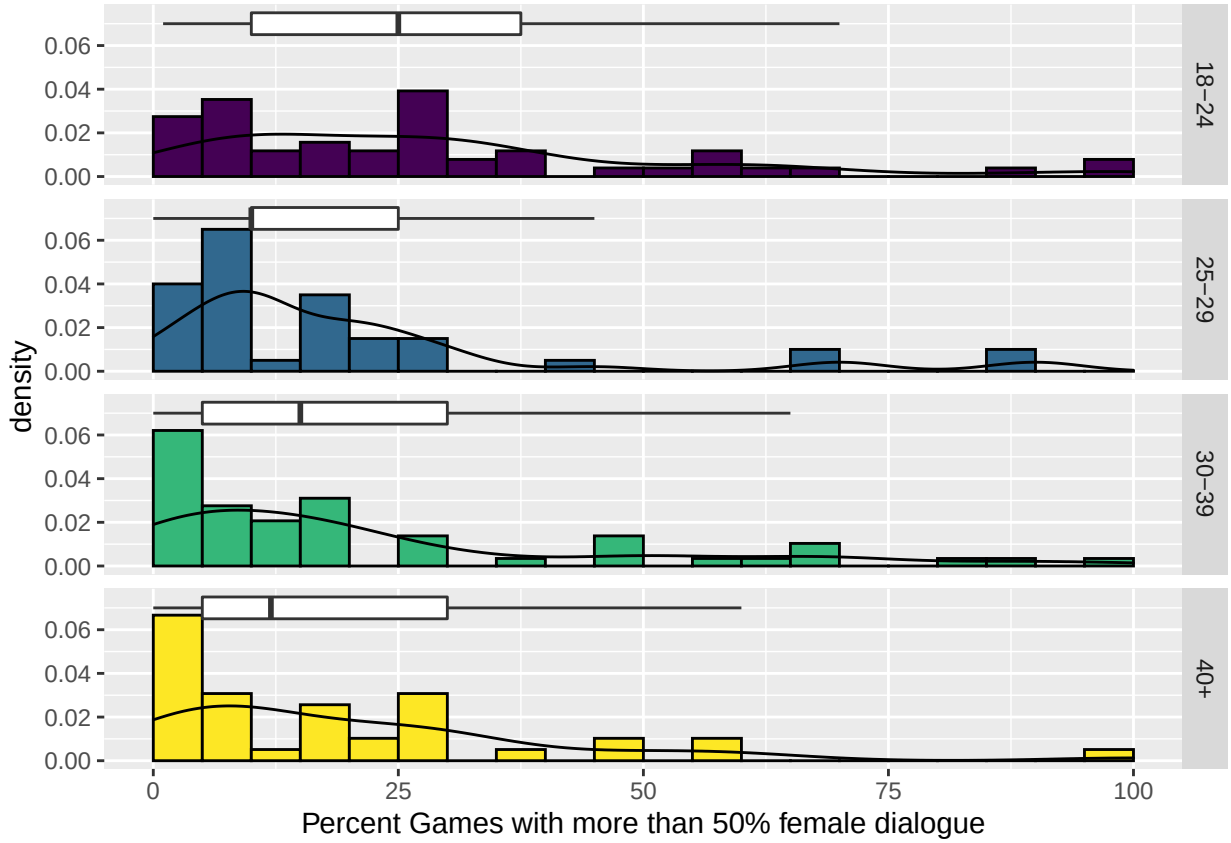

```
ggplot(d,aes(y=percentFemaleDialogue,x=age2,fill=age2)) +
  geom_boxplot() +
  xlab("Age") + ylab("Percent Female Dialogue") +
```

```
  theme(legend.position = "none") +
  geom_hline(yintercept = truePercentFemaleDialogue*100, color="red",linetype="dashed")
```

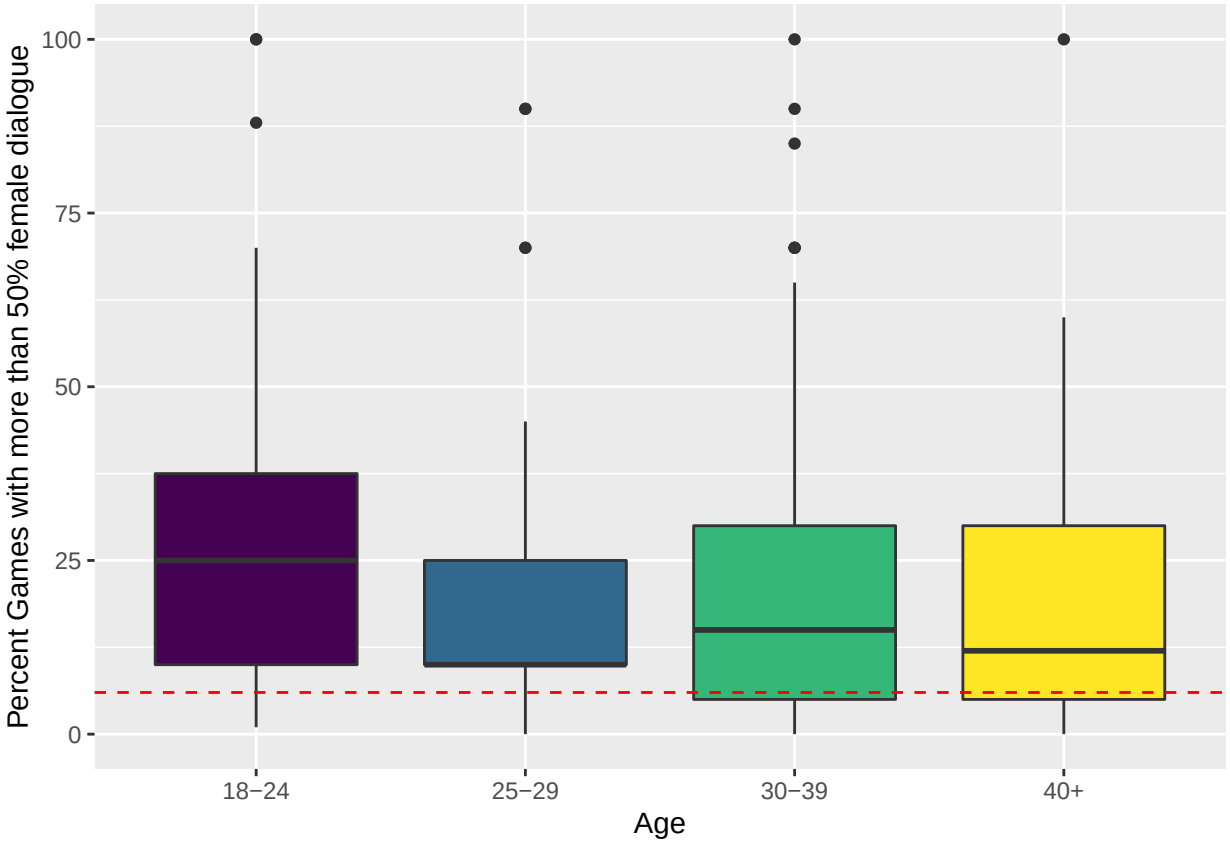


Similar plots for the proportion of games with more than 50% dialogue (red dotted line shows the value in the corpus).

```
ggplot(d, aes(x=percentGamesMoreFemale)) +
  geom_histogram(aes(y = ..density.., fill=age2),breaks=seq(0,100,5), color='black')+
  geom_density() +
  xlim(0,100) +
  facet_grid(rows=vars(age2)) +
  geom_boxplot(width=0.01,position=position_nudge(y=0.07),outlier.alpha = 0) +
  xlab("Percent Games with more than 50% female dialogue") +
  theme(legend.position = "none")
```




```
ggplot(d,aes(y=percentGamesMoreFemale,x=age2,fill=age2)) +
  geom_boxplot() +
  xlab("Age") + ylab("Percent Games with more than 50% female dialogue") +
  theme(legend.position = "none") +
  geom_hline(yintercept = truePercentGamesMoreThan50*100, color="red",linetype="dashed")
```


