959_extension

Dasha Castillo

2023-05-11

9.59 Replication Project

Below, we'll be loading the data of the replication of the Wu & Gibson (2021) experiment.

200 participants were able to take the survey for the experiment.

```r
# loading data for "english replication" aka eng_rep
en_ext <- read_csv("https://raw.githubusercontent.com/dashacastillo/959_wu_gibson_replication/main/9.59
```

```
## Rows: 803 Columns: 10
## -- Column specification ---------------------------------------------------------
## Delimiter: ","
## chr  (7): participant_id, experiment, cond, literal_response, response, engl...
## dbl  (2): item, presentation_order
## dttm (1): timestamp
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
en_ext$response = tolower(en_ext$response)
en_ext
```

```
## # A tibble: 803 x 10
##    participant~1 exper~2  item cond  prese~3 liter~4 respo~5 timestamp
##    <chr>         <chr>   <dbl> <chr>   <dbl> <chr>   <chr>   <dttm>
##  1 5c4c5bc98897~ size        4 top_~       0 Yes     small ~ 2023-05-11 04:46:09
##  2 5c4c5bc98897~ size        2 top_~       1 No      large ~ 2023-05-11 04:46:09
##  3 5c4c5bc98897~ size        1 bott~       2 Yes     small ~ 2023-05-11 04:46:09
##  4 5c4c5bc98897~ size        3 bott~       3 No      large ~ 2023-05-11 04:46:09
##  5 5ea9ee70bd6e~ size        4 bott~       0 No      diamond 2023-05-11 04:48:40
##  6 5ea9ee70bd6e~ size        1 bott~       1 Yes     tirang~ 2023-05-11 04:48:40
##  7 5ea9ee70bd6e~ size        2 top_~       2 No      rectan~ 2023-05-11 04:48:40
##  8 5ea9ee70bd6e~ size        3 top_~       3 Yes     square  2023-05-11 04:48:40
##  9 5d89c4dd68f9~ size        4 top_~       0 Yes     small ~ 2023-05-11 04:57:33
## 10 5d89c4dd68f9~ size        2 top_~       1 No      big re~ 2023-05-11 04:57:33
## # ... with 793 more rows, 2 more variables: english_first_language <chr>,
## #   country_of_birth <chr>, and abbreviated variable names 1: participant_id,
## #   2: experiment, 3: presentation_order, 4: literal_response, 5: response
```

```r
# loading data for "spanish replication" aka sp_rep
sp_ext <- read_csv("https://raw.githubusercontent.com/dashacastillo/959_wu_gibson_replication/main/9.59)
```

```
## Rows: 784 Columns: 10
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (7): participant_id, experiment, cond, literal_response, response, engl...
## dbl  (2): item, presentation_order
## dttm (1): timestamp
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
sp_ext$response = tolower(sp_ext$response)
sp_ext
```

```
## # A tibble: 784 x 10
##    participant~1 exper~2  item cond  prese~3 liter~4 respo~5 timestamp
##    <chr>         <chr>   <dbl> <chr>   <dbl> <chr>   <chr>   <dttm>
##  1 610cca28e74f~ size        4 top_~       0 Yes     estrel~ 2023-05-11 04:58:25
##  2 610cca28e74f~ size        1 bott~       1 No      triáng~ 2023-05-11 04:58:25
##  3 610cca28e74f~ size        2 top_~       2 No      rectán~ 2023-05-11 04:58:25
##  4 610cca28e74f~ size        3 bott~       3 Yes     cruz    2023-05-11 04:58:25
##  5 613aa39321ee~ size        2 top_~       0 Yes     rectan~ 2023-05-11 04:58:57
##  6 613aa39321ee~ size        1 bott~       1 Yes     triang~ 2023-05-11 04:58:57
##  7 613aa39321ee~ size        3 top_~       2 No      cuadra~ 2023-05-11 04:58:57
##  8 613aa39321ee~ size        4 bott~       3 No      rombo   2023-05-11 04:58:57
##  9 5f31de301a46~ size        2 bott~       0 Yes     círculo 2023-05-11 04:59:25
## 10 5f31de301a46~ size        1 bott~       1 No      triáng~ 2023-05-11 04:59:25
## # ... with 774 more rows, 2 more variables: english_first_language <chr>,
## #   country_of_birth <chr>, and abbreviated variable names 1: participant_id,
## #   2: experiment, 3: presentation_order, 4: literal_response, 5: response
```

