Experiment 1: Main Analyses

Bethany Gardner

2022-07-07

Contents

Setup

Variable names:

- Experiment: exp1
- Type
  - d = data
  - m = model
  - est = log odds estimate from model
  - OR = odds ratio converted from est
- Analysis
  - count =sums of response types
  - cond = effect of Condition (Last vs First+Full)
  - nameGender = effects of Condition (First vs Full) and Name Gender Rating
- Subset
  - all = including *other* responses
  - noOther = excluding *other* responses

- FF = First and Full Name conditions only
- Last = Last Name condition only

Load data and select columns used in model. See data/exp1_data_about.txt for more details.

```
exp1_d <- read.csv("../data/exp1_data.csv",
                   stringsAsFactors=TRUE) %>%
  rename("Participant"="SubjID", "Item"="NameShown") %>%
  select(Participant, SubjGender, Condition, GenderRating,
         Item, He, She, Other)
str(exp1_d)
```

```
## 'data.frame':    9564 obs. of  8 variables:
##  $ Participant : Factor w/ 457 levels "R_01wgzz7ygaVl8aJ",..: 278 278 278 278 278 278 278 278 278 278
##  $ SubjGender  : Factor w/ 5 levels "female","genderfluid",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ 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.

```
exp1_d %<>% mutate(GenderRatingCentered=
    scale(GenderRating, scale=FALSE))
```

Set contrasts for name conditions.

```
contrasts(exp1_d$Condition) = cbind(
  "last vs first/full"=c(.33,.33,-0.66),
  "first vs full"=c(-.5,.5,0))
contrasts(exp1_d$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).

```
exp1_d_FF <- exp1_d %>% filter(Condition!="last")
exp1_d_FF$Condition <- droplevels(exp1_d_FF$Condition)
contrasts(exp1_d_FF$Condition) = cbind(
  "first vs full"=c(-.5,.5)) #add contrast back
contrasts(exp1_d_FF$Condition)
```

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

Data Summary

Responses by condition.

```
exp1_d %<>% mutate(ResponseAll=case_when(
       He==1 ~ "He",
       She==1 ~ "She",
       Other==1 ~ "Other"))

exp1_d_count <- exp1_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(exp1_d_count, 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.

```
exp1_m_cond <- glmer(
  She ~ Condition + (1|Participant) + (1|Item),
  data=exp1_d, family=binomial)
summary(exp1_m_cond)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: She ~ Condition + (1 | Participant) + (1 | Item)
##    Data: exp1_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/
## Cndtnvfrst/ -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).

Odds Ratios: Intercept

```
exp1_est_cond_intercept <- exp1_m_cond %>%
  tidy() %>%
  filter(term=="(Intercept)") %>%
  select(estimate) %>% as.numeric()

exp(exp1_est_cond_intercept)
```

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

```
exp(-exp1_est_cond_intercept)
```

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

```
#Save this for the table comparing all 4 experiments
exp1_OR_all_I <- exp(-exp1_est_cond_intercept) %>%
  round(2)
```

0.24x less likely to use to use *she* overall. Easier to interpret: 4.17x more likely to use *he* or *other* overall, $p<.001$

Odds Ratios: Last vs First+Full

```
exp1_est_cond_LFF <- exp1_m_cond %>%
  tidy() %>%
  filter(term=="Conditionlast vs first/full") %>%
  select(estimate) %>% as.numeric()
exp(exp1_est_cond_LFF)
```

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

```
#Save this for the table comparing all 4 experiments
exp1_OR_all_LFF <- exp(-exp1_est_cond_LFF) %>%
  round(2)
```

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, p<.001

Odds Ratios: Last Only

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

```
exp1_d %<>% mutate(Condition_Last=case_when(
  Condition=="first" ~ 1,
  Condition=="full" ~ 1,
  Condition=="last" ~ 0))
exp1_d$Condition_Last %<>% as.factor()
```

```
exp1_m_L <- glmer(
  She ~ Condition_Last + (1|Participant) + (1|Item),
  data=exp1_d, family=binomial)
summary(exp1_m_L)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: She ~ Condition_Last + (1 | Participant) + (1 | Item)
##    Data: exp1_d
##
##      AIC      BIC   logLik deviance df.resid
##   6405.2   6433.9  -3198.6   6397.2     9560
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -9.0006 -0.3022 -0.1440  0.2163  9.8409
##
## Random effects:
##  Groups      Name        Variance Std.Dev.
##  Participant (Intercept) 1.030    1.015
##  Item        (Intercept) 7.274    2.697
## Number of obs: 9564, groups:  Participant, 457; Item, 104
```

```
##
## Fixed effects:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -3.2926     0.6020  -5.469 4.52e-08 ***
## Condition_Last1   2.9424     0.6766   4.349 1.37e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## Condtn_Lst1 -0.890
```

```
exp1_est_L <- exp1_m_L %>%
  tidy() %>%
  filter(term=="(Intercept)") %>%
  select(estimate) %>% as.numeric()

exp(exp1_est_L)
```

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

```
exp(-exp1_est_L)
```

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

```
#Save this for the table comparing all 4 experiments
exp1_OR_all_L <- exp(-exp1_est_L) %>%
  round(2)
```

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, p<.001

Odds Ratios: First and Full Only

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

```
exp1_d %<>% mutate(Condition_FF=case_when(
  Condition=="first" ~ 0,
  Condition=="full" ~ 0,
  Condition=="last" ~ 1))
exp1_d$Condition_FF %<>% as.factor()
```

```
exp1_m_FF <- glmer(
  She ~ Condition_FF + (1|Participant) + (1|Item),
  data=exp1_d, family=binomial)
summary(exp1_m_FF)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
```

```
## Formula: She ~ Condition_FF + (1 | Participant) + (1 | Item)
##    Data: exp1_d
##
##      AIC      BIC   logLik deviance df.resid
##   6405.2   6433.9  -3198.6   6397.2     9560
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -9.0006 -0.3022 -0.1440  0.2163  9.8408
##
## Random effects:
##  Groups      Name        Variance Std.Dev.
##  Participant (Intercept) 1.030    1.015
##  Item        (Intercept) 7.274    2.697
## Number of obs: 9564, groups:  Participant, 457; Item, 104
##
## Fixed effects:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.3502     0.3088  -1.134    0.257
## Condition_FF1  -2.9424     0.6766  -4.349 1.37e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## Conditn_FF1 -0.456
```

```
exp1_est_FF <- exp1_m_FF %>%
  tidy() %>%
  filter(term=="(Intercept)") %>%
  select(estimate) %>% as.numeric()

exp(exp1_est_FF)
```

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

```
exp(-exp1_est_FF)
```

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

```
#Save this for the table comparing all 4 experiments
exp1_OR_all_FF <- exp(-exp1_est_FF) %>%
  round(2)
```

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, p=.26

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.

```
exp1_m_nameGender <- glmer(
  She ~ Condition * GenderRatingCentered +
      (1|Participant) + (1|Item),
  exp1_d_FF, family=binomial)
summary(exp1_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: exp1_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:GndRC  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).