

Experiment 2 Analysis

1
2
3

4 Experiment 2: Honorific Greetings

```
g_fin <- read.csv(here('data/greetings_reviewer2_data.csv'))
g_fin$subject_id <- as.factor(g_fin$subject_id)

g_pp_tab <- g_fin %>%
  distinct(subject_id, bebe_meses)%>%
  summarize(min=min(bebe_meses),
            max=max(bebe_meses),
            M=mean(bebe_meses),
            SD=sd(bebe_meses))

G_N <- nrow(g_fin %>%
  distinct(subject_id, bebe_meses))
G_MIN_AGE <- op(g_pp_tab$min)
G_MAX_AGE <- op(g_pp_tab$max)
G_MEAN_AGE <- op(g_pp_tab$M)
G_SD_AGE <- op(g_pp_tab$SD)
```

```
g_mama_tab <- demo %>%
  filter(subject_id %in% g_fin$subject_id,
         !is.na(mama_edad)) %>%
  summarize(min=min(mama_edad),
            max=max(mama_edad),
            M=mean(mama_edad),
            SD=sd(mama_edad))

G_MOT_MIN_AGE <- op(g_mama_tab$min)
G_MOT_MAX_AGE <- op(g_mama_tab$max)
G_MOT_M_AGE <- op(g_mama_tab$M)
G_MOT_SD_AGE <- op(g_mama_tab$SD)
```

5 **Participants.** We report on data from 18 infants (*range*= 5.33 – 14.93*mos*, M_{age} =
6 10.77*mos*, $SD_{\text{age}} = 2.67\text{mos}$) tested in the Canada Chica *paraje* of a Tseltal Maya commu-
7 nity in Tenejapa, Chiapas, in Southern Mexico.

8 Mothers ranged in age from 14.00 to 42.00*yrs* ($M_{\text{age}} = 26.81\text{yrs}$, $SD_{\text{age}} = 7.80\text{yrs}$)

```

g_fin$stimulus_set <- ''

g_fin[g_fin$old_trial_name %in% c('old_woman-YOUNG_MAN',
                                  'OLD_WOMAN-young_man'), ]$stimulus_set <- 'A'

g_fin[g_fin$old_trial_name %in% c('young_man-OLD_WOMAN',
                                  'YOUNG_MAN-old_woman'), ]$stimulus_set <- 'B'

g_fin[g_fin$old_trial_name %in% c('old_man-YOUNG_WOMAN',
                                  'OLD_MAN-young_woman'), ]$stimulus_set <- 'C'

g_fin[g_fin$old_trial_name %in% c('young_woman-OLD_MAN',
                                  'YOUNG_WOMAN-old_man'), ]$stimulus_set <- 'D'

g_fin$target_stimulus <- paste(g_fin$target_noun, g_fin$stimulus_set, sep='')
g_fin$non_target_stimulus <- ''

g_fin[g_fin$target_stimulus=='young_manA',]$non_target_stimulus <- 'old_womanA'
g_fin[g_fin$target_stimulus=='old_womanA',]$non_target_stimulus <- 'young_manA'

g_fin[g_fin$target_stimulus=='young_manB',]$non_target_stimulus <- 'old_womanB'
g_fin[g_fin$target_stimulus=='old_womanB',]$non_target_stimulus <- 'young_manB'

g_fin[g_fin$target_stimulus=='young_womanC',]$non_target_stimulus <- 'old_manC'
g_fin[g_fin$target_stimulus=='old_manC',]$non_target_stimulus <- 'young_womanC'

g_fin[g_fin$target_stimulus=='young_womanD',]$non_target_stimulus <- 'old_manD'
g_fin[g_fin$target_stimulus=='old_manD',]$non_target_stimulus <- 'young_womanD'

g_target_df <- g_fin %>%
  filter(noun_pair %in% c('old_man-young_woman', 'old_woman-young_man'),
         subject_id %in% g_subjects) %>%
  dplyr::select(subject_id,
                bebe_meses,
                age_centered,
                age_group,
                stimulus_set,
                target_stimulus,
                noun_pair,
                post1_target_prop) %>%
  mutate(merge_on_noun = target_stimulus)

g_non_target_df <- g_fin %>%

