Study 2A - New Attributes

October 06, 2025

Items

Read Data

```
# Set this to TRUE if you have API access, FALSE if using CSV
USE_API <- T

if(USE_API) {
  ## Pull directly from Qualtrics API
  qual_data <- fetch_survey(surveyID='SV_8vM8mfohSNgdTxQ',  # New Study ID
                            start_date = "2025-10-02",
                    force_request = T)
} else {
  # Read the processed data directly from CSV
  d0 <- read.csv('Study2A.csv', check.names = F)
  num_excluded <- unique(d0$num_excluded_total)
}

# Define the categories based on the JavaScript feedback code
women <- c('Zadie Smith', 'Isabel Allende', 'Jane Austen',
           'Joyce Carol Oates', 'Lucy Maud', 'Sylvia Plath')

multi_genre <- c('Herman Melville', 'Joyce Carol Oates', 'Cormac McCarthy', 'Ray
↪  Bradbury',
                 'Kurt Vonnegut', 'George Orwell', 'Isabel Allende', 'Gabriel Garcia
                 ↪  Marquez',
                 'Lucy Maud', 'Neil Gaiman', 'J.R.R. Tolkien')

sold_30m <- c('Charles Dickens', 'Gabriel Garcia Marquez', 'George Orwell', 'Isabel
↪  Allende',
              'J.D. Salinger', 'J.R.R. Tolkien', 'Lucy Maud', 'F. Scott Fitzgerald')

classic_50plus <- c('Jane Austen', 'Charles Dickens', 'Herman Melville', 'Nathaniel
↪  Hawthorne',
                    'Jack London', 'J.D. Salinger', 'F. Scott Fitzgerald', 'Ernest
                    ↪  Hemingway',
                    'John Steinbeck', 'George Orwell', 'J.R.R. Tolkien', 'Lucy Maud',
                    'Ray Bradbury', 'Kurt Vonnegut', 'Gabriel Garcia Marquez',
                    'Michael Crichton', 'Sylvia Plath')

if(USE_API) {
  # Process the API data with new category definitions
  d0 <- qual_data |>
    filter(!is.na(`choice-7`), !is.na(PROLIFIC_PID)) |>
    mutate(
      # Gender feedback detection
      gender_feedback = as.numeric(grepl("were women", feedbackItem1) |
                    grepl("were women", feedbackItem2) |
                    grepl("were women", feedbackItem3)),

      # Attribute feedback detection based on JavaScript code
      forms_shown = as.numeric(grepl("spanning multiple genres", feedbackItem1) |
                  grepl("spanning multiple genres", feedbackItem2) |
                  grepl("spanning multiple genres", feedbackItem3)),

      sold30m_shown = as.numeric(grepl("30M\\+ copies sold", feedbackItem1) |
```

```
                  grepl("30M\\+ copies sold", feedbackItem2) |
                  grepl("30M\\+ copies sold", feedbackItem3)),

    classic_shown = as.numeric(grepl("remained in continuous print for over 50 years",
    ↪  feedbackItem1) |
                  grepl("remained in continuous print for over 50 years",
                  ↪  feedbackItem2) |
                  grepl("remained in continuous print for over 50 years",
                  ↪  feedbackItem3)),

    # Picks based on categories
    female_pick = case_when(`choice-7` %in% women ~ 1, TRUE ~ 0),
    forms_pick = case_when(`choice-7` %in% multi_genre ~ 1, TRUE ~ 0),
    sold30m_pick = case_when(`choice-7` %in% sold_30m ~ 1, TRUE ~ 0),
    classic_pick = case_when(`choice-7` %in% classic_50plus ~ 1, TRUE ~ 0),

    # Demographics
    gender_code = case_when(gender=="Man" ~ 1, TRUE ~ 0),
    race_code = case_when(str_detect(race, "White / Caucasian") ~ 1, TRUE ~ 0),

    # Process political ideology (1-7 scale)
    poli_numeric = case_when(
      poli == "1Extremely liberal" ~ 1,
      poli == "2Liberal" ~ 2,
      poli == "3Slightly liberal" ~ 3,
      poli == "4Moderate; middle of the road" ~ 4,
      poli == "5Slightly conservative" ~ 5,
      poli == "6Conservative" ~ 6,
      poli == "7Extremely conservative" ~ 7,
      TRUE ~ NA_integer_
    ),

    # Center political ideology for interaction
    poli_centered = poli_numeric - 4,  # Center at moderate (4)

    # Process political party affiliation
    party_democrat = case_when(poli_party == "Democrat" ~ 1, TRUE ~ 0),
    party_independent = case_when(poli_party == "Independent" ~ 1, TRUE ~ 0),
    party_republican = case_when(poli_party == "Republican" ~ 1, TRUE ~ 0),

    base_gender = rowSums(across(`choice-1`:`choice-6`, ~ . %in% women))
  ) |>
  mutate(
    across(c(I1:E3),
           ~ case_when(
             . == "Strongly disagree" ~ 1, . == "Disagree" ~ 2, . == "Somewhat
↪  disagree" ~ 3,
             . == "Neither agree nor disagree" ~ 4, . == "Somewhat agree" ~ 5, . ==
↪  "Agree" ~ 6,
             . == "Strongly agree" ~ 7, TRUE ~ NA_integer_))) |>
  mutate(
    internal1Z = (I1 - mean(I1, na.rm = TRUE)) / sd(I1, na.rm = TRUE),
    internal2Z = (I2 - mean(I2, na.rm = TRUE)) / sd(I2, na.rm = TRUE),
    internal3Z = (I3 - mean(I3, na.rm = TRUE)) / sd(I3, na.rm = TRUE),
```

```
      internal4Z = (I4 - mean(I4, na.rm = TRUE)) / sd(I4, na.rm = TRUE),
      internal = (internal1Z + internal2Z + internal3Z + internal4Z) / 4,
      external1Z = (E1 - mean(E1, na.rm = TRUE)) / sd(E1, na.rm = TRUE),
      external2Z = (E2 - mean(E2, na.rm = TRUE)) / sd(E2, na.rm = TRUE),
      external3Z = (E3 - mean(E3, na.rm = TRUE)) / sd(E3, na.rm = TRUE),
      external = (external1Z + external2Z + external3Z) / 3
    ) |>
    dplyr::select(gender_feedback, forms_shown, sold30m_shown, classic_shown,
                  female_pick, forms_pick, sold30m_pick, classic_pick, base_gender,
↪   `choice-1`:`choice-7`,
                  race, gender, age, gender_code, race_code, poli, poli_numeric,
↪   poli_centered,
                  poli_party, party_democrat, party_independent, party_republican,
                  internal1Z:external) |>
    slice(1:1000) # pre-registered sample size

  # Calculate the number of excluded participants
  num_excluded <- nrow(qual_data) - nrow(d0)

  # Save num_excluded in d0
  d0$num_excluded_total <- num_excluded  # As a column

  # Write the API-pulled data into a CSV file
  write.csv(d0, 'Study2A.csv', row.names = FALSE, quote = TRUE)
}
```

