table2_part2_20251110.R

sahail

2025-11-10

```r
options(scipen = 8)

library(tidyverse)
```

```
## -- Attaching core tidyverse packages ---------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.2
## v ggplot2   4.0.0     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.1.0
## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(broom)

## real world data scenarios

# RWD a)
rwd_a <- crossing(trt = c("A", "B"),
        guess = c("A", "B", "DK")) %>%
  mutate(n = c(557, 427, 756,
               418, 573, 764)) %>%
  mutate(w = case_when(guess == "DK" ~ 0,
                       trt == guess ~ 1,
                       trt != guess ~ -1)) %>%
  uncount(n) %>%
  lm(w ~ 1, data = .) %>%
  tidy(., conf.int = TRUE)
rwd_a
```

```
## # A tibble: 1 x 7
##   term        estimate std.error statistic  p.value conf.low conf.high
##   <chr>          <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 (Intercept)   0.0815    0.0126      6.45 1.27e-10   0.0568     0.106
```

```r
# RWD b)
rwd_b <- crossing(trt = c("A", "B"),
        guess = c("A", "B", "DK")) %>%
```

```
  mutate(n = c(15, 18, 0,
             13, 18, 5)) %>%
  mutate(w = case_when(guess == "DK" ~ 0,
                     trt == guess ~ 1,
                     trt != guess ~ -1)) %>%
  uncount(n) %>%
  lm(w ~ 1, data = .) %>%
  tidy(., conf.int = TRUE)

rwd_b
```

```
## # A tibble: 1 x 7
##   term        estimate std.error statistic p.value conf.low conf.high
##   <chr>          <dbl>     <dbl>     <dbl>   <dbl>    <dbl>     <dbl>
## 1 (Intercept)   0.0290     0.117     0.248   0.805   -0.204     0.262
```

```
# RWD c)
rwd_c1 <- crossing(trt = c("A", "B"),
        guess = c("def A", "poss A", "poss B", "def B", "DK")) %>%
  mutate(guess = factor(guess, c("def A", "poss A", "poss B", "def B", "DK"))) %>%
  arrange(trt, guess) %>%
  mutate(n = c(38, 44, 21, 4, 170,
             11, 16, 21, 8, 83)) %>%
  mutate(w = case_when(guess == "DK" ~ 0,
                     trt == "A" & guess == "def A" ~ 1,
                     trt == "B" & guess == "def B" ~ 1,
                     trt == "B" & guess == "def A" ~ -1,
                     trt == "A" & guess == "def B" ~ -1,
                     trt == "A" & guess == "poss A" ~ 0.5,
                     trt == "B" & guess == "poss B" ~ 0.5,
                     trt == "B" & guess == "poss A" ~ -0.5,
                     trt == "A" & guess == "poss B" ~ -0.5)) %>%
  uncount(n) %>%
  lm(w ~ 1, data = .) %>%
  tidy(., conf.int = TRUE)
rwd_c1
```

```
## # A tibble: 1 x 7
##   term        estimate std.error statistic      p.value conf.low conf.high
##   <chr>          <dbl>     <dbl>     <dbl>        <dbl>    <dbl>     <dbl>
## 1 (Intercept)    0.108    0.0217      4.97 0.000000959   0.0654     0.151
```

```