Data Organization

The main goal of the experiment is to see the redundancy in describing the shapes. One way to see whether or not a redundant describer was given is to find all of the unique values of the responses and then making a new colum, Redundant, in which a value of 1 refers to whether a redundant modification was used, and a value of 0 refers to a non-redundant description.

```r
# all unique english descriptions from the responses
en_descs <- unique(en_ext$response)

# made the executive decision that if the response contains shape word(s) and size word(s), then it goe
# this includes if there are also other adjectives
en_red <- en_descs[c(1:4, 9:12, 14:17, 28:31, 34, 35, 40:42, 44, 46, 48:51, 54:57,
                     60:65, 68:72, 75, 76, 83:87, 89, 90, 98, 99, 100, 101, 104, 106,
                     108:113, 117, 126:128, 131:134)]

en_nonred <- en_descs[c(5:8, 13, 18:27, 32, 33, 36:39, 43, 45, 47, 52, 53, 58, 59, 66,
                        67, 73, 78:80, 88, 93, 102, 103, 105, 107, 118:125, 129, 130)]
```

```r
en_throwaway <- en_descs[c()]

en_unsure <- en_descs[c(74, 77, 81, 82, 91, 92, 94, 95, 96, 97, 115, 116, 135:138)]


# all unique spanish descriptions from the responses
sp_descs <- unique(sp_ext$response)

# made the executive decision that if the response contains shape word(s) and size word(s), then it goes
# this includes if there are also other adjectives
sp_red <- sp_descs[c(12, 14, 16:21, 26:29, 32:35, 38, 39, 41:44, 48:56, 61, 62, 65,
                     69:76, 81, 82, 85, 90, 93:96, 104:106, 108:113, 120:127, 134, 136,
                     138, 139, 141, 142, 147:149, 151, 152, 155:159, 171:174)]

sp_nonred <- sp_descs[c(1:11, 22:25, 30, 31, 36, 37, 40, 45:47, 57, 63, 66:68, 98:100,
                        107, 114, 117:119, 128:133, 135, 137, 140, 143, 150, 160, 165:170)]

sp_throwaway <- sp_descs[c(83, 84, 86, 87, 89, 91, 92, 97, 101:103, 115, 116, 175:177)]

sp_unsure <- sp_descs[c(13, 15, 58:60, 64, 77:80, 88, 144:146, 153, 154, 161:164)]
```