Variable Names

Variable	Description
gender_feedback	Binary indicator of whether a participant was randomly assigned to gender feedback condition.
female_pick	Binary indicator of whether a participant selected a female author for their seventh selection
forms_pick	Binary indicator of whether a participant selected an author who wrote in multiple genres for their seventh selection.
sold30m_pick	Binary indicator of whether a participant selected an author with at least one book with 30M+ copies sold for their seventh selection.
classic_pick	Binary indicator of whether a participant selected an author with a book in print for 50+ years for their seventh selection.
base_gender	Count of the number of female authors selected in the initial six authors.
choice-1 to choice-7	The selected authors
gender	Self-selected gender.
race	Self-selected race.
age	Self-entered age.
poli	Political ideology on 7-point scale (1 = extremely liberal, 7 = extremely conservative).
poli_centered	Centered political ideology (0 = moderate).
poli_party	Political party affiliation (Democrat, Republican, Independent).
party_democrat	Indicator for Democrat party affiliation.
party_independent	Indicator for Independent party affiliation.
party_republican	Indicator for Republican party affiliation (reference category).
gender_code	Dummy code for gender (male = 1).
race_code	Dummy code for race (white = 1).
internal1Z-4Z	Individual standardized scale items for Internal Motivation to Respond Without Prejudice.
external1Z-3Z	Individual standardized scale items for External Motivation to Respond Without Prejudice.
internal	Aggregated scale items for Internal Motivation to Respond Without Prejudice.
external	Aggregated scale items for External Motivation to Respond Without Prejudice.

Demographics

```
## Excluded Participants: 126
```

```
##                         Percentage gender
## 1                            Woman   50.7
## 2                              Man   48.0
## 3                       Non-binary    1.1
## 4 Another gender not listed here:    0.2
```

```
##                           Percentage Race
## 1 American Indian or Alaskan Native  0.5
## 2          Asian / Pacific Islander  8.0
## 3         Black or African American 10.8
## 4                 Hispanic / Latinx  4.8
## 5                 White / Caucasian 75.9
```

```
## # A tibble: 1 x 2
##   mean_age sd_age
##      <dbl>  <dbl>
## 1     43.0   13.3
```

```
## Mean Political Ideology (1-7 scale):  3.54
```

```
## SD Political Ideology:  1.85
```

```
##    Percentage Party
## 1    Democrat  40.5
## 2  Republican  28.0
## 3 Independent  31.5
```

```
## Mean (num of initial women selected):  1.52
```

```
## SD (num of initial women selected):  1.22
```

```
## Percentage (initial women selected):  0.2533333
```

```
## SD (initial women selected):  0.2033333
```

```
## # A tibble: 2 x 2
##   gender_feedback  mean
##             <dbl> <dbl>
## 1               0 0.256
## 2               1 0.251
```

```
##
##  Welch Two Sample t-test
##
## data:  base_gender/6 by gender_feedback
## t = 0.41373, df = 997.51, p-value = 0.6792
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
```

```
## 95 percent confidence interval:
##  -0.01991557  0.03055695
## sample estimates:
## mean in group 0 mean in group 1
##       0.2558217       0.2505010
```

Cronbach's Alpha

```r
# Calculating Cronbach's Alpha for the Internal subscale
internal_items <- d0[, c("internal1Z", "internal2Z", "internal3Z", "internal4Z")]
alpha_internal <- alpha(internal_items)

cat("Cronbach's Alpha for Internal Subscale: ", alpha_internal$total$raw_alpha, "\n")
```

```
## Cronbach's Alpha for Internal Subscale:  0.925711
```

```r
# Calculating Cronbach's Alpha for the External subscale
external_items <- d0[, c("external1Z", "external2Z", "external3Z")]
alpha_external <- alpha(external_items)
cat("Cronbach's Alpha for External Subscale: ", alpha_external$total$raw_alpha, "\n")
```

```
## Cronbach's Alpha for External Subscale:  0.8907377
```

Primary Analysis

```r
# primary model
r1 <- lm(female_pick ~ gender_feedback, data=d0)

# Display the summary with robust standard errors
robust_summary(r1)
```

```
##
## Call:
## lm(formula = female_pick ~ gender_feedback, data = d0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4148 -0.4148 -0.2954  0.5852  0.7046
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.29541    0.02042  14.464  < 2e-16 ***
## gender_feedback  0.11942    0.03009   3.968 7.75e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4752 on 998 degrees of freedom
## Multiple R-squared:  0.01557,    Adjusted R-squared:  0.01458
## F-statistic: 15.79 on 1 and 998 DF,  p-value: 7.605e-05
```

```r
robust_confint(r1)
```

```
##                     2.5 %    97.5 %
## (Intercept)     0.2553313 0.3354870
## gender_feedback 0.0603693 0.1784717
```