rwd_c2 <- crossing(trt = c("A", "B"),
        guess = c("def A", "poss A", "poss B", "def B", "DK")) %>%
  mutate(guess = factor(guess, c("def A", "poss A", "poss B", "def B", "DK"))) %>%
  arrange(trt, guess) %>%
  mutate(n = c(38, 44, 21, 4, 170,
             11, 16, 21, 8, 83)) %>%
  mutate(w = case_when(guess == "DK" ~ NA,
                     trt == "A" & guess == "def A" ~ 1,
                     trt == "B" & guess == "def B" ~ 1,
                     trt == "B" & guess == "def A" ~ -1,
```

```
                    trt == "A" & guess == "def B" ~ -1,
                    trt == "A" & guess == "poss A" ~ 0.5,
                    trt == "B" & guess == "poss B" ~ 0.5,
                    trt == "B" & guess == "poss A" ~ -0.5,
                    trt == "A" & guess == "poss B" ~ -0.5)) %>%
  uncount(n) %>%
  lm(w ~ 1, data = .) %>%
  tidy(., conf.int = TRUE)
rwd_c2
```

```
## # A tibble: 1 x 7
##   term        estimate std.error statistic    p.value conf.low conf.high
##   <chr>          <dbl>     <dbl>     <dbl>      <dbl>    <dbl>     <dbl>
## 1 (Intercept)    0.276    0.0530      5.21 0.000000561    0.171     0.381
```

```
tmp <- crossing(trt = c("A", "B"),
                guess = c("def A", "poss A", "poss B", "def B", "DK")) %>%
  mutate(guess = factor(guess, c("def A", "poss A", "poss B", "def B", "DK"))) %>%
  arrange(trt, guess) %>%
  mutate(n = c(38, 44, 21, 4, 170,
               11, 16, 21, 8, 83)) %>%
  mutate(w = case_when(guess == "DK" ~ 0,
                    trt == "A" & guess == "def A" ~ 1,
                    trt == "B" & guess == "def B" ~ 1,
                    trt == "B" & guess == "def A" ~ -1,
                    trt == "A" & guess == "def B" ~ -1,
                    trt == "A" & guess == "poss A" ~ 0.5,
                    trt == "B" & guess == "poss B" ~ 0.5,
                    trt == "B" & guess == "poss A" ~ -0.5,
                    trt == "A" & guess == "poss B" ~ -0.5)) %>%
  uncount(n)
rwd_c3 <- tmp %>%
  group_by(trt) %>%
  nest() %>%
  mutate(out = map(data, ~ lm(w ~ 1, data = .)),
         res = map(out, ~ tidy(., conf.int = TRUE))) %>%
  select(-data, -out) %>%
  unnest(res)
rwd_c3
```

```
## # A tibble: 2 x 8
## # Groups:   trt [2]
##   trt   term        estimate std.error statistic  p.value conf.low conf.high
##   <chr> <chr>          <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 A     (Intercept)  0.164      0.0258    6.37   7.71e-10   0.114     0.215
## 2 B     (Intercept) -0.00360    0.0384   -0.0937 9.25e- 1  -0.0795    0.0723
```

```
# RWD d)
rwd_d1 <- crossing(trt = c("A", "B"),
        guess = c("def A", "poss A", "poss B", "def B", "DK")) %>%
  mutate(guess = factor(guess, c("def A", "poss A", "poss B", "def B", "DK"))) %>%
  arrange(trt, guess) %>%
```

```r