Percent Games with more than 50% female dialogue
100
75
50
25
0
18−24
25−29
30−39
40+
Age

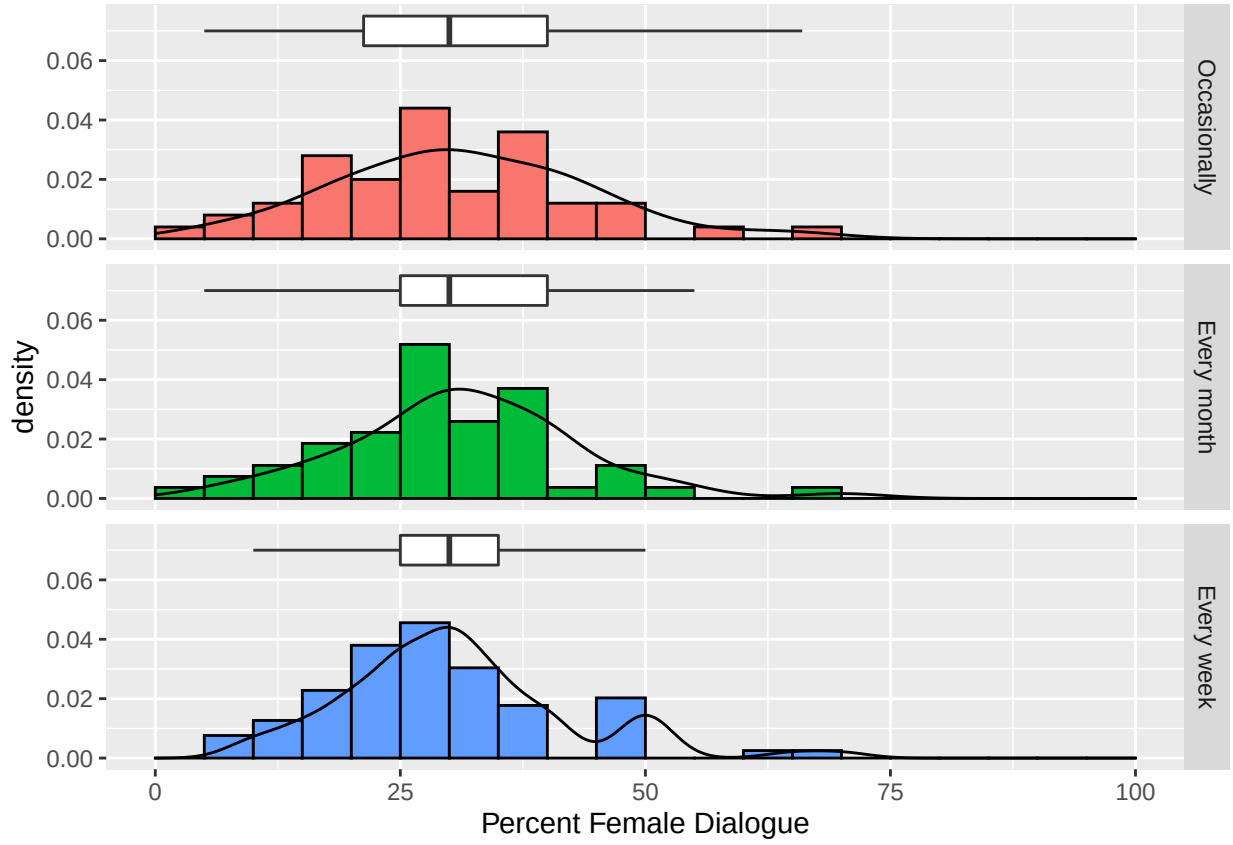
Variation by experience

Below we plot the data by experience. One hypothesis is that more experienced players will have more accurate guesses. Since many players over-estimate the measures, this might lead to more experienced players having lower guesses. It might also predict a smaller variation for more experienced players

Female dialogue

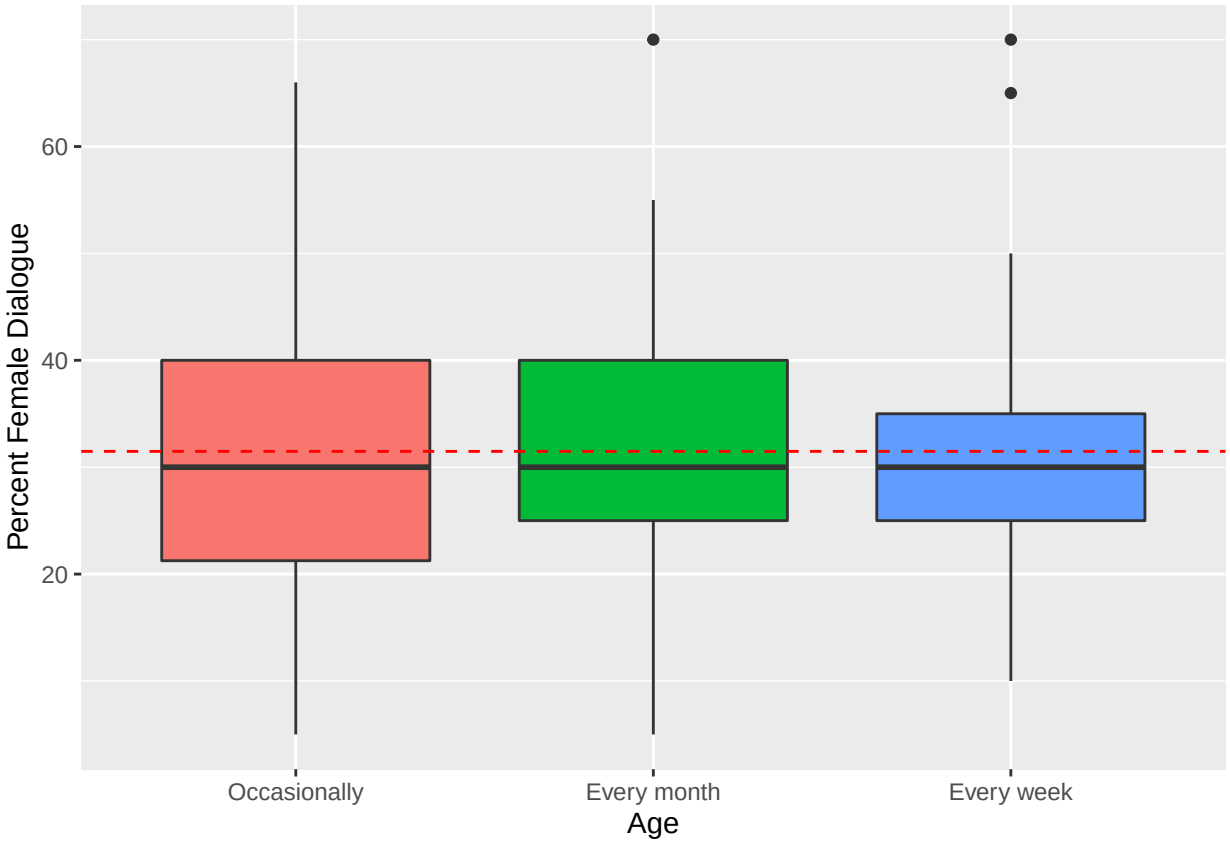
```
d$rpgExperience = factor(d$rpgExperience,
    levels = c("I've never played an RPG", "I occasionally play RPGs",
        'I play an RPG at least once a month',"I play an RPG at least once a week"),
        labels = c("Never","Occasionally","Every month","Every week"))

ggplot(d[d$rpgExperience!="Never",],
    aes(x=percentFemaleDialogue)) +
  geom_histogram(aes(y = ..density.., fill=rpgExperience),breaks=seq(0,100,5), color='black')+
  geom_density() +
  xlim(0,100) +
  facet_grid(rows=vars(rpgExperience)) +
  geom_boxplot(width=0.01,position=position_nudge(y=0.07),outlier.alpha = 0) +
  xlab("Percent Female Dialogue") +
  theme(legend.position = "none")
```