Robustness

```
## which feedback was shown with gender, remove constant due to collinearity
r2 <- lm(female_pick ~ gender_feedback + classic_shown + sold30m_shown + forms_shown - 1,
↪  data=d0)

# Display the summary with robust standard errors
robust_summary(r2)
```

```
##
## Call:
## lm(formula = female_pick ~ gender_feedback + classic_shown +
##     sold30m_shown + forms_shown - 1, data = d0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4775 -0.3704 -0.2954  0.6101  0.7046
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## gender_feedback  0.21567    0.02595   8.312 3.05e-16 ***
## classic_shown    0.12114    0.03224   3.758 0.000181 ***
## sold30m_shown    0.03355    0.03173   1.058 0.290519
## forms_shown      0.14071    0.03194   4.406 1.17e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4745 on 996 degrees of freedom
## Multiple R-squared:  0.3682, Adjusted R-squared:  0.3657
## F-statistic: 145.1 on 4 and 996 DF,  p-value: < 2.2e-16
```

```
robust_confint(r2)
```

```
##                      2.5 %     97.5 %
## gender_feedback  0.16475808 0.26658672
## classic_shown    0.05788556 0.18440339
## sold30m_shown   -0.02870736 0.09581435
## forms_shown      0.07804062 0.20338181
```

```
## robust to demographic controls

r3 <- lm(female_pick ~ gender_feedback + gender_code + race_code + age, data=d0)

# Display the summary with robust standard errors
robust_summary(r3)
```

```
##
## Call:
## lm(formula = female_pick ~ gender_feedback + gender_code + race_code +
##     age, data = d0)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4752 -0.3786 -0.2831  0.5812  0.7698
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.377758   0.059094   6.392 2.51e-10 ***
## gender_feedback  0.120648   0.030186   3.997 6.89e-05 ***
## gender_code     -0.053819   0.030197  -1.782    0.075 .
## race_code       -0.018170   0.036863  -0.493    0.622
## age             -0.001008   0.001163  -0.867    0.386
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.475 on 995 degrees of freedom
## Multiple R-squared:  0.01956,    Adjusted R-squared:  0.01562
## F-statistic: 4.964 on 4 and 995 DF,  p-value: 0.0005779
```

```
robust_confint(r3)
```

```
##                       2.5 %      97.5 %
## (Intercept)      0.261794280 0.493722488
## gender_feedback  0.061412673 0.179883741
## gender_code     -0.113076555 0.005437751
## race_code       -0.090507926 0.054168284
## age             -0.003288957 0.001273909
```

```
## logistic regression
# Fit the logistic regression model
r4 <- glm(female_pick ~ gender_feedback, family = binomial, data=d0)
summary(r4)
```

```
##
## Call:
## glm(formula = female_pick ~ gender_feedback, family = binomial,
##     data = d0)
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.86926    0.09793  -8.877  < 2e-16 ***
## gender_feedback  0.52522    0.13359   3.932 8.43e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1301.0  on 999  degrees of freedom
## Residual deviance: 1285.3  on 998  degrees of freedom
## AIC: 1289.3
##
## Number of Fisher Scoring iterations: 4
```

```r
# Odds ratio
tidy_r4 <- tidy(r4, exponentiate = TRUE, conf.int = T)
print(tidy_r4)
```

```
## # A tibble: 2 x 7
##   term            estimate std.error statistic  p.value conf.low conf.high
##   <chr>              <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 (Intercept)        0.419    0.0979     -8.88 6.89e-19    0.345     0.507
## 2 gender_feedback    1.69     0.134       3.93 8.43e- 5    1.30      2.20
```

Interaction Analysis

```r
## interaction of base gender
# primary model
r_interaction <- lm(female_pick ~ gender_feedback*base_gender, data=d0)


# Display the summary with robust standard errors
robust_summary(r_interaction)
```

```
##
## Call:
## lm(formula = female_pick ~ gender_feedback * base_gender, data = d0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4987 -0.3313 -0.2908  0.5571  0.7803
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  0.310471   0.032335   9.602  < 2e-16 ***
## gender_feedback              0.188241   0.046932   4.011  6.5e-05 ***
## base_gender                 -0.009813   0.015939  -0.616    0.538
## gender_feedback:base_gender -0.045997   0.022363  -2.057    0.040 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4732 on 996 degrees of freedom
## Multiple R-squared:  0.02614,    Adjusted R-squared:  0.02321
## F-statistic: 8.913 on 3 and 996 DF,  p-value: 7.865e-06
```

