

# Experiment 1: Main Analyses

Bethany Gardner

05/05/2022

## Contents

|   |          |
|---|----------|
| <b>Setup</b>                            | <b>1</b> |
| <b>Data Summary</b>                     | <b>2</b> |
| <b>Model 1: Condition</b>               | <b>3</b> |
| Convert to Odds Ratios . . . . .        | 3        |
| <b>Model 2: Condition * Name Gender</b> | <b>5</b> |

## Setup

Load data and select columns used in model. See data/exp1\_data\_about.txt for more details.

```
d <- read.csv("../data/exp1_data.csv", stringsAsFactors=TRUE) %>%  
  rename("Participant"="SubjID", "Item"="NameShown") %>%  
  select(Participant, Condition, GenderRating, Item, He, She, Other)  
str(d)
```

```
## 'data.frame':   9564 obs. of  7 variables:  
## $ Participant : Factor w/ 457 levels "R_01wgzz7ygaVl8aJ",...: 278 278 278 278 278 278 278 278 278 278 ...  
## $ Condition   : Factor w/ 3 levels "first","full",...: 1 1 1 1 1 1 1 1 1 1 ...  
## $ GenderRating: num  1.21 1.24 1.28 2.12 2.41 2.61 3.61 3.75 3.87 4.22 ...  
## $ Item        : Factor w/ 104 levels "Ashley","Ashley Cook",...: 64 11 43 18 95 29 88 71 79 92 ...  
## $ He          : int   1 1 1 1 0 1 1 0 1 1 ...  
## $ She         : int   0 0 0 0 1 0 0 1 0 0 ...  
## $ Other       : int   0 0 0 0 0 0 0 0 0 0 ...
```

Center gender rating for names: Original scale from 1 to 7, with 1 as most masculine and 7 as most feminine. Mean-centered with higher still as more feminine.

```
d %>% mutate(GenderRatingCentered=scale(d$GenderRating, scale=FALSE))
```

Set contrasts for name conditions.

```
contrasts(d$Condition) = cbind("last vs first/full"=c(.33,.33,-0.66),
                              "first vs full"=c(-.5,.5,0))
contrasts(d$Condition)
```

```
##      last vs first/full first vs full
## first           0.33         -0.5
## full            0.33           0.5
## last           -0.66           0.0
```

Subset for gender rating effects (First and Full conditions only).

```
d.FF <- d %>% filter(Condition!="last")
d.FF$Condition <- droplevels(d.FF$Condition)
contrasts(d.FF$Condition) = cbind("first vs full"=c(-.5,.5)) #add contrast back
contrasts(d.FF$Condition)
```

```
##      first vs full
## first          -0.5
## full           0.5
```

## Data Summary

Responses by condition.

```
d %>% mutate(ResponseAll=case_when(
  He==1 ~ "He",
  She==1 ~ "She",
  Other==1 ~ "Other"))

d.count_responses <- d %>% group_by(Condition, ResponseAll) %>%
  summarise(n=n()) %>%
  pivot_wider(names_from=ResponseAll,
              values_from=n) %>%
  mutate(She_HeOther = She / (He+Other),
         She_He = She / He)

kable(d.count_responses, digits=3, align='c')
```

| Condition | He   | Other | She  | She_HeOther | She_He |
|-----------|------|-------|------|-------------|--------|
| first     | 1572 | 225   | 1395 | 0.776       | 0.887  |
| full      | 1514 | 131   | 1535 | 0.933       | 1.014  |
| last      | 2616 | 325   | 251  | 0.085       | 0.096  |

- First name condition has second-most *she* responses
- Full name condition has most *she* responses
- Last name condition has fewest *she* responses

## Model 1: Condition

Effect of Condition (first name, last name, full name) on likelihood of a *she* response, as opposed to a *he* or *other* response. Participant and Item are included as random intercepts, with items defined as the unique first, last and first + last name combinations. Because the condition manipulations were fully between-subject and between-item, fitting a random slope model was not possible.

```
m.cond <- glmer(She ~ Condition + (1|Participant) + (1|Item),
               data=d, family=binomial)
m.cond_tidy <- tidy(m.cond)
summary(m.cond)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##   Family: binomial   ( logit )
## Formula: She ~ Condition + (1 | Participant) + (1 | Item)
##   Data: d
##
##           AIC          BIC    logLik deviance df.resid
##    6406.5     6442.3   -3198.2   6396.5     9559
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -8.9619 -0.3029 -0.1438  0.2164 10.0122
##
## Random effects:
##   Groups       Name            Variance Std.Dev.
##   Participant (Intercept) 1.029      1.014
##   Item          (Intercept) 7.234      2.690
## Number of obs: 9564, groups: Participant, 457; Item, 104
##
## Fixed effects:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -1.4284    0.3076  -4.644 3.42e-06 ***
## Conditionlast vs first/full    2.8241    0.7016   4.026 5.69e-05 ***
## Conditionfirst vs full         0.6197    0.6998   0.886  0.376
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Cvfrs/
## Cndtnvfirst/ -0.181
## Cndtnfrstvf -0.360 -0.239
```

- Fewer *she* responses overall
- First+Full have more *she* responses than Last. Full has more *she* responses than First (n.s. but matches ratios).