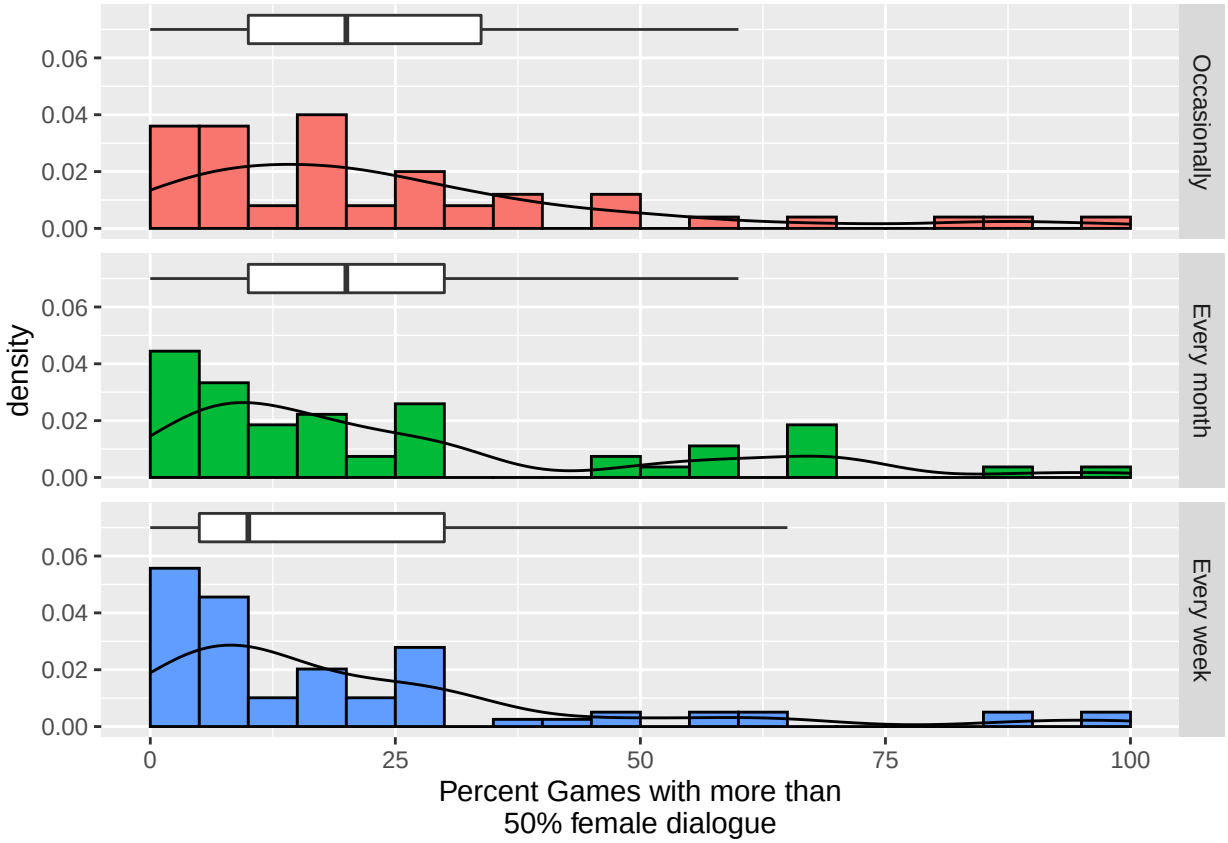

```
ggplot(d[d$rpgExperience!="Never",],
    aes(y=percentFemaleDialogue,x=rpgExperience,fill=rpgExperience)) +
  geom_boxplot() +
```

```
  xlab("Age") + ylab("Percent Female Dialogue") +
  theme(legend.position = "none") +
  geom_hline(yintercept = truePercentFemaleDialogue*100, color="red",linetype="dashed")
```

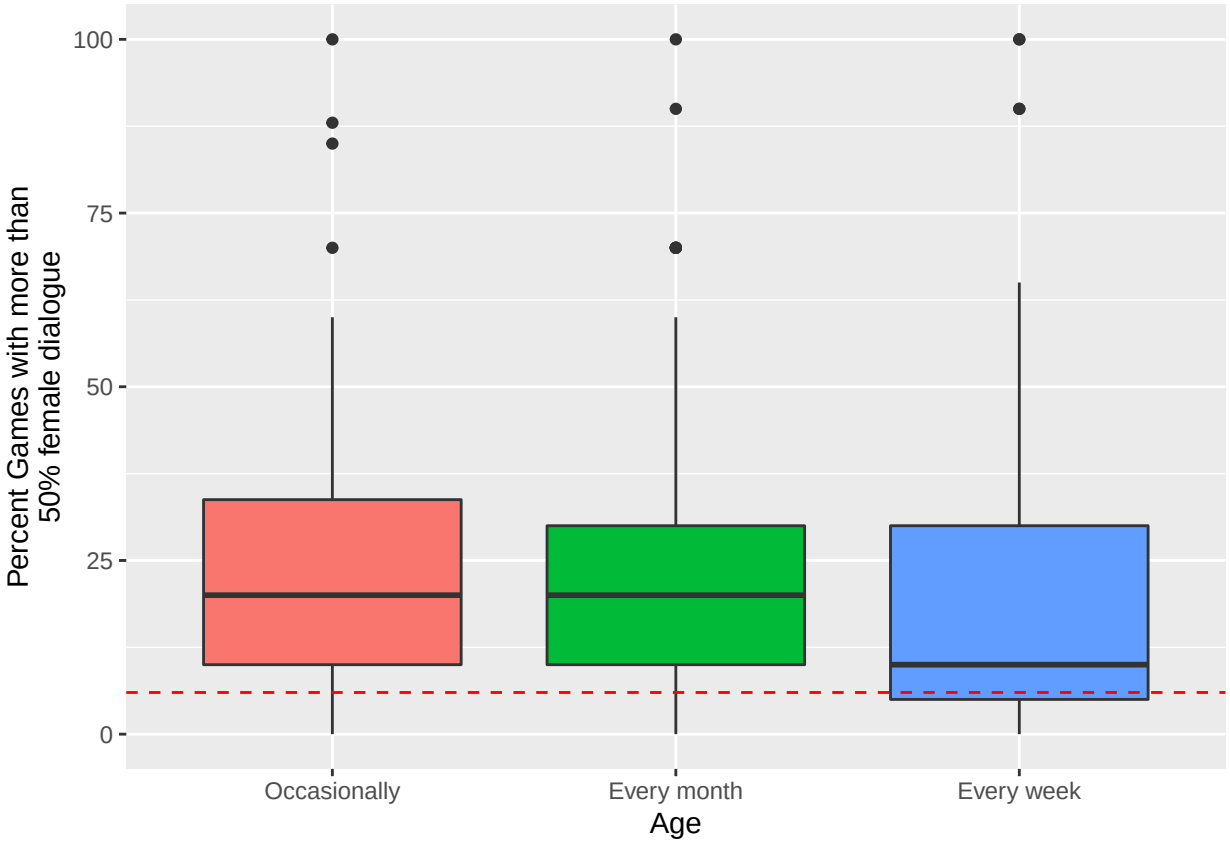


Games with more than 50% female dialogue

```
ggplot(d[d$rpgExperience!="Never",],
    aes(x=percentGamesMoreFemale)) +
  geom_histogram(aes(y = ..density.., fill=rpgExperience),breaks=seq(0,100,5), color='black')+
  geom_density() +
  xlim(0,100) +
  facet_grid(rows=vars(rpgExperience)) +
  geom_boxplot(width=0.01,position=position_nudge(y=0.07),outlier.alpha = 0) +
  xlab("Percent Games with more than\n50% female dialogue") +
  theme(legend.position = "none")
```




```
ggplot(d[d$rpgExperience!="Never",],
    aes(y=percentGamesMoreFemale,x=rpgExperience,fill=rpgExperience)) +
  geom_boxplot() +
  xlab("Age") + ylab("Percent Games with more than\n50% female dialogue") +
  theme(legend.position = "none") +
  geom_hline(yintercept = truePercentGamesMoreThan50*100, color="red",linetype="dashed")
```