Moderation by Political Ideology

```
# Test for moderation by political ideology
r_poli_moderation <- lm(female_pick ~ gender_feedback * poli_centered, data=d0)

# Display the summary with robust standard errors
cat("Moderation by Political Ideology\n")
```

```
## Moderation by Political Ideology
```

```
robust_summary(r_poli_moderation)
```

```
##
## Call:
## lm(formula = female_pick ~ gender_feedback * poli_centered, data = d0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4979 -0.3443 -0.2855  0.5665  0.7733
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     0.28551    0.02064  13.830  < 2e-16 ***
## gender_feedback                 0.11583    0.03039   3.811 0.000147 ***
## poli_centered                  -0.01960    0.01065  -1.841 0.065966 .
## gender_feedback:poli_centered -0.01260    0.01576  -0.799 0.424248
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4732 on 996 degrees of freedom
## Multiple R-squared:  0.02619,    Adjusted R-squared:  0.02326
## F-statistic:  8.93 on 3 and 996 DF,  p-value: 7.676e-06
```

```
robust_confint(r_poli_moderation)
```

```
##                                    2.5 %      97.5 %
## (Intercept)                    0.24500042 0.326024937
## gender_feedback                0.05618713 0.175474510
## poli_centered                 -0.04049028 0.001295435
## gender_feedback:poli_centered -0.04353560 0.018332429
```

```
# Simple slopes analysis at different levels of political ideology
# Liberal (1 SD below mean, approximately liberal)
liberal_slope <- coef(r_poli_moderation)["gender_feedback"] +
                 coef(r_poli_moderation)["gender_feedback:poli_centered"] * (-2)

# Moderate (at mean, centered = 0)
moderate_slope <- coef(r_poli_moderation)["gender_feedback"]

# Conservative (1 SD above mean, approximately conservative)
```

```r
conservative_slope <- coef(r_poli_moderation)["gender_feedback"] +
                      coef(r_poli_moderation)["gender_feedback:poli_centered"] * (2)

cat("\n\nSimple Slopes Analysis:\n")
```

```
##
##
## Simple Slopes Analysis:
```

```r
cat("Effect of gender feedback for liberals (ideology = 2): ", round(liberal_slope * 100,
↪  2), "%\n")
```

```
## Effect of gender feedback for liberals (ideology = 2):  14.1 %
```

```r
cat("Effect of gender feedback for moderates (ideology = 4): ", round(moderate_slope *
↪  100, 2), "%\n")
```

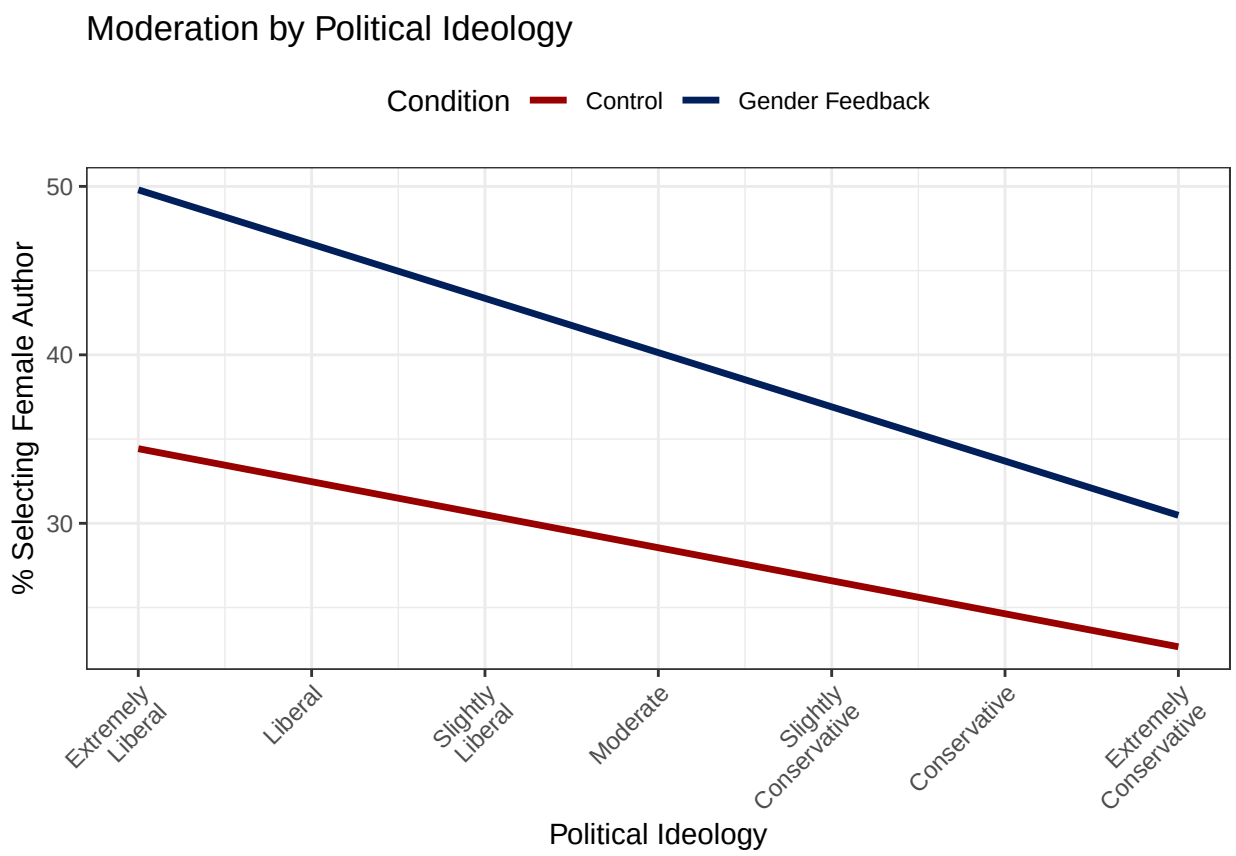
```
## Effect of gender feedback for moderates (ideology = 4):  11.58 %
```

```r
cat("Effect of gender feedback for conservatives (ideology = 6): ",
↪  round(conservative_slope * 100, 2), "%\n")
```

```
## Effect of gender feedback for conservatives (ideology = 6):  9.06 %
```

```r
# Visualization of interaction
library(ggplot2)
# Create prediction data
pred_data <- expand.grid(
  gender_feedback = c(0, 1),
  poli_centered = seq(-3, 3, by = 0.5)  # From very liberal to very conservative
)
pred_data$female_pick <- predict(r_poli_moderation, newdata = pred_data)
pred_data$poli_label <- pred_data$poli_centered + 4  # Convert back to 1-7 scale

# Plot interaction
p_interaction <- ggplot(pred_data, aes(x = poli_label, y = female_pick * 100,
                                       color = factor(gender_feedback))) +
  geom_line(size = 1.2) +
  scale_color_manual(values = c("0" = "#990000", "1" = "#011F5B"),
                     labels = c("0" = "Control", "1" = "Gender Feedback")) +
  scale_x_continuous(breaks = 1:7,
                     labels = c("Extremely\nLiberal", "Liberal", "Slightly\nLiberal",
                                "Moderate", "Slightly\nConservative", "Conservative",
                                "Extremely\nConservative")) +
  labs(x = "Political Ideology",
       y = "% Selecting Female Author",
       title = "Moderation by Political Ideology",
       color = "Condition") +
  theme_bw() +
  theme(legend.position = "top",
```

```
        axis.text.x = element_text(angle = 45, hjust = 1))

print(p_interaction)
```



