E4-analysis-calibration

Rachel Ryskin

```
data.files <- list.files('data/run-2', full.names = TRUE)
data.tables <- lapply(data.files, function(file){
  data.table <- fromJSON(file)
  return(data.table)
})
all.data <- bind_rows(data.tables)
```

```
all.data.calib = all.data %>%
  filter(trial_type == "webgazer-validate") %>%
  dplyr::select(subject, trial_index,  percent_in_roi, average_offset) %>%
  tidyr::unnest(percent_in_roi)

summary.data.calib = all.data.calib %>%
  group_by(subject, trial_index) %>%
  summarize(mean_percent_in_roi = mean(percent_in_roi)) %>%
  group_by(subject) %>%
  mutate(calib_num = row_number())
```

```
## `summarise()` has grouped output by 'subject'. You can override using the
## `.groups` argument.
```

```
ggplot(summary.data.calib)+
  geom_line(aes(x = calib_num, y = mean_percent_in_roi, color=subject))+
  theme_bw()+
  theme(legend.position = "none")
```

```
summary.data.calib.wide = summary.data.calib %>%
  select(-trial_index) %>%
  pivot_wider(id_cols = subject, names_from=calib_num, values_from = mean_percent_in_roi )

# correlation between initial and halfway calibration
ggplot(summary.data.calib.wide %>% filter(is.na(`3`)), aes(x = `1`, y = `2`))+
  geom_point()+
  geom_smooth(method = 'lm')+
  theme_bw()+
  theme(legend.position = "none")

# correlation between 2 successive calibration attempts
ggplot(summary.data.calib.wide %>% filter(!is.na(`3`)), aes(x = `1`, y = `2`))+
  geom_point()+
  geom_smooth(method = 'lm')+
  theme_bw()+
  theme(legend.position = "none")
```

```r
# correlation between second attempt calibration and halfway
ggplot(summary.data.calib.wide %>% filter(!is.na(`3`)), aes(x = `2`, y = `3`))+
  geom_point()+
  geom_smooth(method = 'lm')+
  theme_bw()+
  theme(legend.position = "none")
```

```r
calib.by.subj = summary.data.calib %>%
  group_by(subject) %>%
  summarize(mean_percent_in_roi = mean(mean_percent_in_roi))

eyetracking.effects.by.subj = read_csv( "output/E4_eye-tracking_data.csv") %>%
  rename("condition" = compatibility) %>%
  filter(time.window == "post-instrument-onset", condition != "filler" ) %>%
  group_by(condition, time.window, subject) %>%
  summarize(M = mean(prop.fixations.animal)) %>%
  pivot_wider(names_from = condition, values_from = M) %>%
  mutate(bias_effect = modifier - instrument) %>%
  left_join(calib.by.subj, by = "subject")
```

```
## Rows: 14667 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (4): subject, sound, compatibility, time.window
## dbl (3): trialID, prop.fixations.animal, prop.fixations.instrument
##
## 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.
## `summarise()` has grouped output by 'condition', 'time.window'. You can override using the `.groups`
```

```r
ggplot(eyetracking.effects.by.subj, aes(x = mean_percent_in_roi, y = modifier))+
  geom_point()+
  geom_smooth(method = "lm")
```

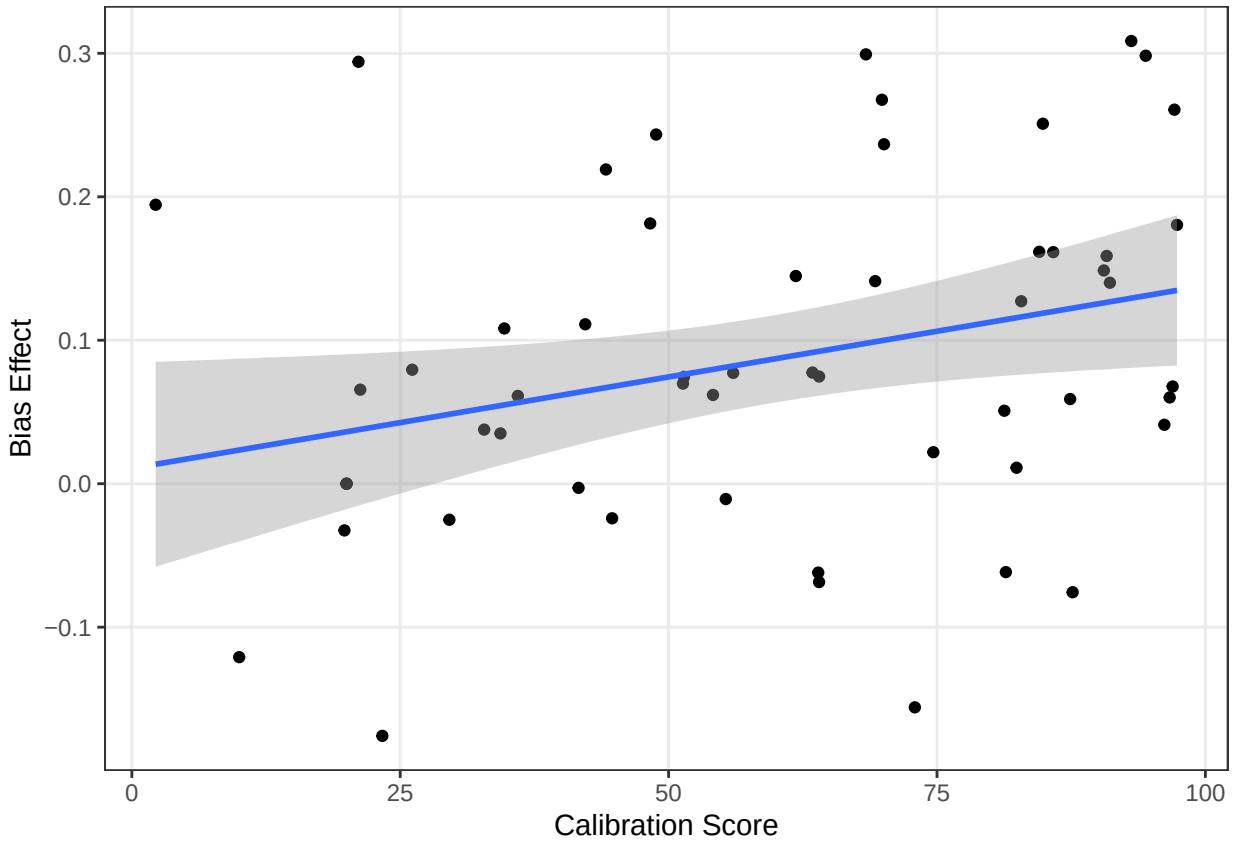
```r
write_csv(calib.by.subj,"output/calib_by_subj.csv")
```

```r
m_calib = broom::tidy(cor.test(eyetracking.effects.by.subj$mean_percent_in_roi, eyetracking.effects.by.s
#m_calib
```

```r
m_calib_r = m_calib %>% pull(estimate )
```

```r
ggplot(eyetracking.effects.by.subj, aes(x = mean_percent_in_roi, y = bias_effect))+
  geom_point()+
  geom_smooth(method = "lm") +
  labs(x="Calibration Score", y="Bias Effect") +
  theme_bw()+
  theme(panel.grid.minor = element_blank())
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



Participants' calibration quality, measured as the mean percentage of fixations that landed within 200 pixels of the calibration point, varied substantially (between 2.22, 97.36%). The quality of a participant's calibration significantly correlated with the participant's effect size (*Pearson's* $r = 0.2947347$, $p < 0.05$). The difference in target animal fixation proportions between modifier and instrument conditions was higher for participants with better calibration (see Figure @ref(fig:E4-calib-corr-plot)).