Adding a New Column to Show if Redundancy Exists in Response

```r
# first delete the unsure and throwaway items!
# we can't analyze what we can't put in a binary

en_extension_data = subset(en_ext,
                           response!="black circle" &
                             response!="black heart" &
                             response!="a black dot" &
                             response!="a black 5 sided star" &
                             response!="square with a star in it" &
                             response!="a triable inside a square" &
                             response!="a square with a plus inside of it" &
                             response!="black triangle" &
                             response!="black dot" &
                             response!="black square" &
                             response!="only a filled in circle" &
                             response!="black cross" &
                             response!="black square in red square upper right" &
                             response!="black diamond in black square upper left" &
                             response!="black heart in red square upper right" &
                             response!="black circle in red square lower right")
# summary(en_replication_data)

en_ext_redundancy = mutate(en_extension_data, redundant = if_else(en_extension_data$response %in% en_re
# test to see
# en_rep_redundancy[c("response", "redundant")]


# first delete the unsure and throwaway items!
# we can't analyze what we can't put in a binary
```

```r
sp_extension_data = subset(sp_ext,
                           response!="desconocido" &
                             response!="imagen no disponible" &
                             response!="no carga la imagen" &
                             response!="imagen no es visible" &
                             response!="no se ve la imagen" &
                             response!="no cargo" &
                             response!="the image is not loading, i have tried several times but it's o
                             response!="not loading" &
                             response!="no se visualizo la imagen" &
                             response!="no se visualizo  lq imagen" &
                             response!="no se ve" &
                             response!="figura geometrica de 5 picos que representa algo en el cielo de
                             response!="figura geometrica sin inicio ni fin" &
                             response!="4" &
                             response!="2" &
                             response!="3" &
                             response!="rombo negro" &
                             response!="cuadro negro" &
                             response!="circulo negro" &
                             response!="cruz negra" &
                             response!="corazón negri" &
                             response!="cuadrado negro" &
                             response!="círculo negro." &
                             response!="estrella negra." &
                             response!="rectángulo negro" &
                             response!="círculo sólido negro" &
                             response!="corazón sólido negro" &
                             response!="estrella sólida negra" &
                             response!="rectangulo negro" &
                             response!="triangulo negro" &
                             response!="circle" &
                             response!="square" &
                             response!="diamond" &
                             response!="heart")
# summary(sp_replication_data)

sp_ext_redundancy = mutate(sp_extension_data, redundant = if_else(sp_extension_data$response %in% sp_re
# test to see
# sp_rep_redundancy[c("response", "redundant")]
```
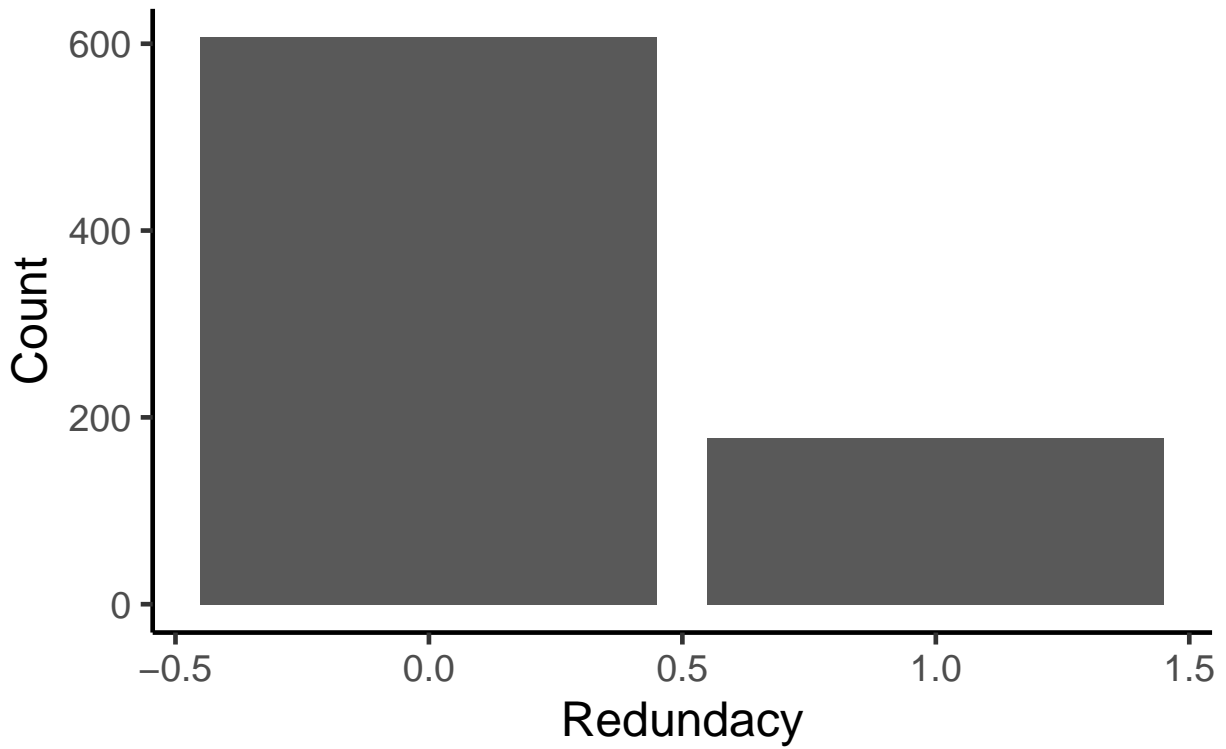
Making Figures

```r
ggplot(en_ext_redundancy, aes(x = redundant)) +
  geom_bar() +
  ggtitle("Size Redundancy in English") +
  xlab("Redundacy") + ylab("Count")
```

Size Redundancy in English


```
ggplot(sp_ext_redundancy, aes(x = redundant)) +
  geom_bar() +
  ggtitle("Size Redundancy in Spanish") +
  xlab("Redundacy") + ylab("Count")
```

Size Redundancy in Spanish