Moderation by Political Party

```
# Test for moderation by political party affiliation
# Republican is the reference category
r_party_moderation <- lm(female_pick ~ gender_feedback * party_democrat +
                                       gender_feedback * party_independent, data=d0)

# Display the summary with robust standard errors
cat("Moderation by Political Party Affiliation\n")
```

```
## Moderation by Political Party Affiliation
```

```
robust_summary(r_party_moderation)
```

```
##
## Call:
## lm(formula = female_pick ~ gender_feedback * party_democrat +
##     gender_feedback * party_independent, data = d0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4465 -0.3662 -0.2885  0.5758  0.7246
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       0.27536    0.03830   7.189 1.28e-12 ***
## gender_feedback                   0.09083    0.05590   1.625    0.104
## party_democrat                    0.03865    0.05018   0.770    0.441
## party_independent                 0.01310    0.05291   0.248    0.805
## gender_feedback:party_democrat    0.01940    0.07363   0.263    0.792
## gender_feedback:party_independent 0.06724    0.07766   0.866    0.387
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4755 on 994 degrees of freedom
## Multiple R-squared:  0.0184, Adjusted R-squared:  0.01346
## F-statistic: 3.727 on 5 and 994 DF,  p-value: 0.002393
```

```
robust_confint(r_party_moderation)
```

```
##                                      2.5 %    97.5 %
## (Intercept)                       0.20019849 0.3505262
## gender_feedback                  -0.01886163 0.2005314
## party_democrat                   -0.05982011 0.1371148
## party_independent                -0.09073614 0.1169346
## gender_feedback:party_democrat   -0.12509587 0.1638917
## gender_feedback:party_independent -0.08515978 0.2196487
```

```r
# Simple effects analysis for each party
# Effect for Republicans (reference category)
republican_effect <- coef(r_party_moderation)["gender_feedback"]

# Effect for Democrats
democrat_effect <- coef(r_party_moderation)["gender_feedback"] +
                   coef(r_party_moderation)["gender_feedback:party_democrat"]

# Effect for Independents
independent_effect <- coef(r_party_moderation)["gender_feedback"] +
                      coef(r_party_moderation)["gender_feedback:party_independent"]

cat("\n\nSimple Effects Analysis:\n")
```

```
##
##
## Simple Effects Analysis:
```

```r
cat("Effect of gender feedback for Republicans: ", round(republican_effect * 100, 2),
↪  "%\n")
```

```
## Effect of gender feedback for Republicans:  9.08 %
```

```r
cat("Effect of gender feedback for Democrats: ", round(democrat_effect * 100, 2), "%\n")
```

```
## Effect of gender feedback for Democrats:  11.02 %
```

```r
cat("Effect of gender feedback for Independents: ", round(independent_effect * 100, 2),
↪  "%\n")
```

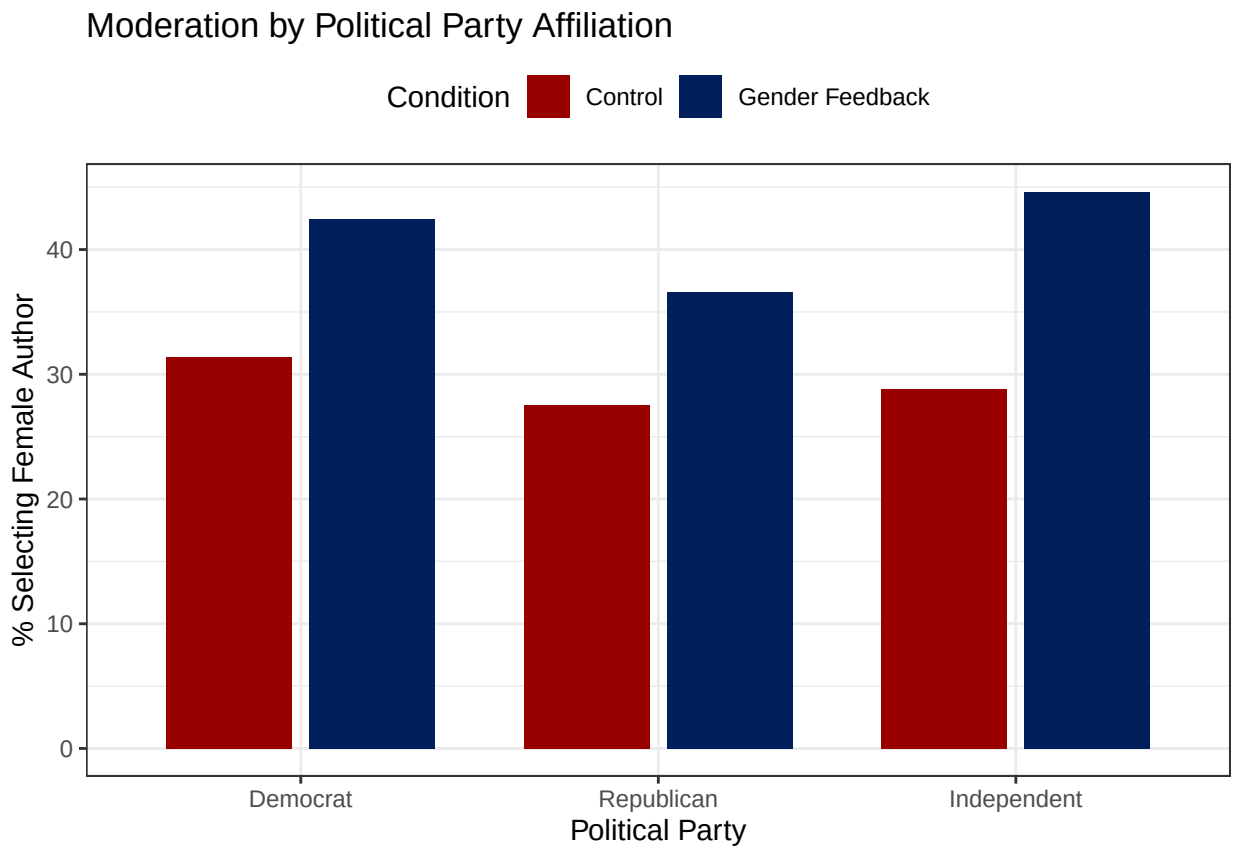
```
## Effect of gender feedback for Independents:  15.81 %
```

```r
# Compute means for each group and condition
party_means <- d0 |>
  group_by(poli_party, gender_feedback) |>
  summarize(
    mean_female_pick = mean(female_pick, na.rm = TRUE),
    n = n()
  ) |>
  filter(!is.na(poli_party))

print(party_means)
```

```
## # A tibble: 6 x 4
## # Groups:   poli_party [3]
##   poli_party  gender_feedback mean_female_pick     n
##   <ord>                 <dbl>            <dbl> <int>
## 1 Democrat                  0            0.314   207
## 2 Democrat                  1            0.424   198
```

```
## 3 Republican                0            0.275   138
## 4 Republican                1            0.366   142
## 5 Independent               0            0.288   156
## 6 Independent               1            0.447   159
```