```

```

filter(noun_pair %in% c('old_man-young_woman', 'old_woman-young_man'),
       subject_id %in% g_subjects) %>%
dplyr::select(subject_id,
              bebe_meses,
              age_centered,
              age_group,
              noun_pair,
              stimulus_set,
              non_target_stimulus,
              post1_nontarget_prop) %>%
mutate(merge_on_noun=non_target_stimulus)

g_target_nontarget_props_df <- merge(g_target_df, g_non_target_df,
                                     by=c('subject_id',
                                           'noun_pair',
                                           'stimulus_set',
                                           'merge_on_noun',
                                           'age_centered',
                                           'bebe_meses',
                                           'age_group'), all=T) %>%

group_by(subject_id, bebe_meses, age_centered, age_group,
          noun_pair, stimulus_set, merge_on_noun) %>%
summarize(post1_target_prop = na.mean(post1_target_prop),
          post1_nontarget_prop = na.mean(post1_nontarget_prop))

g_diffs_df <- g_target_nontarget_props_df %>%
dplyr::select(subject_id, bebe_meses, age_centered, age_group,
              noun_pair, stimulus_set, merge_on_noun,
              post1_target_prop,
              post1_nontarget_prop) %>%
mutate(diff = post1_target_prop - post1_nontarget_prop) %>%
group_by(subject_id,
          bebe_meses, age_centered, age_group,
          noun_pair, stimulus_set) %>%
summarize(stim_pair_diff = mean(diff)) %>%
group_by(subject_id, bebe_meses, age_centered, age_group, noun_pair) %>%
summarize(noun_pair_diff = na.mean(stim_pair_diff)) %>%
filter(!is.na(noun_pair))

g_diffs_df$age_std <- g_diffs_df$age_centered/as.numeric(G_SD_AGE)

write.csv(g_diffs_df,
          here('data/r_analysis_dfs', 'g_bs2012_diffs.csv')
          )

```

9 Mean Difference Score Analysis.

```

set.seed(36)

g_by_item_tab <- g_diffs_df %>%
  group_by(noun_pair, subject_id, age_centered) %>%
  summarize(subj_item_mean = na.mean(noun_pair_diff)) %>%
  filter(!is.na(subj_item_mean)) %>%
  ungroup() %>%
  group_by(noun_pair) %>%
  summarize(n = n(),
            M = na.mean(subj_item_mean),
            cil=M-ci.low(subj_item_mean),
            cih = M+ci.high(subj_item_mean),
            min=min(subj_item_mean, na.rm=T),
            max=max(subj_item_mean, na.rm=T))

G_ALL_ITEM_MEAN <- mean(g_by_item_tab$M)
G_ALL_ITEM_CIL <- G_ALL_ITEM_MEAN - ci.low(g_by_item_tab$M)
G_ALL_ITEM_CIH <- G_ALL_ITEM_MEAN + ci.high(g_by_item_tab$M)
G_ALL_ITEM_MIN <- min(g_by_item_tab$M)
G_ALL_ITEM_MAX <- max(g_by_item_tab$M)
N_G_POSITIVE_ITEMS <- sum(g_by_item_tab$M > 0)

```

```

set.seed(36)

write(apa_table(g_by_item_tab,
               caption='Experiment 2 Item Means'),
      here('supplement/tables/exp_2/g_by_item_tab.tex'))

xtable2kable(g_by_item_tab)

```

10 *By Item.*

```

11 ## # A tibble: 2 x 7
12 ##   noun_pair      n      M      cil      cih      min      max
13 ##   <chr>      <int> <dbl> <dbl> <dbl> <dbl> <dbl>
14 ## 1 old_man-young_woman    13 0.0848 -0.138 0.285 -0.713 0.538
15 ## 2 old_woman-young_man    15 0.182  0.0648 0.305 -0.191 0.540

```

```
# Bayes factor
ttestBF(g_by_item_tab$M)
```

16 *Non-parametric.*

```
17 ## Bayes factor analysis
18 ## -----
19 ## [1] Alt., r=0.707 : 1.2 ±0%
20 ##
21 ## Against denominator:
22 ##    Null, mu = 0
23 ## ---
24 ## Bayes factor type: BFoneSample, JZS
```

```
# Standard deviation
stdev=sd(g_by_item_tab$M, na.rm=T)
# Mean
mean_data=na.mean(g_by_item_tab$M)
# Effect size
G_ITEM_D=abs(mean_data/stdev)
```

```
g_by_item_sub_tab <- g_diffs_df %>%
  group_by(noun_pair, subject_id) %>%
  summarize(subj_item_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  filter(!is.na(subj_item_mean))

g_item_wilcox <- wilcox.test(g_by_item_sub_tab$subj_item_mean,
                           mu=0,
                           alternative="two.sided")

G_ITEM_P_WILCOX <- reportP(as.numeric(g_item_wilcox['p.value']))
```

```
g_item_binom <- binom.test(2, 2, p=.5)

G_ITEM_P_BINOM <- reportP(g_item_binom$p.value)
```

```
set.seed(36)

g_diffs_df$noun_pair_label <- ''
g_diffs_df[
  g_diffs_df$noun_pair=='old_woman-young_man',]$noun_pair_label <-
```

```

'old woman -\nyoung man'
g_diffs_df[
  g_diffs_df$noun_pair=='old_man-young_woman',]$noun_pair_label <-
  'old man -\nyoung woman'

g_plot2 <- g_diffs_df %>%
  group_by(noun_pair, noun_pair_label, subject_id) %>%
  summarize(subj_item_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  filter(!is.na(subj_item_mean)) %>%
  group_by(noun_pair, noun_pair_label) %>%
  mutate(n = n(),
         item_mean = na.mean(subj_item_mean),
         min=min(subj_item_mean, na.rm=T),
         max=max(subj_item_mean, na.rm=T),
         ci.low=item_mean-ci.low(subj_item_mean),
         ci.high = item_mean+ci.high(subj_item_mean)) %>%
  ggplot(., aes(x=noun_pair_label, y=item_mean, color=noun_pair)) +
    geom_hline(yintercept=0) +
    geom_jitter(aes(y=subj_item_mean), alpha=.5) +
    geom_pointrange(aes(y=item_mean, ymin=ci.low, ymax=ci.high,
                       fill=noun_pair), shape=23, size=.5) +
  sb.density.theme +
  theme(axis.text.x = element_text(angle=60, vjust=1, hjust=1, size=10),
        legend.position = "none",
        axis.title = element_text(colour="gray30", size=11, face="bold")) +
  ylab("Increase in Target Looking") +
  xlab("Honorific-Pair") +
  ylim(-1, 1) +
  scale_discrete_manual(aesthetics=c("color", "fill"),
                       values=greeting_pair_fills)