Percent Games with more than
50% female dialogue
100
75
50
25
0
Occasionally
Every month
Every week
Age

Overall predictions

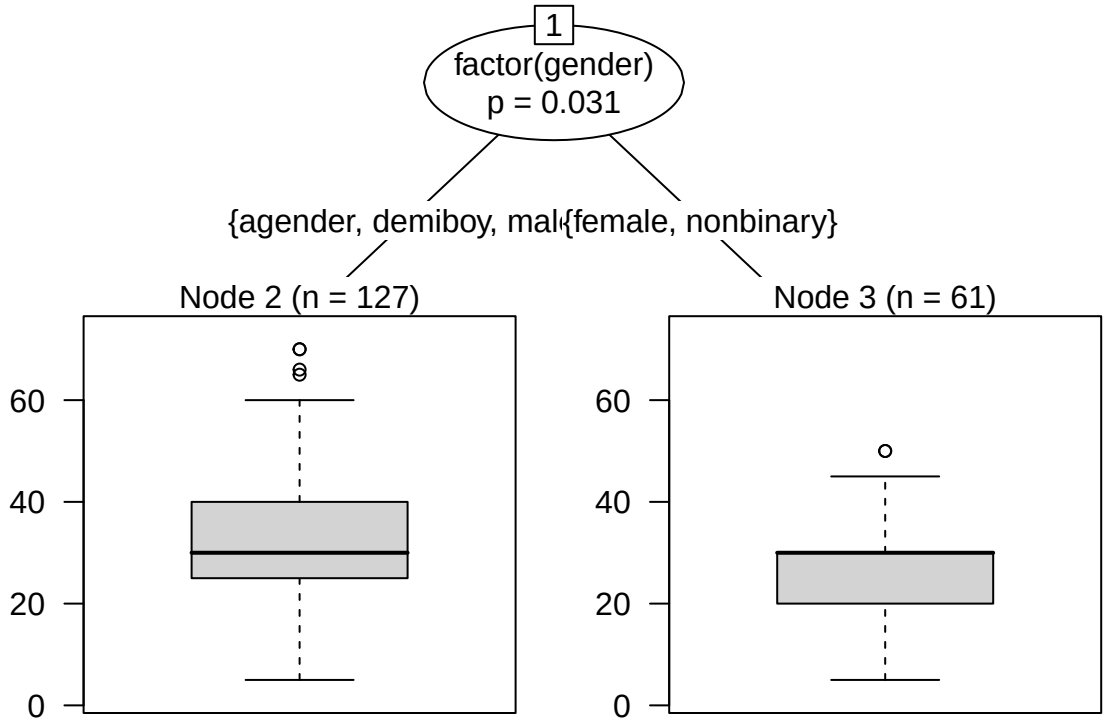
We use a regression model to identify what factors predict player guesses. For guesses of the percentage of female dialogue, player gender is the only significant predicting factor:

```
m0 = lm(percentFemaleDialogue ~
          age2 + framing + gender2 + rpgExperience,
        data=d)
summary(m0)
```

```
##
## Call:
## lm(formula = percentFemaleDialogue ~ age2 + framing + gender2 +
##     rpgExperience, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -28.667  -7.250  -0.685   4.915  38.285
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                17.12566    6.05862   2.827 0.005354 **
## age2.L                     -2.32361    1.89664  -1.225 0.222476
## age2.Q                      0.91678    1.86233   0.492 0.623255
## age2.C                     -1.56501    1.83536  -0.853 0.395202
## framingmale                 0.07414    1.89685   0.039 0.968875
## gender2male                 6.62937    1.90247   3.485 0.000649 ***
## gender2other                2.54725    6.00804   0.424 0.672201
## rpgExperienceOccasionally   9.83990    6.11429   1.609 0.109675
## rpgExperienceEvery month    9.56655    6.05323   1.580 0.116150
## rpgExperienceEvery week     9.40978    5.98262   1.573 0.117888
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.27 on 148 degrees of freedom
##   (30 observations deleted due to missingness)
## Multiple R-squared:  0.1145, Adjusted R-squared:  0.06062
## F-statistic: 2.126 on 9 and 148 DF,  p-value: 0.0307
```

We can use a decision tree that takes all the gender categories into account. This splits the data into two groups. The first is 'female' and 'nonbinary', who on average guess lower than the second group, 'male', 'agender' and 'demiboy'.

```
tree = ctree(percentFemaleDialogue ~
          age2 + factor(framing) + factor(gender) + rpgExperience,
          data=d)
plot(tree)
```



Question framing seems to be the main predictor for percentage of games with more than 50% female dialogue. The guesses for the proportion of games with more than 50% dialogue are higher for participants asked to estimate the proportion of games with more *male* dialogue than female (their guesses were then inverted for the analyses above and below). This suggests that there is a framing bias.

```
m0B = lm(percentGamesMoreFemale ~
           framing,
         data=d)
summary(m0B)
```