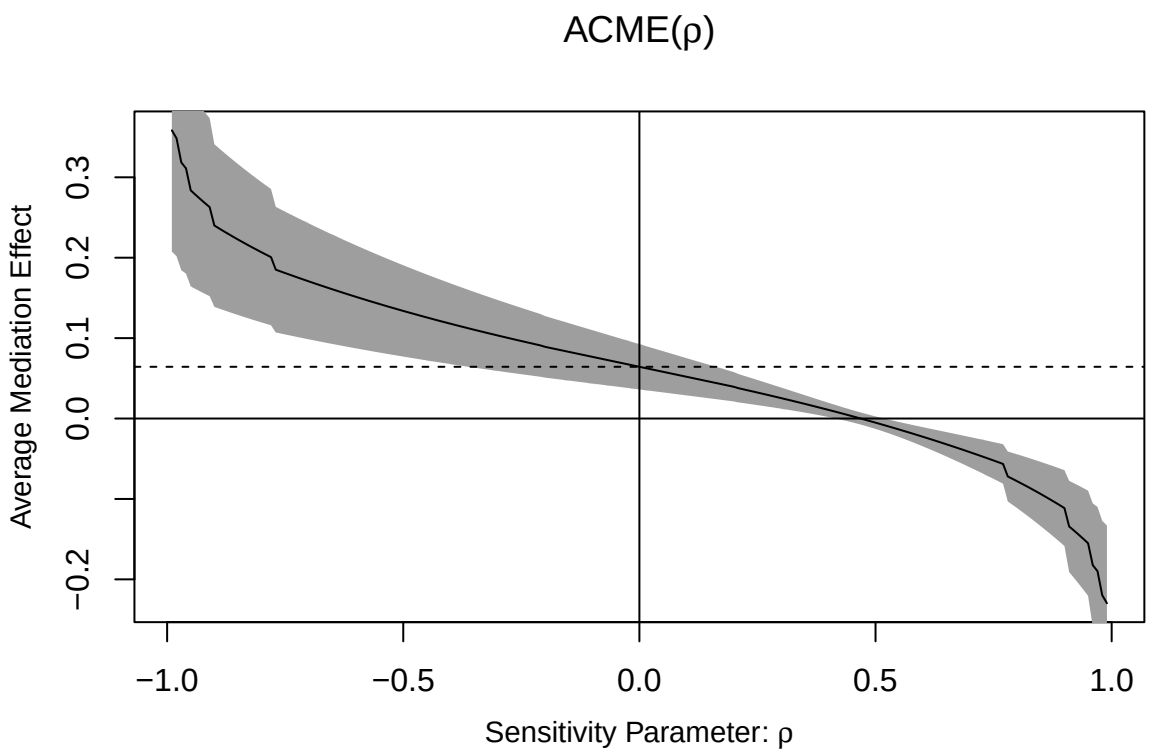
```r
# Visualization of interaction
p_party_interaction <- ggplot(party_means, aes(x = poli_party, y = mean_female_pick *
↪  100,
                                               fill = factor(gender_feedback))) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.8), width = 0.7) +
  scale_fill_manual(values = c("0" = "#990000", "1" = "#011F5B"),
                    labels = c("0" = "Control", "1" = "Gender Feedback")) +
  labs(x = "Political Party",
       y = "% Selecting Female Author",
       title = "Moderation by Political Party Affiliation",
       fill = "Condition") +
  theme_bw() +
  theme(legend.position = "top")

print(p_party_interaction)
```