```

```
set.seed(36)

g_sub_df <- g_diffs_df %>%
  filter(!is.na(noun_pair_diff)) %>%
  group_by(subject_id, bebe_meses, age_centered) %>%
  summarize(g_M = na.mean(noun_pair_diff),
            n=n())

G_N_FOR_MEAN_DIFF_ANALYSES <- nrow(g_sub_df)

g_positive_scorers_age_tab <- g_sub_df %>%
  filter(g_M>0) %>%
  ungroup() %>%
  summarize(mean=mean(bebe_meses),
            sd=sd(bebe_meses),
            ci.low=mean-ci.low(bebe_meses),
            ci.high=mean+ci.high(bebe_meses))

G_N_POSITIVE_SUBS <- sum(g_sub_df$g_M>0)
G_N_TOTAL_SUBS <- nrow(g_sub_df)
G_PS_MEAN_AGE <- g_positive_scorers_age_tab$mean
G_PS_CILO_AGE <- g_positive_scorers_age_tab$ci.low
G_PS_CIHI_AGE <- g_positive_scorers_age_tab$ci.high
G_PS_SD_AGE <- g_positive_scorers_age_tab$sd
```



```
set.seed(36)

g_by_sub_overall_tab <- g_diffs_df %>%
  group_by(subject_id) %>%
  summarize(subj_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  filter(!is.na(subj_mean)) %>%
  summarize(group_subject_mean = mean(subj_mean),
            min=min(subj_mean),
            max=max(subj_mean),
            ci.low=group_subject_mean-ci.low(subj_mean),
            ci.high = group_subject_mean+ci.high(subj_mean))

G_ALL_SUB_MEAN <- g_by_sub_overall_tab$group_subject_mean
G_ALL_SUB_CIL <- g_by_sub_overall_tab$ci.low
G_ALL_SUB_CIH <- g_by_sub_overall_tab$ci.high
G_ALL_SUB_MIN <- g_by_sub_overall_tab$min
G_ALL_SUB_MAX <- g_by_sub_overall_tab$max
```

```

g_diffs_df %>%
  filter(!is.na(noun_pair_diff)) %>%
  group_by(subject_id, bebe_meses) %>%
  summarize(subj_item_mean = na.mean(noun_pair_diff),
             n_items = n()) %>%
  mutate(all_pairs = n_items==2) %>%
  filter(!is.na(subj_item_mean)) %>%
  ggplot(.) +
  geom_point(aes(x=bebe_meses, y=subj_item_mean),
             color=rf_color, fill=rf_fill,
             shape=24, size=2.5) +
  geom_linerange(aes(x=bebe_meses, ymin=0, ymax=subj_item_mean),
                color=rf_color) +
  geom_hline(yintercept=0) +
  sb.density.theme +
  theme(
    axis.title = element_text(colour="gray30", size=11, face="bold"),
    axis.text = element_text(colour="gray30", size=11),
    axis.ticks = element_line(colour="gray30")) +
  ylab("Increase in Target Looking") +
  xlab("Child Age (months)") +
  xlim(4.5, 16.0) +
  scale_x_continuous(breaks=c(4,5,6,7,8,9,10,11,12,13,14,15)) +
  ylim(-.4, .6)

```

```

ggsave(here('supplement/plots/exp_2/pdfs', 'g_meandiffs_bysub.pdf'),
       device='pdf', width=2.5, height=1.5, units='in', scale=2)
ggsave(here('supplement/plots/exp_2/pngs', 'g_meandiffs_bysub.png'),
       device='png', width=2.5, height=1.5, units='in', scale=2)

```

25 *By Subject.*

```

# Bayes factor
ttestBF(g_sub_df$g_M)