```
eyetracking.window.3 = read_csv( "output/E4_eye-tracking_data.csv") %>%
  filter(time.window == "post-instrument-onset", compatibility != "filler" ) %>%
  mutate(condition = factor(compatibility, levels = c('instrument', 'equibiased', 'modifier'))) %>%
  left_join(calib.by.subj, by = "subject")
```

```
## Rows: 14667 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (4): subject, sound, compatibility, time.window
## dbl (3): trialID, prop.fixations.animal, prop.fixations.instrument
##
## 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.
```

```
# Add orthogonal contrasts to model
contrasts(eyetracking.window.3$condition) <- cbind(c(-2/3, 1/3, 1/3), c(0, -1/2, 1/2))
```

```
model.time.window.3 <- lmer(prop.fixations.animal ~ condition + (1 +condition | subject) + (1 | trialID)
                            control = lmerControl(optimizer = "bobyqa",
                                                  optCtrl = list(maxfun = 2e6)))
summary(model.time.window.3)
```

```
eyetracking.window.3.good.calib = eyetracking.window.3 %>%
  filter(mean_percent_in_roi >= 50)
```

```
# Add orthogonal contrasts to model
contrasts(eyetracking.window.3.good.calib$condition) <- cbind(c(-2/3, 1/3, 1/3), c(0, -1/2, 1/2))
```

```
m.good.calib <- lmer(prop.fixations.animal ~ condition + (1  | subject) + (1 | trialID), data=eyetracki
                           control = lmerControl(optimizer = "bobyqa",
                                                 optCtrl = list(maxfun = 2e6)))
#summary(m.good.calib)

m_calib_good = broom.mixed::tidy(m.good.calib)

m_calib_good_c1 = m_calib_good %>% filter(term == "condition1")
m_calib_good_c2 = m_calib_good %>% filter(term == "condition2")
```

Re-analysis After Exclusions Replicating the linear mixed-effects analysis (in the post-instrument onset time window only) on a subset of 35 participants with calibration quality >50% suggests that the effect of verb bias condition was larger in this subset than in the full dataset. Participants looked more at the target animal in the modifier-biased condition and the equi-biased conditions relative to the instrument-biased condition ($b = 0.1$, $SE = 0.02$, $p < 0.001$) but not significantly so in the modifier biased condition relative to the equi-biased condition ($b = 0.02$, $SE = 0.02$, $p = 0.29$).

```
eyetracking.window.3.great.calib = eyetracking.window.3 %>%
  filter(mean_percent_in_roi >= 75)
```

```
# Add orthogonal contrasts to model
contrasts(eyetracking.window.3.great.calib$condition) <- cbind(c(-2/3, 1/3, 1/3), c(0, -1/2, 1/2))
```

```
m.great.calib <- lmer(prop.fixations.animal ~ condition + (1  | subject) + (1 | trialID), data=eyetrack
                           control = lmerControl(optimizer = "bobyqa",
                                                 optCtrl = list(maxfun = 2e6)))
#summary(m.great.calib)
m_calib_great = broom.mixed::tidy(m.great.calib)

m_calib_great_c1 = m_calib_great %>% filter(term == "condition1")
m_calib_great_c2 = m_calib_great %>% filter(term == "condition2")
```