## Convert to Odds Ratios

### Intercept

```
m.cond_intercept <- m.cond_tidy %>% filter(term=="(Intercept)") %>%
  select(estimate) %>% as.numeric()

exp(m.cond_intercept)
```

```
## [1] 0.2396996
```

```
exp(-m.cond_intercept)
```

```
## [1] 4.171888
```

0.24x less likely to use *she* overall. Easier to interpret: 4.17x more likely to use *he* or *other* overall.

### Condition: Last vs First+Full

```
m.cond_LFF <- m.cond_tidy %>%
  filter(term=="Conditionlast vs first/full") %>%
  select(estimate) %>% as.numeric()
exp(m.cond_LFF)
```

```
## [1] 16.846
```

16.85x more likely to use *she* in First + Full compared to Last. -> 16.85 times more likely to use *he* and *other* in Last than in First + Full.

### Condition: Last Only

Dummy code with Last Name as 0, so that intercept is the Last Name condition only.

```
d %<>% mutate(Condition_Last=case_when(
  Condition=="first" ~ 1,
  Condition=="full" ~ 1,
  Condition=="last" ~ 0))
d$Condition_Last %<>% as.factor()

m.last <- glmer(She ~ Condition_Last + (1|Participant) + (1|Item),
  data=d, family=binomial)
m.last_tidy <- tidy(m.last)
```

```
m.cond_last <- m.last_tidy %>%
  filter(term=="(Intercept)") %>%
  select(estimate) %>% as.numeric()

exp(m.cond_last)
```

```
## [1] 0.0371584
```

```
exp(-m.cond_last)
```

```
## [1] 26.91181
```

0.04x times less likely to use *she* in the Last Name condition → 26.91x more likely to use *he* and *other* in the Last Name condition.

### Condition: First and Full Only

Dummy code with First and Full Name as 0, so that intercept is average for these two conditions.

```
d %<>% mutate(Condition_FF=case_when(
  Condition=="first" ~ 0,
  Condition=="full" ~ 0,
  Condition=="last" ~ 1))
d$Condition_FF %<>% as.factor()

m.ff <- glmer(She ~ Condition_FF + (1|Participant) + (1|Item),
  data=d, family=binomial)
m.ff_tidy <- tidy(m.ff)
```

```
m.cond_ff <- m.ff_tidy %>%
  filter(term=="(Intercept)") %>%
  select(estimate) %>% as.numeric()

exp(m.cond_ff)
```

```
## [1] 0.7045676
```

```
exp(-m.cond_ff)
```

```
## [1] 1.41931
```

0.70x times less likely to use *she* in the First and Full Name conditions → 1.42x more likely to use *he* and *other* in the First and Full Name conditions.

## Model 2: Condition \* Name Gender

Effects of Condition (first name, full name) and the first name's Gender Rating (centered, positive=more feminine) on the likelihood of a *she* response, as opposed to a *he* or *other* response. In Experiment 1, the Last Name condition does not include any instances of the gendered first name, so only the First and Full Name conditions are analyzed here. Participant and Item are again included as random intercepts.

```
m.namegender <- glmer(She ~ Condition * GenderRatingCentered +
  (1|Participant) + (1|Item),
  data=d.FF, family=binomial)
m.namegender_tidy <- tidy(m.namegender)
summary(m.namegender)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: She ~ Condition * GenderRatingCentered + (1 | Participant) +
## (1 | Item)
```

```

## Data: d.FF
##
##      AIC      BIC   logLik deviance df.resid
##  4657.4   4698.0  -2322.7   4645.4     6366
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -9.1567 -0.3548 -0.0551  0.3126 14.3200
##
## Random effects:
## Groups      Name      Variance Std.Dev.
## Participant (Intercept) 0.889    0.9429
## Item        (Intercept) 0.501    0.7078
## Number of obs: 6372, groups: Participant, 305; Item, 83
##
## Fixed effects:
##
##              Estimate Std. Error z value
## (Intercept)    -0.51325    0.11987  -4.282
## Conditionfirst vs full    0.53204    0.23993   2.218
## GenderRatingCentered     1.59330    0.07253  21.967
## Conditionfirst vs full:GenderRatingCentered -0.17492    0.13917  -1.257
##
##              Pr(>|z|)
## (Intercept)    1.86e-05 ***
## Conditionfirst vs full    0.0266 *
## GenderRatingCentered    < 2e-16 ***
## Conditionfirst vs full:GenderRatingCentered  0.2088
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) Cndtvf GndrRC
## Cndtnfrstvf -0.346
## GndrRtngCnt -0.179  0.122
## Cvfll:GndrRC  0.111 -0.172 -0.409

```

- More *she* responses as first names become more feminine.
- Difference between First and Full is now significant (as compared to condition-only model).