```

26 *Non-parametric.*

27 ## Bayes factor analysis

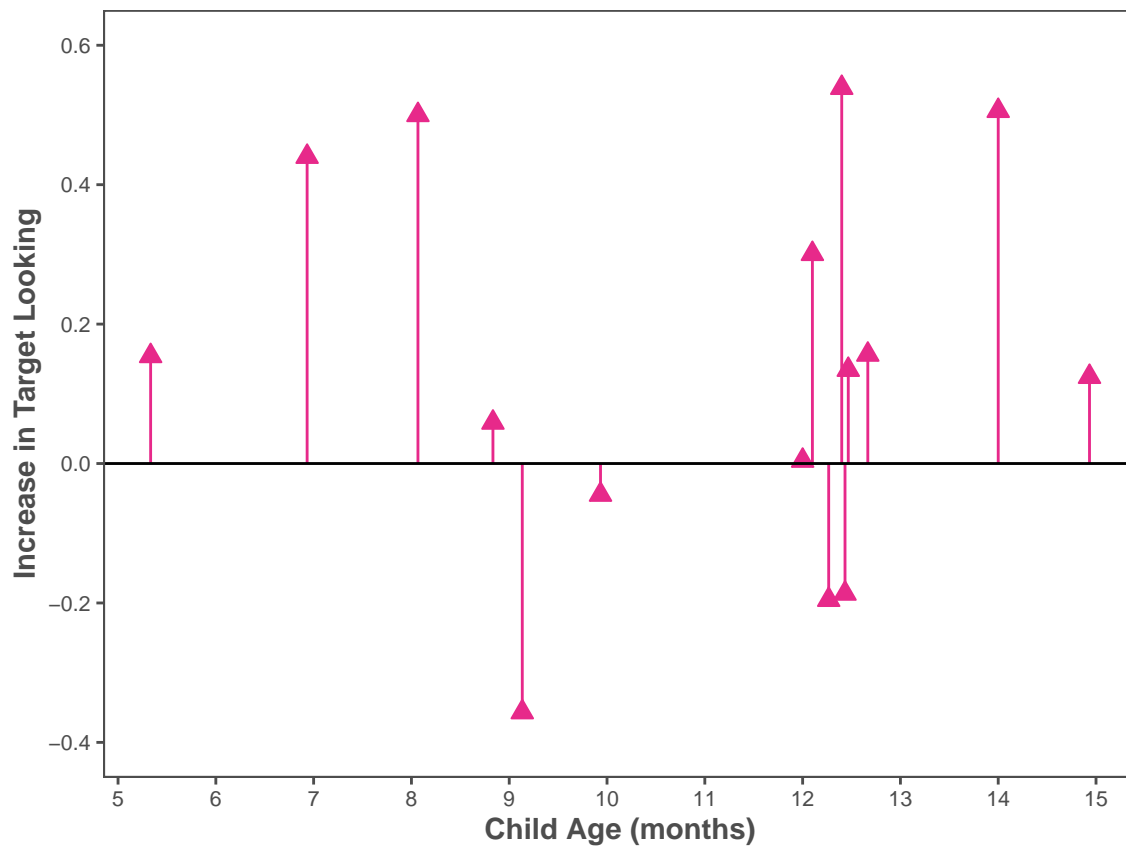


Figure 1

```

28 ## -----
29 ## [1] Alt., r=0.707 : 1.3 ±0.02%
30 ##
31 ## Against denominator:
32 ##   Null, mu = 0
33 ## ---
34 ## Bayes factor type: BFoneSample, JZS

```

```

# Standard deviation
stdev=sd(g_sub_df$g_M)
# Mean
mean_data=mean(g_sub_df$g_M)
# Effect size
G_SUBJ_D=abs(mean_data/stdev)

```

```

g_subj_wilcox <- wilcox.test(g_sub_df$g_M, mu=0,
                           alternative="two.sided")

```

```

G_SUBJ_P_WILCOX <- reportP(g_subj_wilcox$p.value)

```

```
g_subj_binom <- binom.test(G_N_POSITIVE_SUBS, G_N_FOR_MEAN_DIFF_ANALYSES, p=.5)
G_SUBJ_P_BINOM <- reportP(g_subj_binom$p.value)
```

```
set.seed(36)

g_model0 <- lmer(noun_pair_diff ~ 0 + (1|subject_id), REML = FALSE,
  g_diffs_df)
g_model1 <- lmer(noun_pair_diff ~ 1 + (1|subject_id), REML = FALSE,
  g_diffs_df)

g_b0_anova <- anova(g_model1, g_model0)
G_B0_CHISQ <- g_b0_anova$Chisq[2]
G_B0_P <- reportP(g_b0_anova$`Pr(>Chisq)`[2])

g_all_subs_intercept <- as.data.frame(
  cbind(b=fixef(g_model1),
    ci.low=confint(g_model1)[3,1],
    ci.high=confint(g_model1)[3,2])
)

G_B0_EST <- op(g_all_subs_intercept$b)
G_B0_CIL <- op(g_all_subs_intercept$ci.low)
G_B0_CIH <- op(g_all_subs_intercept$ci.high)

G_B0_TT_DF <- op(as.numeric(unlist(summary(g_model1)['coefficients'])[3]))
G_B0_TT_STAT <- op(as.numeric(
  unlist(summary(g_model1)['coefficients'])[4])
)
G_B0_TT_P <- reportP(as.numeric(
  unlist(summary(g_model1)['coefficients'])[5])
)

#noun pair not enough variance to warrant random intercept
write(texreg(g_model1),
  file=here('supplement/tables/exp_2/g_lmer0_tab.tex'))
```

35 **MLM Intercept.** 2/2 honorific-pairs showed positive mean difference scores. The
 36 mean across items was positive (*range*: 0.08–0.18, $M = 0.13$ [0.08, 0.18]; $p < .05$, Wilcoxon
 37 test; $p = 0.500$, binomial test; $d = 1.95$).

