Study 1 (Gender NPR Study)

December 12, 2025

Items

Variable Names

Variable	Description
treatment	Binary indicator of whether a participant was randomly assigned to treatment condition (shown women feedback).
set_num	Indicator of which feedback set was shown (1 or 2, with different percentage values).
women_feedback	Binary indicator of whether women feedback was shown to participant.
women_count	Count of women selected across the three choices (0-3).
women_proportion	Proportion of women selected (DV: ranges from 0 to 1).
age_feedback	Binary indicator of whether age feedback was shown.
age_proportion	Proportion of experts under 50 years old selected.
location_feedback	Binary indicator of whether location feedback was shown.
location_proportion	Proportion of experts based on West Coast selected.
university_feedback	Binary indicator of whether university feedback was shown.
university_proportion	Proportion of experts working at a university selected.
choice-1 to choice-3	The selected AI experts
gender	Self-selected gender.
race	Self-selected race.
age	Self-entered age.
gender_code	Dummy code for gender (male = 1).
race_code	Dummy code for race (white = 1).

Demographics

```
## Excluded Participants: 495
```

```
##                         Percentage gender
## 1                            Woman  55.21
## 2                              Man  43.89
## 3                       Non-binary   0.90
## 4 Another gender not listed here:   0.00
```

```
##                           Percentage  Race
## 1 American Indian or Alaskan Native  0.80
## 2          Asian / Pacific Islander  7.21
## 3         Black or African American 13.13
## 4                 Hispanic / Latinx  6.51
## 5                 White / Caucasian 72.34
```

```
## # A tibble: 1 x 2
##   mean_age sd_age
##      <dbl>  <dbl>
## 1     43.8   13.2
```

```
## Treatment condition:  50 %
```

```
## Control condition:  50 %
```

```
## Set 1:  52 %
```

```
## Set 2:  48 %
```

```
## Mean proportion of women selected:  0.399
```

```
## SD proportion of women selected:  0.268
```

```
## Treatment × Set interaction: B = 0.041, p = 0.219
```

```
## Main effect of stimulus set: B = 0.001, p = 0.948
```

Primary Analysis

```r
# Primary model: Effect of treatment on proportion of women selected
# As preregistered: includes treatment (gender feedback) and Set1 indicator
r1 <- lm(women_proportion ~ treatment + set_num, data=d0)

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

```
##
## Call:
## lm(formula = women_proportion ~ treatment + set_num, data = d0)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.47288 -0.13955 -0.01445  0.19378  0.67401
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.30420    0.02638  11.530  < 2e-16 ***
## treatment    0.12510    0.01650   7.580 7.87e-14 ***
## set_num      0.02179    0.01654   1.317    0.188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2608 on 997 degrees of freedom
## Multiple R-squared:  0.05577,    Adjusted R-squared:  0.05387
## F-statistic: 29.44 on 2 and 997 DF,  p-value: 3.778e-13
```

```r
robust_confint(r1)
```

```
##                   2.5 %     97.5 %
## (Intercept)  0.25242392 0.35597448
## treatment    0.09271652 0.15748842
## set_num     -0.01067624 0.05425673
```

Robustness

```
## Model 2: Which feedback was shown with women, remove constant due to collinearity
r2 <- lm(women_proportion ~ women_feedback + age_feedback + location_feedback +
↪  university_feedback - 1, data=d0)

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

```
##
## Call:
## lm(formula = women_proportion ~ women_feedback + age_feedback +
##     location_feedback + university_feedback - 1, data = d0)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.46970 -0.13636 -0.00333  0.19697  0.66333
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## women_feedback       0.23440    0.01476  15.881  < 2e-16 ***
## age_feedback         0.11945    0.01892   6.315 4.07e-10 ***
## location_feedback    0.11584    0.01947   5.950 3.70e-09 ***
## university_feedback  0.10137    0.01660   6.107 1.46e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2611 on 996 degrees of freedom
## Multiple R-squared:  0.7061, Adjusted R-squared:  0.705
## F-statistic: 598.4 on 4 and 996 DF,  p-value: < 2.2e-16
```

```
robust_confint(r2)
```

```
##                         2.5 %    97.5 %
## women_feedback      0.20543707 0.2633632
## age_feedback        0.08233186 0.1565759
## location_feedback   0.07763895 0.1540469
## university_feedback 0.06879546 0.1339442
```

```
## Model 3: Robust to demographic controls
r3 <- lm(women_proportion ~ women_feedback + gender_code + race_code + age, data=d0)

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

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

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.51243 -0.16417 -0.04477  0.16924  0.72109
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     3.949e-01  3.307e-02  11.942  < 2e-16 ***
## women_feedback  1.196e-01  1.635e-02   7.318 5.19e-13 ***
## gender_code    -9.763e-02  1.647e-02  -5.927 4.24e-09 ***
## race_code      -1.522e-02  1.850e-02  -0.823    0.411
## age            -4.289e-05  6.469e-04  -0.066    0.947
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2565 on 993 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.08684,    Adjusted R-squared:  0.08317
## F-statistic: 23.61 on 4 and 993 DF,  p-value: < 2.2e-16
```

```
robust_confint(r3)
```

```
##                      2.5 %       97.5 %
## (Intercept)     0.330039664  0.459830682
## women_feedback  0.087555376  0.151719828
## gender_code    -0.129955492 -0.065308997
## race_code      -0.051534867  0.021088453
## age            -0.001312263  0.001226475
```

```
## logistic regression
# Fit the logistic regression model using the third (final) choice
r4 <- glm(women_choice3 ~ women_feedback, family = binomial, data=d0)

# 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)       0.524    0.0941     -6.86 7.04e-12    0.435     0.630
## 2 women_feedback    1.31     0.131       2.09 3.69e- 2    1.02      1.70
```

```
summary(r4)
```

```
##
## Call:
## glm(formula = women_choice3 ~ women_feedback, family = binomial,
##     data = d0)
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)    -0.64552    0.09414  -6.857 7.04e-12 ***
## women_feedback  0.27328    0.13093   2.087   0.0369 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1324.1  on 999  degrees of freedom
## Residual deviance: 1319.8  on 998  degrees of freedom
## AIC: 1323.8
##
## Number of Fisher Scoring iterations: 4
```