Secondary Analysis

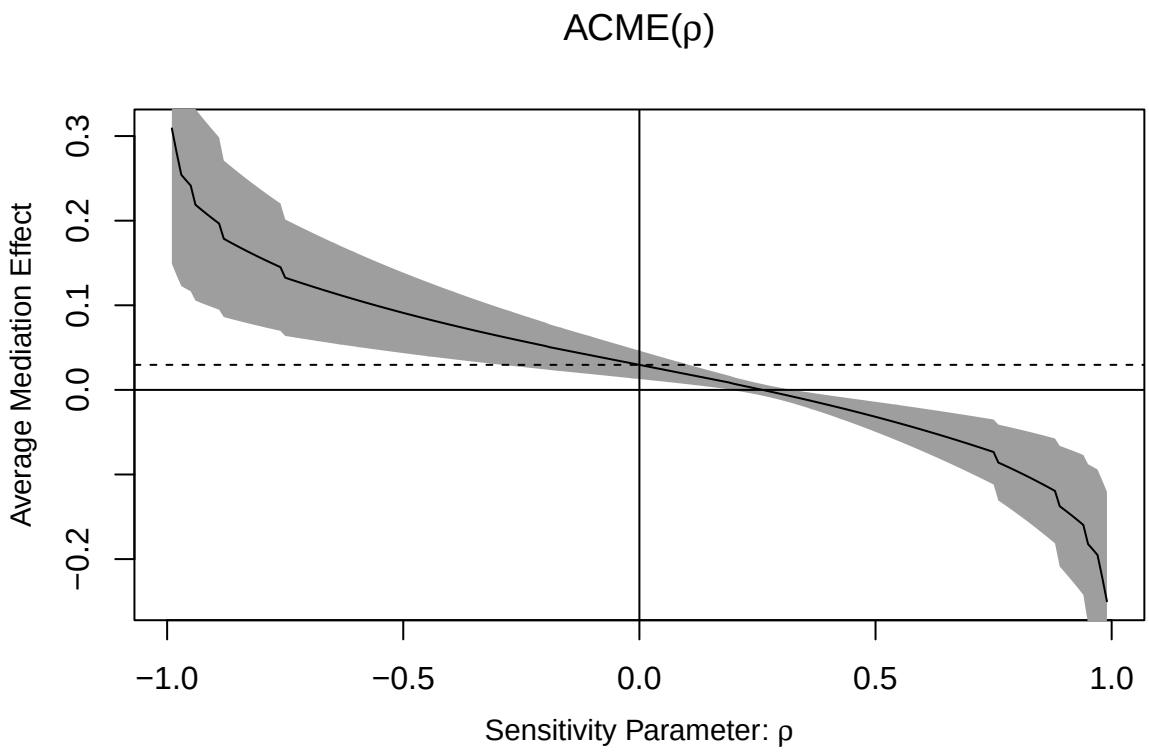
Mediation

```
## Sobel test for Internal Motivation

## $statistic
## internal
## 17.04891
##
## $p_value
## internal
##        0
##
## $se
##   internal
## 0.01431617

##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##                 Estimate 95% CI Lower 95% CI Upper p-value
## ACME           0.0643434    0.0369389    0.0936615  <2e-16 ***
## ADE            0.0550770    0.0018497    0.1074007  0.0418 *
## Total Effect   0.1194205    0.0594097    0.1776128  0.0002 ***
## Prop. Mediated 0.5387974    0.3252691    0.9710058  0.0002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 1000
##
##
## Simulations: 10000
```




```
## Sobel test for External Motivation

## $statistic
## external
## 8.443706
##
## $p_value
## external
##        0
##
## $se
##   external
## 0.01625515

##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##                Estimate 95% CI Lower 95% CI Upper p-value
## ACME           0.029583     0.013617     0.047779  <2e-16 ***
## ADE            0.089837     0.031995     0.145893  0.0014 **
## Total Effect   0.119420     0.060807     0.177152  <2e-16 ***
## Prop. Mediated 0.247724     0.113456     0.510151  <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 1000
##
##
## Simulations: 10000
```




```
## Correlation Between Internal and External:  0.7792483

## P-value:  9.693804e-205

## Combined Multiple Mediation Model Results

##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##                  Estimate 95% CI Lower 95% CI Upper p-value
## ACME            0.0906828    0.0518569    0.1307623  <2e-16 ***
## ADE             0.0563528    0.0036533    0.1078551  0.0372 *
## Total Effect    0.1470355    0.0819772    0.2109770  <2e-16 ***
## Prop. Mediated 0.6167404    0.4011862    0.9614873  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Sample Size Used: 1000
##
##
## Simulations: 10000

##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##                  Estimate 95% CI Lower 95% CI Upper p-value
## ACME            -0.0276151   -0.0457692   -0.0121979  <2e-16 ***
## ADE              0.0563528    0.0047562    0.1095582  0.0318 *
## Total Effect     0.0287377   -0.0260970    0.0852633  0.3122
## Prop. Mediated  -0.9609342  -10.6201618   10.3787817  0.3122
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 1000
##
##
## Simulations: 10000
```