38 11/15 subjects showed a positive mean difference score ($M_{\text{age}} = 10.88$ [9.14, 12.58],

39 $SD_{age} = 3.09$). The mean across subjects was positive (*range*: $-0.36 - 0.54$, $M = 0.14$
 40 $[0.00, 0.28]$; $p = 0.095$, Wilcoxon test, $p = 0.118$, binomial test, $d = 0.52$).

41 A linear mixed effects model suggested that infants' positive mean performance was
 42 reliable, accounting for subject- and item-level variability ($\beta_0 = 0.14$, 95% *CI*: $[-0.01, 0.29]$,
 43 $t(15.23) = 1.98$, $p = 0.066$, $\chi^2(1) = 3.53$, $p = 0.060$).

```
set.seed(36)

g_by_subj_age_tab <- g_diffs_df %>%
  filter(!is.na(noun_pair_diff)) %>%
  group_by(subject_id, bebe_meses, age_centered) %>%
  summarize(subj_item_mean = na.mean(noun_pair_diff),
            n_subjects = n(),
            min = min(noun_pair_diff, na.rm=T),
            max = max(noun_pair_diff, na.rm=T),
            ci.low = subj_item_mean-ci.low(noun_pair_diff),
            ci.high = subj_item_mean+ci.high(noun_pair_diff))

g_age_cor_test <- cor.test(g_by_subj_age_tab$bebe_meses,
                          g_by_subj_age_tab$subj_item_mean,
                          method="kendall")

G_AGE_CORR <- as.numeric(g_age_cor_test$estimate)
G_AGE_CORR_P <- g_age_cor_test$p.value
G_AGE_CORR_T <- g_age_cor_test$statistic
```

44 **Correlation with Age.** Mean scores were not significantly correlated with infants'
 45 ages in months ($\tau = 0.05$, $p = 0.85$).

```
g_age_lm <- lm(subj_item_mean ~ 1 + age_centered, g_by_subj_age_tab)
summary(g_age_lm)
```

```
46 ##
47 ## Call:
48 ## lm(formula = subj_item_mean ~ 1 + age_centered, data = g_by_subj_age_tab)
49 ##
50 ## Residuals:
51 ##      Min       1Q   Median       3Q      Max
52 ## -0.5040 -0.1624 -0.0037  0.2242  0.4009
53 ##
54 ## Coefficients:
55 ##              Estimate Std. Error t value Pr(>|t|)
```

```
56 ## (Intercept) 0.14180 0.07403 1.92 0.078 .
57 ## age_centered -0.00281 0.02809 -0.10 0.922
58 ## ---
59 ## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
60 ##
61 ## Residual standard error: 0.29 on 13 degrees of freedom
62 ## Multiple R-squared: 0.00077, Adjusted R-squared: -0.0761
63 ## F-statistic: 0.01 on 1 and 13 DF, p-value: 0.922
```

```
write(texreg(g_age_lm),
       file=here('supplement/tables/exp_2/g_age_lm_tab.tex'))
```

```

g_supp_pre <- g_fin %>%
  select(pre_target_sum_ms, pre_nontarget_sum_ms,
         subject_id, bebe_meses, age_centered, noun_pair, path)

g_supp_pre$phase <- "pre-naming"
g_supp_pre$target_bins <- round(g_supp_pre$pre_target_sum_ms/20,0)
g_supp_pre$nontarget_bins <- round(g_supp_pre$pre_nontarget_sum_ms/20, 0)

g_supp_pre <- g_supp_pre %>%
  select(subject_id, bebe_meses, age_centered, noun_pair, phase,
         target_bins, nontarget_bins, path)

g_supp_post <- g_fin %>%
  select(post1_target_sum_ms, post1_nontarget_sum_ms,
         subject_id, bebe_meses, age_centered, noun_pair, path)

g_supp_post$phase <- "post-naming"
g_supp_post$target_bins <- round(g_supp_post$post1_target_sum_ms/20, 0)
g_supp_post$nontarget_bins <- round(g_supp_post$post1_nontarget_sum_ms/20, 0)

g_supp_post <- g_supp_post %>%
  select(subject_id, bebe_meses, age_centered, noun_pair, phase,
         target_bins, nontarget_bins, path)

g_supp_stacked_r2 <- rbind(g_supp_pre, g_supp_post)

g_supp_stacked_r2$phase <- as.factor(g_supp_stacked_r2$phase)
g_supp_stacked_r2$phase <- relevel(g_supp_stacked_r2$phase, ref="pre-naming")

g_supp_stacked_r2 %>%
  group_by(subject_id, bebe_meses, age_centered, noun_pair, phase, path) %>%
  summarize(target_bins = target_bins,
            nontarget_bins = nontarget_bins) %>%
  write.csv(., here('data/r_analysis_dfs',
                  'g_prepost_target_looking.csv'))
)