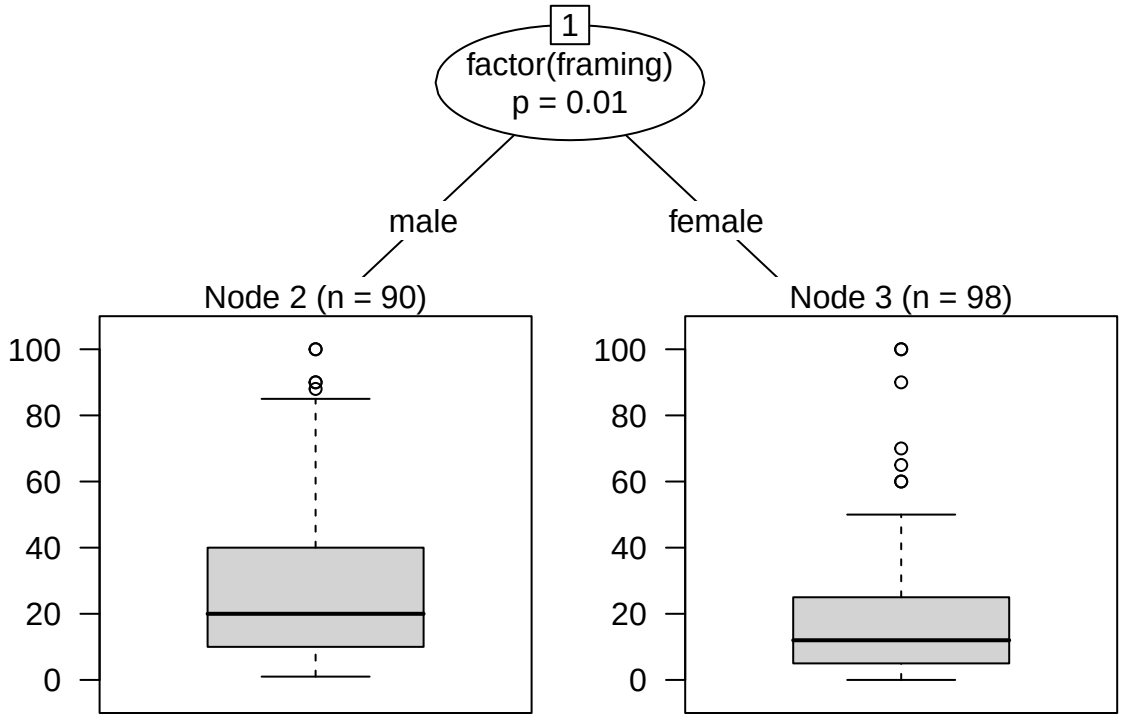
```
##
## Call:
## lm(formula = percentGamesMoreFemale ~ framing, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -28.500 -17.352  -9.102   7.048  80.898
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   19.102      2.340   8.164 4.79e-14 ***
## framingmale   10.398      3.382   3.075  0.00242 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23.16 on 186 degrees of freedom
## Multiple R-squared:  0.04837,    Adjusted R-squared:  0.04325
```

```
## F-statistic: 9.454 on 1 and 186 DF,  p-value: 0.002424
```

```
m1B = lm(percentGamesMoreFemale ~
          age2 + framing + gender2 + rpgExperience,
        data=d)
summary(m1B)
```

```
##
## Call:
## lm(formula = percentGamesMoreFemale ~ age2 + framing + gender2 +
##     rpgExperience, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.939 -15.347  -6.780   7.942  77.728
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   4.604     12.599   0.365  0.71528
## age2.L                       -2.417      3.944  -0.613  0.54097
## age2.Q                        1.743      3.873   0.450  0.65331
## age2.C                       -5.579      3.817  -1.462  0.14592
## framingmale                  12.667      3.944   3.211  0.00162 **
## gender2male                   3.916      3.956   0.990  0.32381
## gender2other                 -8.105     12.493  -0.649  0.51750
## rpgExperienceOccasionally    11.136     12.714   0.876  0.38253
## rpgExperienceEvery month     15.748     12.587   1.251  0.21287
## rpgExperienceEvery week       9.846     12.441   0.791  0.42994
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23.44 on 148 degrees of freedom
##   (30 observations deleted due to missingness)
## Multiple R-squared:  0.1129, Adjusted R-squared:  0.05899
## F-statistic: 2.094 on 9 and 148 DF,  p-value: 0.03349
```

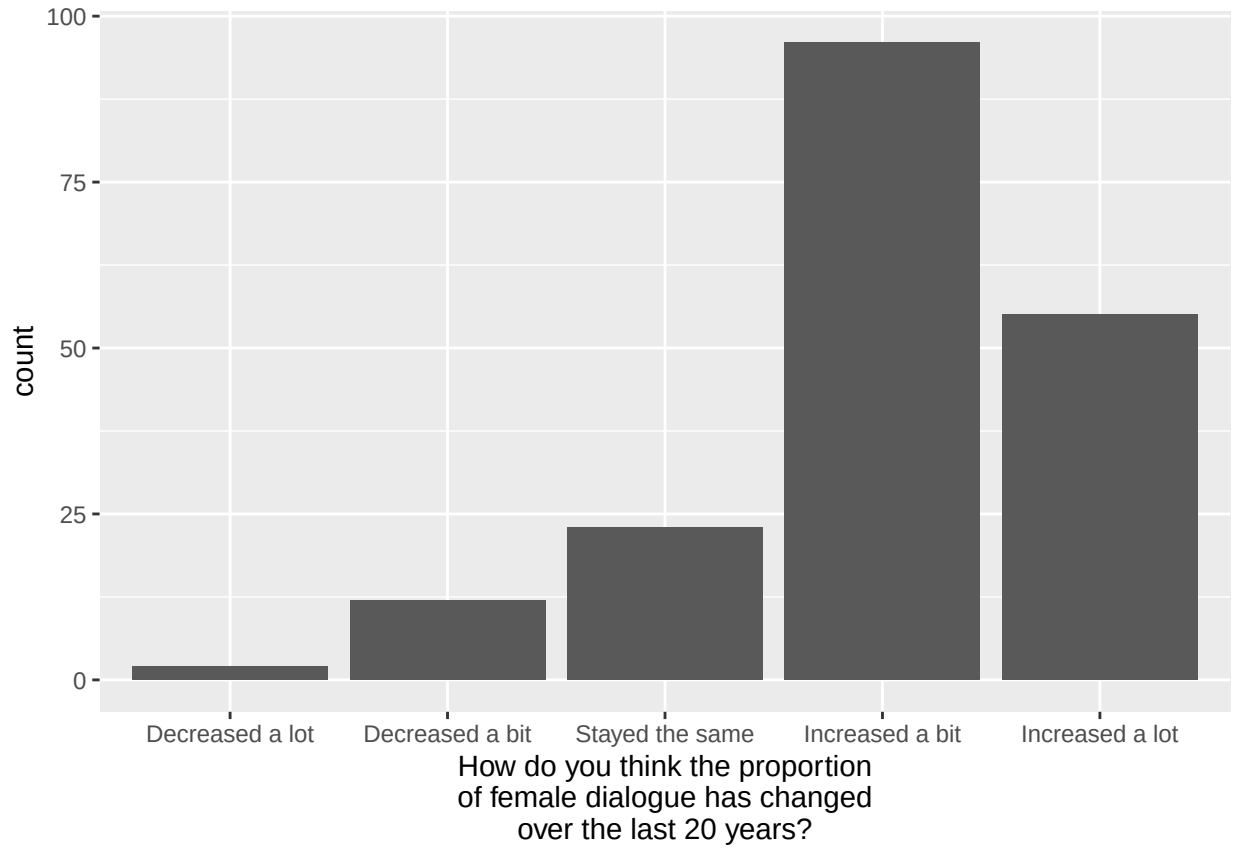
```
treeB = ctree(percentGamesMoreFemale ~
          age2 + factor(framing) + factor(gender) + rpgExperience,
          data=d)
plot(treeB)
```


1
factor(framing)
p = 0.01
male
female
Node 2 (n = 90)
Node 3 (n = 98)
100
80
60
40
20
0
100
80
60
40
20
0

Change over time

Summary of the responses to the question of "How do you think the proportion of female dialogue has changed over the last 20 years?". Most respondents say it has increased a bit.