Figure S1 Code

```r
# Get p-values from regression models
p_gender <- robust_summary(r1)$coefficients["gender_feedback", "Pr(>|t|)"]
p_sold30m <- robust_summary(r_sold30m)$coefficients["sold30m_shown", "Pr(>|t|)"]
p_forms <- robust_summary(r_forms)$coefficients["forms_shown", "Pr(>|t|)"]
p_classic <- robust_summary(r_classic)$coefficients["classic_shown", "Pr(>|t|)"]

# Function to convert p-value to significance stars
get_sig_stars <- function(p) {
  if (p < 0.001) return("***")
  else if (p < 0.01) return("**")
  else if (p < 0.05) return("*")
  else return("n.s.")
}

# Get significance labels
sig_gender <- get_sig_stars(p_gender)
sig_sold30m <- get_sig_stars(p_sold30m)
sig_forms <- get_sig_stars(p_forms)
sig_classic <- get_sig_stars(p_classic)

dfemale_plot <- d0 |>
  dplyr::select(gender_feedback, female_pick) |>
  dplyr::group_by(gender_feedback) |>
  dplyr::summarise(
    n = n(),
    freq = mean(female_pick),
```

```
    sd = sd(female_pick) * 100,
    se = (sd(female_pick) / sqrt(n())) * 100
  ) |>
  dplyr::mutate(
    gender_feedback = case_when(
      gender_feedback == 1 ~ "\"Treatment\"",
      TRUE ~ "\"Control\""
    )
  ) |>
  dplyr::rename(Condition = gender_feedback)


##### sold 30M+

dsold30m_plot <- d0 |>
  dplyr::select(sold30m_shown, sold30m_pick) |>
  dplyr::group_by(sold30m_shown) |>
  dplyr::summarise(
    n = n(),
    freq = mean(sold30m_pick),
    sd = sd(sold30m_pick) * 100,
    se = (sd(sold30m_pick) / sqrt(n())) * 100
  )|>
  mutate(sold30m_shown = case_when(sold30m_shown==1 ~ "\"Treatment\"",
                           TRUE ~ "\"Control\"")) |>
  rename(Condition = sold30m_shown)


##### forms

dforms_plot <- d0 |>
  dplyr::select(forms_shown, forms_pick) |>
  dplyr::group_by(forms_shown) |>
  dplyr::summarise(
    n = n(),
    freq = mean(forms_pick),
    sd = sd(forms_pick) * 100,
    se = (sd(forms_pick) / sqrt(n())) * 100
  )|>
  mutate(forms_shown = case_when(forms_shown==1 ~ "\"Treatment\"",
                           TRUE ~ "\"Control\"")) |>
  rename(Condition = forms_shown)


#### classic

dclassic_plot <- d0 |>
  dplyr::select(classic_shown, classic_pick) |>
  dplyr::group_by(classic_shown) |>
  dplyr::summarise(
    n = n(),
    freq = mean(classic_pick),
    sd = sd(classic_pick) * 100,
    se = (sd(classic_pick) / sqrt(n())) * 100
```

```
  )|>
  mutate(classic_shown = case_when(classic_shown==1 ~ "\"Treatment\"",
                          TRUE ~ "\"Control\"")) |>
  rename(Condition = classic_shown)

## Combine plots

df_combined <- bind_rows(
  dforms_plot %>% mutate(Category = "\nWrote Multiple\nGenres", sig_label = sig_forms),
  dsold30m_plot %>% mutate(Category = "\n30M+ Copies\nSold", sig_label = sig_sold30m),
  dclassic_plot %>% mutate(Category = "\n50+ Years\nin Print", sig_label = sig_classic),
  dfemale_plot %>% mutate(Category = "\nWere Women", sig_label = sig_gender)
, .id = "id") %>%
  mutate(Category = factor(Category, levels = c('\nWrote Multiple\nGenres', '\n30M+
  ↪  Copies\nSold', '\n50+ Years\nin Print', '\nWere Women')))

p_combined <- ggplot(df_combined, aes(x = Condition, y = freq*100, fill = Condition)) +
  geom_bar(stat="identity", width = 0.85, position = position_dodge(width = 0.7)) +
  geom_text(aes(label=paste0(sprintf("%.1f", freq*100),"%")),
            position=position_dodge(width=0.7), vjust=5, size = 5, color = "white") +
  geom_errorbar(aes(ymin=freq*100-se, ymax=freq*100+se), width = .1, position =
  ↪  position_dodge(width = 0.7)) +
  facet_wrap(~factor(Category, c('\nWrote Multiple\nGenres', '\n30M+ Copies\nSold',
  ↪  '\n50+ Years\nin Print', '\nWere Women')), nrow = 1, strip.position = "bottom") +
   geom_segment(data = df_combined %>% filter(Condition == "\"Treatment\""),
                aes(x = 1, xend = 2, y = freq*100 + se + 5, yend = freq*100 + se + 5),
                inherit.aes = FALSE) +
   geom_text(data = df_combined %>% filter(Condition == "\"Treatment\""),
             aes(x = 1.5, y = freq*100 + se + 7, label = sig_label),
             inherit.aes = FALSE, vjust = 0, size = 5) +
  theme_bw() +
  scale_fill_manual(values = c("#990000", "#011F5B"), labels = c("No feedback provided",
  ↪  "Feedback provided"), "Feedback") +
  scale_y_continuous(labels = function(x) paste0(x,"%"), limits = c(0,90)) +
  scale_x_discrete(labels = c("\"Control\"" = "Not\nShown", "\"Treatment\"" = "Shown")) +
  labs(x = "Feedback on % of authors who...", y = "% of New Authors with the Target
  ↪  Attribute",title = "The Effect of Getting Feedback on Your Author Selections") +
  theme(plot.caption = element_text(face = "italic"),
        legend.position = c(0.5, 0.95),
        legend.title = element_blank(),
        legend.direction = "horizontal",
        legend.text = element_text(size = 20),
        legend.key.size = unit(7, 'mm'),
        legend.background = element_rect(fill = "white"),
        panel.grid.minor = element_blank(),
        panel.grid = element_blank(),
        panel.border = element_rect(fill= NA, color = "white"),
        plot.background = element_rect(fill = "white"),
        panel.background = element_rect(fill = "white"),
        axis.title.x = element_text(face="bold", size = 21, vjust = 17),
        plot.title = element_blank(),
        axis.title.y = element_text(size = 20, color = "black"),
        axis.text.x = element_blank(),
        axis.ticks = element_blank(),
```

```
        axis.text.y = element_text(size = 20, color = "black"),
        strip.text = element_text(size = 20, color = "black"),
        strip.background = element_rect(colour = "white", fill = "white"))

#p_combined
# Save the plot with Times New Roman font
ggsave("Figure-S1.pdf", plot = p_combined, width = 10, height = 8, units = "in", device =
↪  cairo_pdf, family = "Times New Roman")
```

System of Simultaneous Equations

```
##                                    Wald.Coefficient      P_Value
## Gender Feedback - Sold30M Feedback        11.484007 7.156490e-04
## Gender Feedback - Forms Feedback          37.919382 8.893642e-10
## Gender Feedback - Classic Feedback         9.005155 2.725614e-03
```