```

64

Pre/Post Looking Logit Model. *Get back to 126 trials*

```

set.seed(36)

g_supp_glmer_null <- glmer(cbind(target_bins, nontarget_bins) ~
                          1 + (1|subject_id),

```

```

                                family=binomial, g_supp_stacked_r2,
                                control=glmerControl(optimizer="bobyqa",
                                optCtrl=list(maxfun=100000)))

g_supp_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
                      phase + (1|subject_id),
                      family=binomial, g_supp_stacked_r2,
                      control=glmerControl(optimizer="bobyqa",
                      optCtrl=list(maxfun=100000)))

g_supp_glmer_tab <- as.data.frame(
  cbind("OR"=op(exp(fixef(g_supp_glmer))),
        "CIL"=op(exp(confint.merMod(g_supp_glmer))[2:3,1]),
        "CIH"=op(exp(confint.merMod(g_supp_glmer))[2:3,2]))

G_POSTNAMING_OR <- g_supp_glmer_tab$OR[2]
G_POSTNAMING_CIL <- g_supp_glmer_tab$CIL[2]
G_POSTNAMING_CIH <- g_supp_glmer_tab$CIH[2]
G_PHASE_WALD_CHISQ <- Anova(g_supp_glmer)['phase', 'Chisq']
G_PHASE_WALD_P <- reportP(
  Anova(g_supp_glmer)['phase', 'Pr(>Chisq)'])

g_supp_age_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
                          phase + age_centered +
                          (1|subject_id),
                          family=binomial, g_supp_stacked_r2,
                          control=glmerControl(optimizer="bobyqa",
                          optCtrl=list(maxfun=100000)))

summary(g_supp_glmer)

```

```

65 ## Generalized linear mixed model fit by maximum likelihood (Laplace
66 ##   Approximation) [glmerMod]
67 ##   Family: binomial ( logit )
68 ## Formula: cbind(target_bins, nontarget_bins) ~ phase + (1 | subject_id)
69 ##   Data: g_supp_stacked_r2
70 ## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
71 ##
72 ##           AIC           BIC    logLik deviance df.resid
73 ##      19929      19939     -9961    19923      227
74 ##
75 ## Scaled residuals:
76 ##      Min       1Q   Median       3Q      Max
77 ## -23.27  -5.88   1.18   6.75  16.68

```



```

78 ##
79 ## Random effects:
80 ##   Groups      Name      Variance Std.Dev.
81 ##   subject_id (Intercept) 0.223    0.472
82 ## Number of obs: 230, groups:  subject_id, 18
83 ##
84 ## Fixed effects:
85 ##               Estimate Std. Error z value Pr(>|z|)
86 ## (Intercept)      0.1191    0.1126    1.06    0.29
87 ## phasepost-naming  0.2619    0.0235   11.14   <2e-16 ***
88 ## ---
89 ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
90 ##
91 ## Correlation of Fixed Effects:
92 ##               (Intr)
93 ## phspst-nmng -0.107

```

```
exp(fixef(g_supp_glmer))
```

```

94 ##      (Intercept) phasepost-naming
95 ##              1.1              1.3

```

```
exp(confint(g_supp_glmer))
```

```

96 ##              2.5 % 97.5 %
97 ## .sig01          1.42    2.0
98 ## (Intercept)      0.89    1.4
99 ## phasepost-naming 1.24    1.4

```

```
Anova(g_supp_glmer)
```

```

100 ## Analysis of Deviance Table (Type II Wald chisquare tests)
101 ##
102 ## Response: cbind(target_bins, nontarget_bins)
103 ##      Chisq Df Pr(>Chisq)
104 ## phase   124  1   <2e-16 ***
105 ## ---
106 ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

#G_POSTNAMING_OR <- op(g_pp_age_glmer_tab$OR[2])
#G_POSTNAMING_CIL <- op(g_pp_age_glmer_tab$CIL[2])
#G_POSTNAMING_CIH <- op(g_pp_age_glmer_tab$CIH[2])

```

```

#G_PHASE_WALD_CHISQ <- op(Anova(g_supp_age_glmer)['phase', 'Chisq'])
#G_PHASE_WALD_P <- op(Anova(g_supp_age_glmer)['phase', 'Pr(>Chisq)'])

g_supp_ageint_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
                             phase*age_centered +
                             (1|subject_id),
                             family=binomial, g_supp_stacked_r2)

confint.merMod(object = g_supp_ageint_glmer, method = "boot")

```

```

107 ##                                2.5 % 97.5 %
108 ## .sig01                        0.253  0.545
109 ## (Intercept)                  -0.127  0.270
110 ## phasepost-naming              0.227  0.320
111 ## age_centered                  -0.180 -0.025
112 ## phasepost-naming:age_centered  0.018  0.051