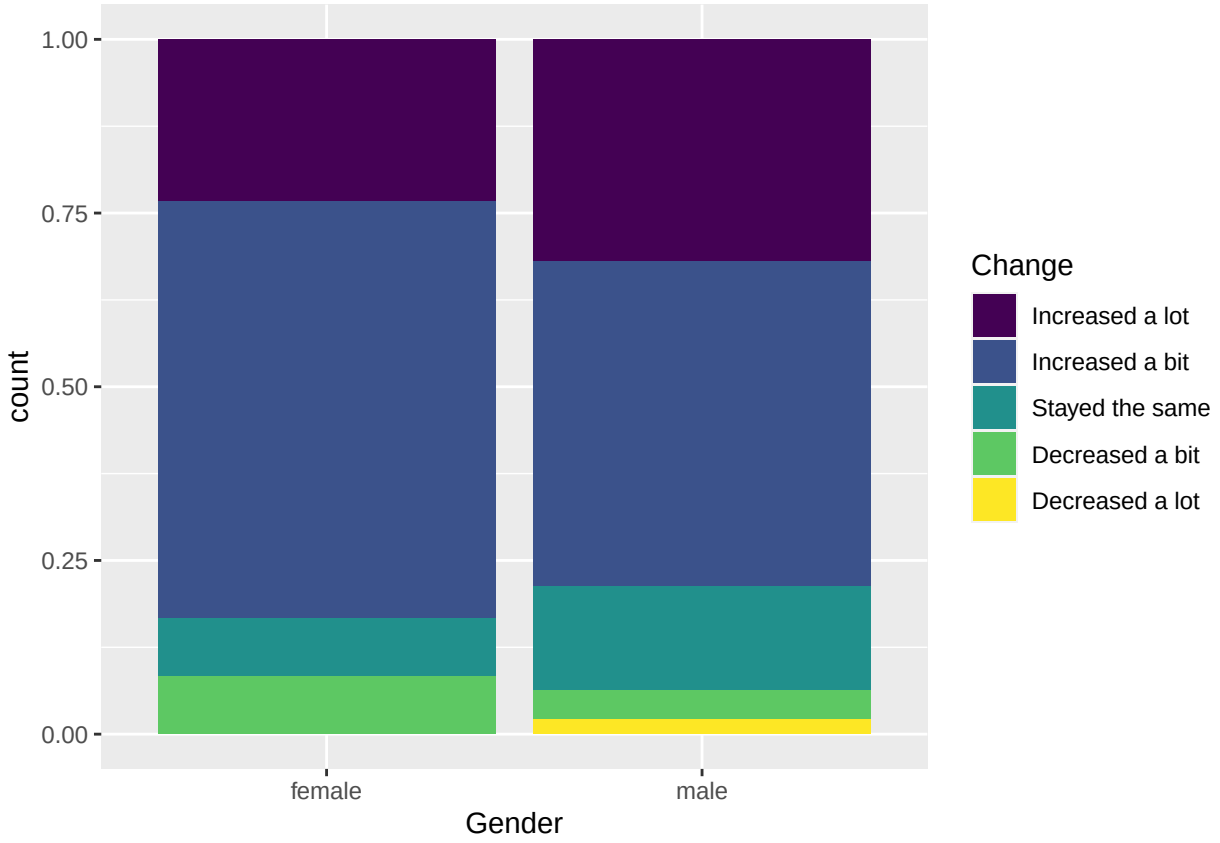
```
ggplot(d,aes(x=FemaleChange20y)) +
  geom_bar() +
  xlab("How do you think the proportion\nof female dialogue has changed\nover the last 20 years?")
```



Variation by gender:

```
d$Change = factor(d$FemaleChange20y,
  levels = c("Increased a lot",
                    "Increased a bit",
                    "Stayed the same",
                    "Decreased a bit",
                    "Decreased a lot"))

ggplot(d[d$gender2 %in% c("female","male"),],
       aes(x=gender2,fill=Change)) +
  geom_bar(position="fill") +
  xlab("Gender")+
  scale_fill_viridis_d()
```



Test whether the responses differ by gender:

```
changeTable = table(d[d$gender2 %in% c("female", "male"),]$FemaleChange20y,
                    d[d$gender2 %in% c("female", "male"),]$gender2)
changeTable
```

```
##
##                    female male
##   Decreased a lot       0    2
##   Decreased a bit       5    4
##   Stayed the same       5   14
##   Increased a bit      36   44
##   Increased a lot      14   30
```

```
fisher.test(changeTable)
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  changeTable
## p-value = 0.2326
## alternative hypothesis: two.sided
```

No significant variation by gender.

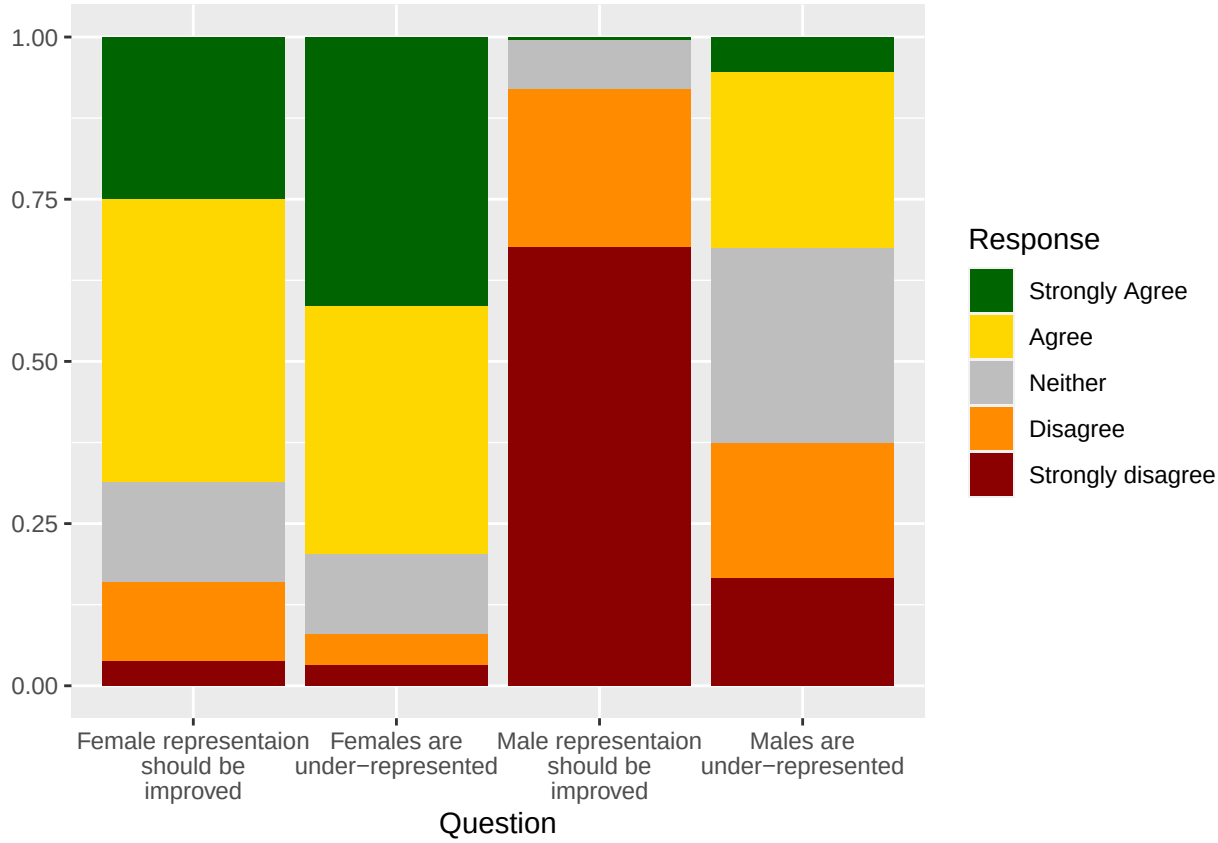
Representation

Responses for the 5-point agreement scales.

```
OpLevels = rev(c("Strongly disagree","Disagree",
           "Neither agree nor disagree","Agree",
           "Strongly Agree"))
OpLables = rev(c("Strongly disagree","Disagree",
           "Neither","Agree",
           "Strongly Agree"))
d$OpFemaleRepImprove = factor(d$OpFemaleRepImprove,levels = OpLevels,labels=OpLables,ordered=T)
d$OpMaleRepImprove = factor(d$OpMaleRepImprove,levels = OpLevels,labels=OpLables,ordered=T)
d$OpFemaleUnderRepresented = factor(d$OpFemaleUnderRepresented,levels = OpLevels,labels=OpLables,ordere
d$OpMaleUnderRepresented = factor(d$OpMaleUnderRepresented,levels = OpLevels,labels=OpLables,ordered=T)

dx = data.frame(
  Response = c(
    d$OpFemaleRepImprove,
    d$OpMaleRepImprove,
    d$OpFemaleUnderRepresented,
    d$OpMaleUnderRepresented),
  Question = c(
    rep("Females are\nunder-represented",nrow(d)),
    rep("Males are\nunder-represented",nrow(d)),
    rep("Female representaion\nshould be\nimproved",nrow(d)),
    rep("Male representaion\nshould be\nimproved",nrow(d))
  )
)

ggplot(dx[!is.na(dx$Response),],
       aes(x=Question,fill=Response)) +
  geom_bar(position="fill") +
  scale_fill_manual(values=c("darkgreen","gold","gray","darkorange","darkred")) +
  theme(axis.title.y = element_blank())
```



The majority of respondents agreed that females are under-represented and representation should be improved. The majority disagreed that male representation should be improved. And there was a mixed response to whether males are under-represented. This suggests that there may be different factors and key issues for improving female representation and male representation.