Secondary Analysis: Other Attributes

```
## Effect of age feedback:

##
## Call:
## lm(formula = age_proportion ~ age_feedback, data = d0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1788 -0.1788 -0.1532  0.1545  0.8212
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.15323    0.01854   8.265 4.42e-16 ***
## age_feedback  0.02562    0.01982   1.293    0.196
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2069 on 998 degrees of freedom
## Multiple R-squared:  0.001665,   Adjusted R-squared:  0.0006649
## F-statistic: 1.665 on 1 and 998 DF,  p-value: 0.1973

##                    2.5 %     97.5 %
## (Intercept)   0.11684568 0.18960593
## age_feedback -0.01327099 0.06450583

##
## Effect of location feedback:

##
## Call:
## lm(formula = location_proportion ~ location_feedback, data = d0)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.25866 -0.25866  0.07467  0.07467  0.74134
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.21393    0.01889  11.325   <2e-16 ***
## location_feedback  0.04473    0.02063   2.168   0.0304 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2405 on 998 degrees of freedom
## Multiple R-squared:  0.004007,   Adjusted R-squared:  0.003009
## F-statistic: 4.015 on 1 and 998 DF,  p-value: 0.04535

##                         2.5 %    97.5 %
## (Intercept)       0.176862344 0.2509984
## location_feedback 0.004251318 0.0852090
```

```
##
## Effect of university feedback:

##
## Call:
## lm(formula = university_proportion ~ university_feedback, data = d0)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.24670 -0.24670  0.08663  0.08663  0.81680
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.18320    0.01515  12.096  < 2e-16 ***
## university_feedback  0.06351    0.01769   3.589 0.000348 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2478 on 998 degrees of freedom
## Multiple R-squared:  0.01192,    Adjusted R-squared:  0.01093
## F-statistic: 12.04 on 1 and 998 DF,  p-value: 0.0005417

##                         2.5 %     97.5 %
## (Intercept)         0.15347514 0.21291604
## university_feedback 0.02878294 0.09822957
```

Figure 3A

System of Simultaneous Equations

```
## SUR Model Coefficients:

## ======================

## Women feedback effect: 0.0956

## Age feedback effect: 0.0378

## Location feedback effect: 0.0275

## University feedback effect: 0.033

##
##
## Wald Tests for Cross-Equation Comparisons:

## ========================================

## Test 1: Women Feedback Effect vs. Age Feedback Effect

## --------------------------------------------------

## Linear hypothesis test (Theil's F test)
##
## Hypothesis:
## womeneq_women_feedback - ageeq_age_feedback = 0
##
## Model 1: restricted model
## Model 2: sur_model
##
##   Res.Df Df     F  Pr(>F)
## 1   3989
## 2   3988  1 6.629 0.01007 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
##
## Test 2: Women Feedback Effect vs. Location Feedback Effect

## ------------------------------------------------------------

## Linear hypothesis test (Theil's F test)
##
## Hypothesis:
## womeneq_women_feedback - locationeq_location_feedback = 0
##
## Model 1: restricted model
## Model 2: sur_model
```

```
##
##   Res.Df Df      F   Pr(>F)
## 1   3989
## 2   3988  1 8.3509 0.003876 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
##
## Test 3: Women Feedback Effect vs. University Feedback Effect

## ------------------------------------------------------------

## Linear hypothesis test (Theil's F test)
##
## Hypothesis:
## womeneq_women_feedback - universityeq_university_feedback = 0
##
## Model 1: restricted model
## Model 2: sur_model
##
##   Res.Df Df      F    Pr(>F)
## 1   3989
## 2   3988  1 12.841 0.0003431 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
##
## Summary of Wald Tests:

## ======================

##                           Test F_Statistic  P_Value Significant
##         Women vs. Age Feedback        6.63 0.010069         Yes
##    Women vs. Location Feedback        8.35 0.003876         Yes
##  Women vs. University Feedback       12.84 0.000343         Yes
```