```

```

g_supp_ageint_glmer_tab <- as.data.frame(
  cbind("OR"=op(exp(fixef(g_supp_ageint_glmer))),
        "CIL"=op(exp(confint.merMod(object = g_supp_ageint_glmer,
                                     method = "boot"))[2:5,1]),
        "CIH"=op(exp(confint.merMod(object = g_supp_ageint_glmer,
                                     method = "boot"))[2:5,2]))))

g_supp_ageint_anova <- anova(g_supp_glmer, g_supp_age_glmer,
                             g_supp_ageint_glmer)

G_POST_AGE_INT_PHASE_OR <- g_supp_ageint_glmer_tab$OR[2]
G_POST_AGE_INT_PHASE_CIL <- g_supp_ageint_glmer_tab$CIL[2]
G_POST_AGE_INT_PHASE_CIH <- g_supp_ageint_glmer_tab$CIH[2]
G_POST_AGE_INT_PHASE_WALD_CHISQ <-
  Anova(g_supp_ageint_glmer)['phase', 'Chisq']
G_POST_AGE_INT_PHASE_WALD_P <-
  reportP(Anova(g_supp_ageint_glmer)['phase', 'Pr(>Chisq)'])

G_POST_AGE_INT_OR <- g_supp_ageint_glmer_tab$OR[4]
G_POST_AGE_INT_CIL <- g_supp_ageint_glmer_tab$CIL[4]
G_POST_AGE_INT_CIH <- g_supp_ageint_glmer_tab$CIH[4]
G_POST_AGE_INT_WALD_CHISQ <-
  op(Anova(g_supp_ageint_glmer)['phase:age_centered', 'Chisq'])
G_POST_AGE_INT_WALD_P <-
  reportP(Anova(g_supp_ageint_glmer)['phase:age_centered', 'Pr(>Chisq)'])

G_POST_AGE_INT_DF <- g_supp_ageint_anova$Df[2]

```

```
G_POST_AGE_INT_CHISQ <- g_supp_ageint_anova$Chisq[2]
G_POST_AGE_INT_P <- reportP(g_supp_ageint_anova$`Pr(>Chisq)`[2])

Anova(g_supp_ageint_glmer)
```

```
113 ## Analysis of Deviance Table (Type II Wald chisquare tests)
114 ##
115 ## Response: cbind(target_bins, nontarget_bins)
116 ##               Chisq Df Pr(>Chisq)
117 ## phase           123.84  1    <2e-16 ***
118 ## age_centered      5.08  1    0.0242 *
119 ## phase:age_centered 13.82  1    0.0002 ***
120 ## ---
121 ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
write(texreg(list(g_supp_glmer, g_supp_age_glmer)),
      here('supplement/tables/exp_2', 'g_prepost_glmers_w_wo_age_tab.tex')
)

write(texreg(list(g_supp_glmer, g_supp_age_glmer, g_supp_ageint_glmer)),
      here('supplement/tables/exp_2', 'g_prepost_3glmers_tab.tex')
)
```

122 Results Text

123 The odds ratio for trial phase (POST-NAMING OR= 1.30, 95% CI: [1.24, 1.36]) suggests
 124 that infants' relative looking time to the competitor faces was responsive to the honorific
 125 term used by their caregivers: infants dedicated a significantly greater proportion of their
 126 looking time to the target face after hearing the honorific ($Wald\chi^2(1) = 123.99$, $p < .001$,
 127 *Cohen d* = 0.65).

128 As in Experiment 1, a model which additionally included infant age and its interaction
 129 with trial phase resulted in a significantly better fit ($\chi^2(1) = 4.49$, $p < .05$), showing
 130 a reliable effect of trial phase ($OR = 1.31$, 95% CI: [1.25, 1.37], $Wald\chi^2(1) = 123.84$,
 131 $p < .001$, *Cohen d* = 0.65) and interaction with age, such that older children showed
 132 a greater increase in the ratio of target:non-target looking after hearing the target word
 133 ($OR = 1.03$, 95% CI: [1.02, 1.05], $Wald\chi^2(1) = 13.82$, $p < .001$, *Cohen d* = 0.081).

134 **GLMERs by Age Group.** From B&S2012:

135 >A separate hierarchical logistic regression model was created for each group of children
 136 (6–9 mo, 10–13 mo, 14–16 mo, and 18–20 mo) for each trial type (paired-picture and scene).
 137 Phase of trial (pretarget utterance vs. posttarget utterance) was included as a fixed-effect
 138 predictor, and subject and item were included as random effects. Each model predicts (the
 139 log of) the ratio of target to distracter looking, as calculated by counting time bins.