Distribution

In the main analysis of the corpus, the proportion of female dialogue as a whole was not particularly surprising to the authors. The authors were more surprised by the distribution, with very few games having more than 50% female dialogue. Question 6 of the survey included detailed questions that attempted to identify what players thought the distribution looked like.

6. This question takes a bit more thought. Imagine we picked 10 popular RPG games. How many games do you think would have ...

- ... between 0% and 10% [female/male] dialogue?
- ... between 10% and 30% [female/male] dialogue?
- ... between 10% and 30% [female/male] dialogue?
- ... between 50% and 70% [female/male] dialogue?
- ... between 70% and 90% [female/male] dialogue?
- ... between 90% and 100% [female/male] dialogue?

Participants were free to specify any number between 0 and 10 in each of their responses, so their totals could add up to more than the 10 we asked for. We need to control for this so that participants specifying the distribution for 100 games doesn't bias the estimate.

Instead of working with raw numbers, we convert each participant's responses into proportions across the scale:

```
f = d[,grepl("f[1-9]0",names(d))]
rs = rowSums(f,na.rm = T)
fNormed = f/rowSums(f,na.rm=T)
fNormed[is.na(fNormed)] = 0
```

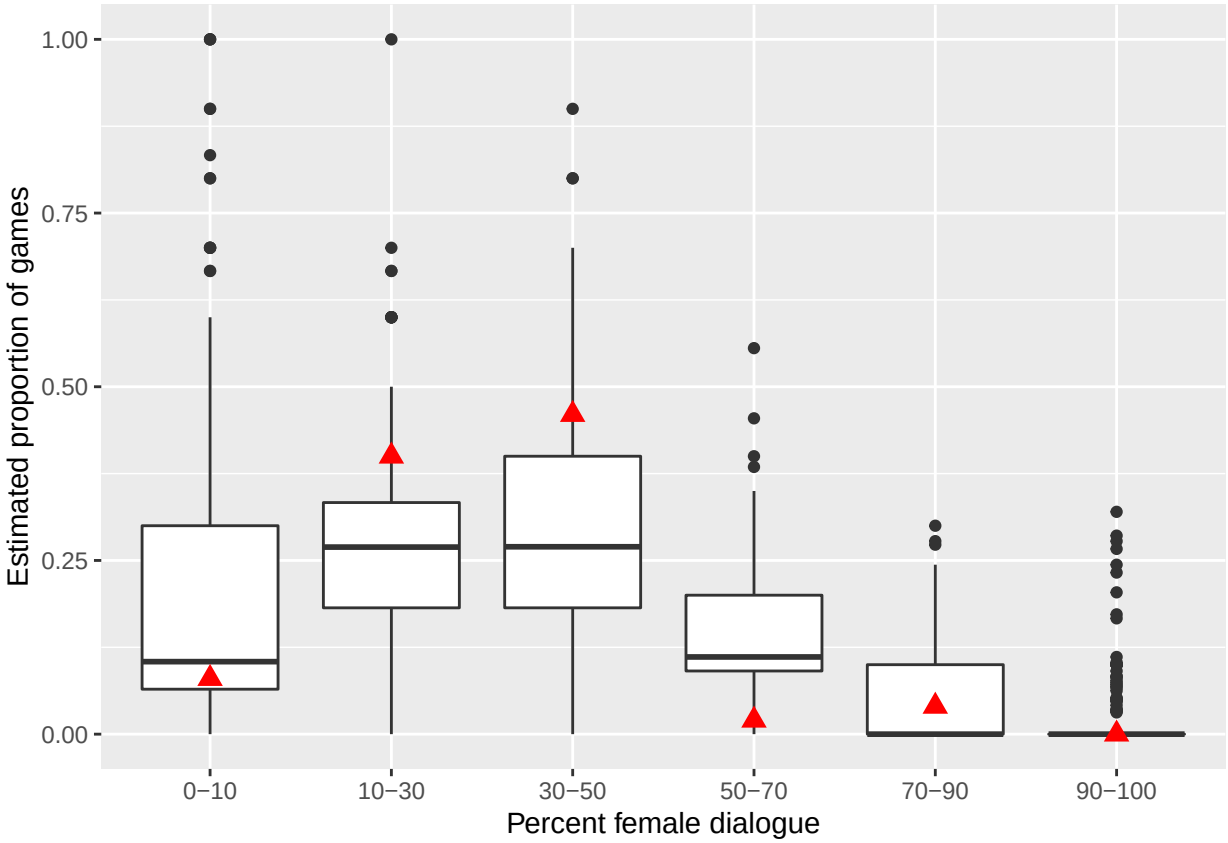
Collect the data:

```
rx = data.frame(
  n = unlist(f),
  p = unlist(fNormed),
  category.boundary = rep(c(10,30,50,70,90,100),each=nrow(f)),
  category.middle = rep(c(5,20,40,60,80,95),each=nrow(f)),
  category.names = rep(c("0-10","10-30","30-50","50-70","70-90","90-100"),each=nrow(f)),
  rowSum = rep(rs,6)
)

brks = c(10,30,50,70,90,101)
tn = table(cut(trueFemaleDialogueDist*100,c(0,brks)))
tp = tn/sum(tn)
trueDist = data.frame(
  p = as.vector(tp),
  category.middle = c(5,20,40,60,80,95),
  category.names = c("0-10","10-30","30-50","50-70","70-90","90-100"))
```

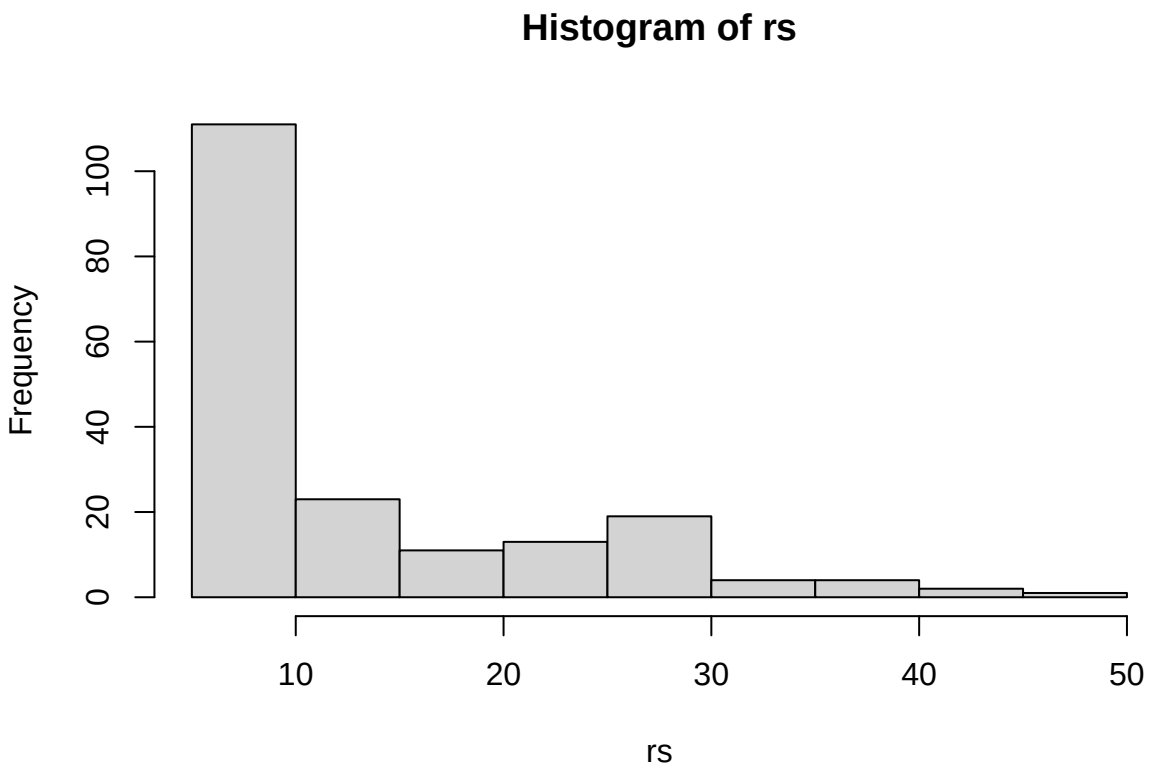
Plot the mean guesses as boxplots, with the true distribution indicated as red triangles

```
ggplot(rx, aes(x=factor(category.names),y=p))+
  geom_boxplot() +
  geom_point(data=trueDist,aes(y=p),stat="identity",
             colour="red",size=3,shape="triangle") +
  xlab("Percent female dialogue") +
  ylab("Estimated proportion of games")
```