  mutate(n = c(0, 4, 3, 0, 32,
               0, 6, 3, 0, 32)) %>%
  mutate(w = case_when(guess == "DK" ~ 0,
                       trt == "A" & guess == "def A" ~ 1,
                       trt == "B" & guess == "def B" ~ 1,
                       trt == "B" & guess == "def A" ~ -1,
                       trt == "A" & guess == "def B" ~ -1,
                       trt == "A" & guess == "poss A" ~ 0.5,
                       trt == "B" & guess == "poss B" ~ 0.5,
                       trt == "B" & guess == "poss A" ~ -0.5,
                       trt == "A" & guess == "poss B" ~ -0.5)) %>%
  uncount(n) %>%
  lm(w ~ 1, data = .) %>%
  tidy(., conf.int = TRUE)
rwd_d1
```

```
## # A tibble: 1 x 7
##   term          estimate std.error statistic p.value conf.low conf.high
##   <chr>            <dbl>     <dbl>     <dbl>   <dbl>    <dbl>     <dbl>
## 1 (Intercept)    -0.0125    0.0251    -0.498   0.620  -0.0625    0.0375
```

```r
rwd_d2 <- crossing(trt = c("A", "B"),
       guess = c("def A", "poss A", "poss B", "def B", "DK")) %>%
  mutate(guess = factor(guess, c("def A", "poss A", "poss B", "def B", "DK"))) %>%
  arrange(trt, guess) %>%
  mutate(n = c(0, 4, 3, 0, 32,
               0, 6, 3, 0, 32)) %>%
  mutate(w = case_when(guess == "DK" ~ 0,
                       trt == "A" & guess == "def A" ~ 1,
                       trt == "B" & guess == "def B" ~ 1,
                       trt == "B" & guess == "def A" ~ -1,
                       trt == "A" & guess == "def B" ~ -1,
                       trt == "A" & guess == "poss A" ~ 0.75,
                       trt == "B" & guess == "poss B" ~ 0.75,
                       trt == "B" & guess == "poss A" ~ -0.75,
                       trt == "A" & guess == "poss B" ~ -0.75)) %>%
  uncount(n) %>%
  lm(w ~ 1, data = .) %>%
  tidy(., conf.int = TRUE)
rwd_d2
```

```
## # A tibble: 1 x 7
##   term          estimate std.error statistic p.value conf.low conf.high
##   <chr>            <dbl>     <dbl>     <dbl>   <dbl>    <dbl>     <dbl>
## 1 (Intercept)    -0.0187    0.0377    -0.498   0.620  -0.0937    0.0562
```

```r
# RWD e)
rwd_e <- crossing(trt = c("A", "B"),
       guess = c("A", "B")) %>%
  mutate(n = c(42, 0,
               19, 2)) %>%
  mutate(w = case_when(guess == "DK" ~ 0,
```

```
                      trt == guess ~ 1,
                      trt != guess ~ -1)) %>%
  uncount(n) %>%
  lm(w ~ 1, data = .) %>%
  tidy(., conf.int = TRUE)
rwd_e
```

```
## # A tibble: 1 x 7
##   term        estimate std.error statistic p.value conf.low conf.high
##   <chr>          <dbl>     <dbl>     <dbl>   <dbl>    <dbl>     <dbl>
## 1 (Intercept)    0.397     0.117      3.40 0.00117    0.164     0.630
```

```
# RWD f)
tmp <- crossing(trt = c("A", "B", "C"),
         guess = c("A", "B", "C", "DK")) %>%