Did people understand the question? What totals did they actual use?

```
hist(rs)
```

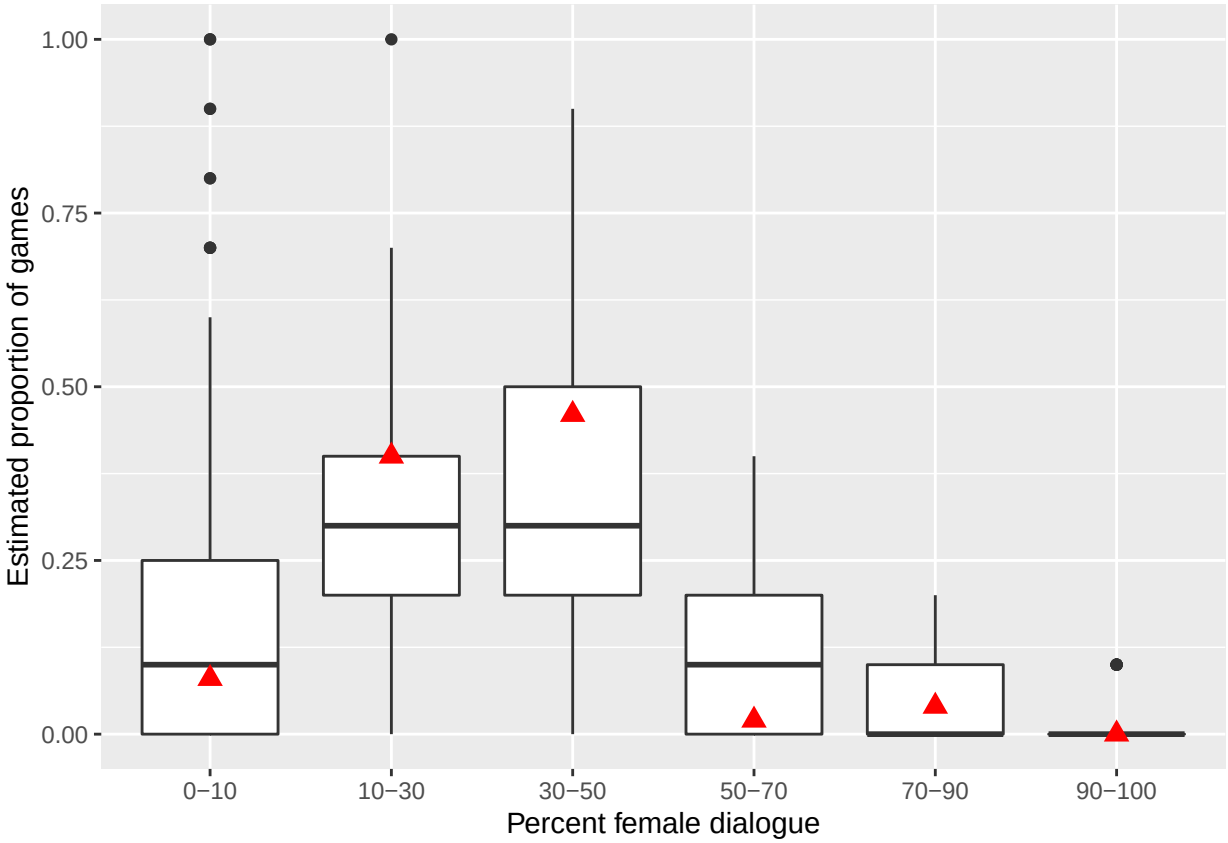



```
# Proportion of participants responding with 10 games
sum(rs==10)/length(rs)
```

```
## [1] 0.5478723
```

Only about half of respondents used 10 items. So let's see the distribution for only participants that specified a total of 10:

```
ggplot(rx[rx$rowSum==10,], aes(x=factor(category.names),y=p))+
  geom_boxplot() +
  geom_point(data=trueDist,aes(y=p),stat="identity",
             colour="red",size=3,shape="triangle")+
  xlab("Percent female dialogue") +
  ylab("Estimated proportion of games")
```



People seem to over-estimate the number of games that have more than 50% dialogue.

For each percentage category except 70-90, the mean guesses are significantly different from the true values:

```
distTest = data.frame()
for(i in 1:nrow(trueDist)){
  mx = rx$category.middle==trueDist[i,]$category.middle
  tt = t.test(rx[mx,]$p,
         mu=trueDist[i,]$p)
  distTest = rbind(distTest,data.frame(
    category = rx[mx,]$category.names[1],
    participantMean = tt$estimate, trueMean = trueDist[i,]$p,
    t = tt$statistic,p = tt$p.value
  ))
}
# Adjust for multiple comparisons
distTest$p = distTest$p * 6
distTest
```

```
##            category participantMean trueMean          t            p
## mean of x      0-10      0.21343244     0.08   7.918398 1.245801e-12
## mean of x1    10-30      0.27086989     0.40 -11.447707 1.468153e-22
## mean of x2    30-50      0.30081765     0.46 -11.725708 2.221112e-23
## mean of x3    50-70      0.13679580     0.02  15.784302 1.884724e-35
## mean of x4    70-90      0.05439421     0.04   2.873463 2.717884e-02
## mean of x5   90-100      0.02369000     0.00   5.626078 3.987537e-07
```

Write out stats for the 50-70 category for the paper

```
tx = distTest[distTest$category=="50-70",]
cat(round(100*tx$participantMean),file="../../results/latexStats/Survey_dist50_70_partGuess.tex")
cat(round(100*tx$trueMean),file="../../results/latexStats/Survey_dist50_70_true.tex")
px = "p < 0.0001"
if(tx$p>0.0001){
  px = paste0("p = ",round(tx$p,4))
}
cat(paste0("t = ",round(tx$t,2),", ",px)
  ,file="../../results/latexStats/Survey_dist50_70_stat.tex")
```

Conclusion

The mean guess for the average percentage of female dialogue was 30.88% (sd = 12.02). This is not significantly different from the empirical estimate. Male respondents estimated a higher percentage of female dialogue than female respondents.

The mean guess for the percentage of games with more than 50% female dialogue was 24.08% (sd = 23.68). This is significantly higher than the empirical estimate. The main predictor of this measure was the whether the survey was framed to ask about female or male dialogue.