  mutate(n = c(41, 66, 30, 44,
               27, 72, 24, 51,
               22, 36, 64, 52)) %>%
  mutate(w = case_when(guess == "DK" ~ 0,
                       trt == guess ~ 1,
                       trt %in% c("A", "B") & guess == "C" ~ -1,
                       trt == "C" & guess %in% c("A", "B") ~ -1,
                       trt == "A" & guess == "B" ~ 0.5,
                       trt == "B" & guess == "A" ~ 0.5)) %>%
  uncount(n)

rwd_f1 <- tmp %>%
  lm(w ~ 1, data = .) %>%
  tidy(., conf.int = TRUE)
rwd_f1
```

```
## # A tibble: 1 x 7
##   term        estimate std.error statistic  p.value conf.low conf.high
##   <chr>          <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 (Intercept)    0.211    0.0322      6.56 1.32e-10    0.148     0.274
```

```
rwd_f2 <- tmp %>%
  group_by(trt) %>%
  nest() %>%
  mutate(out = map(data, ~ lm(w ~ 1, data = .)),
         res = map(out, ~ tidy(., conf.int = TRUE))) %>%
  select(-data, -out) %>%
  unnest(res)
rwd_f2
```

```
## # A tibble: 3 x 8
## # Groups:   trt [3]
##   trt   term        estimate std.error statistic  p.value conf.low conf.high
##   <chr> <chr>          <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 A     (Intercept)   0.243    0.0486      5.01 1.31e- 6    0.147     0.339
## 2 B     (Intercept)   0.353    0.0519      6.81 1.52e-10    0.251     0.456
## 3 C     (Intercept)   0.0345   0.0636      0.542 5.88e- 1  -0.0911     0.160
```
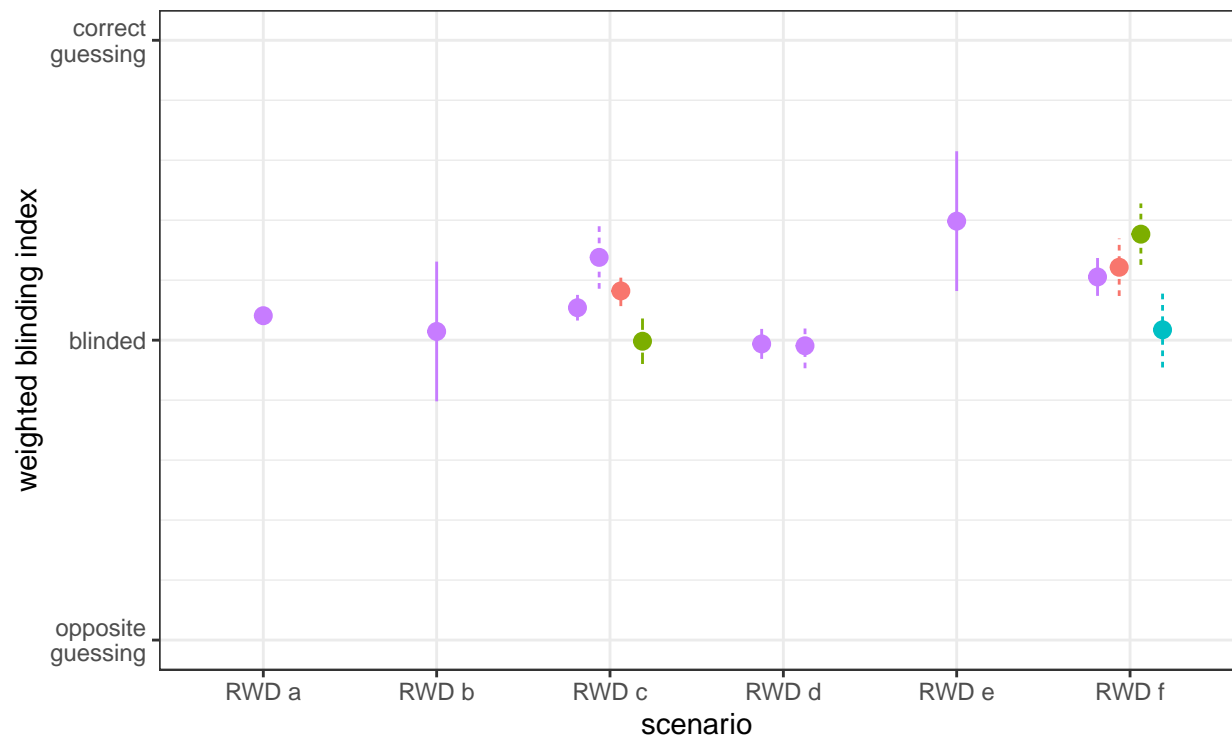
```
### comparison

bind_rows(rwd_a %>% mutate(scenario = "RWD a"),
          rwd_b %>% mutate(scenario = "RWD b"),
          rwd_c1 %>% mutate(scenario = "RWD c", opt = 1),
          rwd_c2 %>% mutate(scenario = "RWD c", opt = 2),
          rwd_c3 %>% mutate(scenario = "RWD c", opt = 3),
          rwd_d1 %>% mutate(scenario = "RWD d", opt = 1),
          rwd_d2 %>% mutate(scenario = "RWD d", opt = 2),
          rwd_e %>% mutate(scenario = "RWD e"),
          rwd_f1 %>% mutate(scenario = "RWD f", opt = 1),
          rwd_f2 %>% mutate(scenario = "RWD f", opt = 2)) %>%
  mutate(opt = as.character(opt)) %>%
  replace_na(list(opt = "1", trt = "overall")) %>%
  mutate(version = paste(opt, trt, sep = "")) %>%
  ggplot(aes(scenario, estimate,
             ymin = conf.low, ymax = conf.high,
             linetype = opt, color = trt)) +
  geom_pointrange(position = position_dodge(width = 0.5)) +
  ylab("weighted blinding index") +
  scale_y_continuous(limits = c(-1, 1),
                     breaks = c(-1, 0, 1),
                     labels = c("opposite\nguessing",
                                "blinded",
                                "correct\nguessing"),
                     minor_breaks = seq(-1, 1, 0.2)) +
  theme_bw() +
  theme(legend.position = "bottom")
```