140

Trial Counts Across Analyses

```

G_TOTAL_TRIALS_N <- nrow(g_fin)

G_MEAN_DIFF_TRIALS_N <- g_target_nontarget_props_df %>%
  dplyr::select(subject_id, age_centered,
                noun_pair, stimulus_set, merge_on_noun,
                post1_target_prop,
                post1_nontarget_prop) %>%
  mutate(diff = post1_target_prop - post1_nontarget_prop) %>%
  group_by(
    subject_id, age_centered, noun_pair, stimulus_set
  ) %>%
  summarize(stim_pair_diff = mean(diff)) %>%
  filter(!is.na(stim_pair_diff)) %>%
  nrow(.) * 2

G_DROPPED_TRIALS <- G_TOTAL_TRIALS_N - G_MEAN_DIFF_TRIALS_N
G_DROPPED_TRIAL_PERCENT <- G_DROPPED_TRIALS/G_TOTAL_TRIALS_N

```

141

25 trials dropped for paired difference score analysis, or 21.74% of non-excluded trials.

```

g_supp_stacked_r2 %>%
  group_by(subject_id) %>%
  summarize(trials=n()/2) %>%
  ungroup() %>%
  summarize(min=min(trials),
            max=max(trials),
            mean=mean(trials),
            med=median(trials),
            mode=DescTools::Mode(trials))

```

142

```
## # A tibble: 1 x 5
```

143

```
##      min    max  mean   med  mode
```

144

```
##   <dbl> <dbl> <dbl> <dbl> <dbl>
```

145

```
## 1      3      8  6.39   6.5    8
```

146

Trial Durations

```

G_MED_TRIAL_DUR_S <- median(g_fin$trialtolookingoffset_dur_s, na.rm=T)
G_MIN_TRIAL_DUR_S <- min(g_fin[g_fin$trial_dur_s<30,]$trialtolookingoffset_dur_s)
G_MAX_TRIAL_DUR_S <- max(g_fin[g_fin$trial_dur_s<30,]$trialtolookingoffset_dur_s)

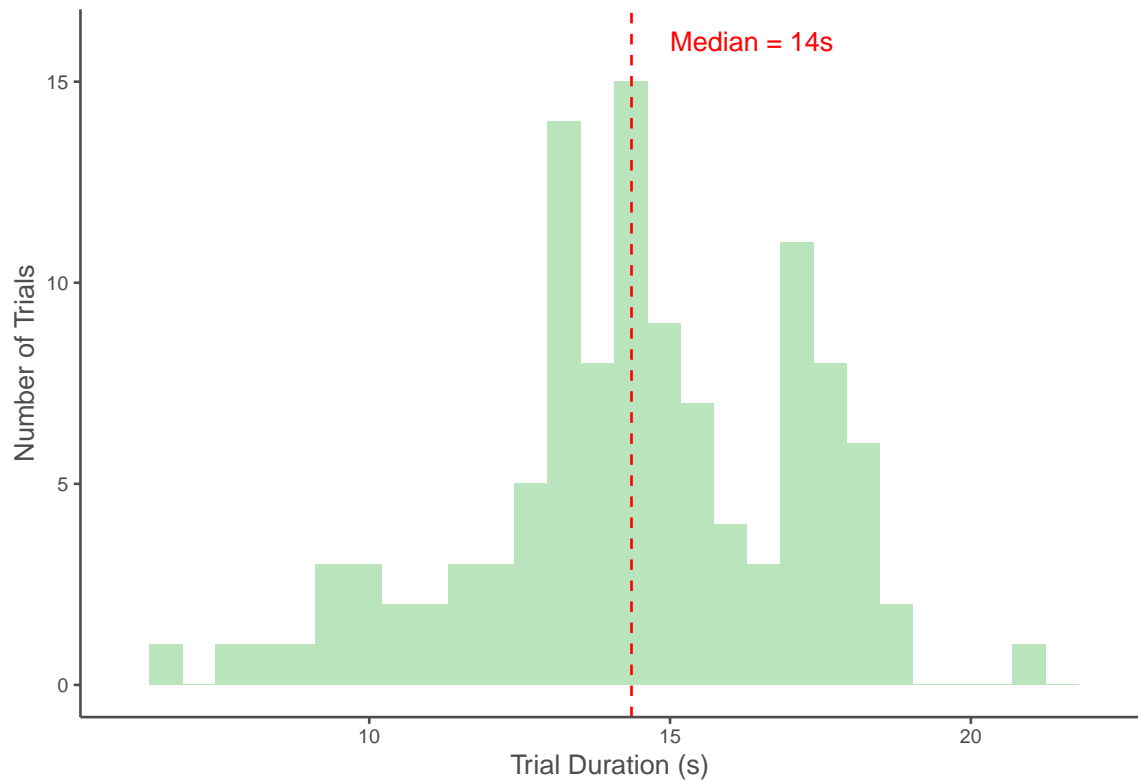
```

```

G_MEAN_TRIAL_DUR_S <- mean(
  g_fin[g_fin$trial_dur_s<30,]$trialextolookingoffset_dur_s)
G_CIL_TRIAL_DUR_S <- G_MEAN_TRIAL_DUR_S - ci.low(
  g_fin[g_fin$trial_dur_s<30,]$trialextolookingoffset_dur_s)
G_CIH_TRIAL_DUR_S <- G_MEAN_TRIAL_DUR_S + ci.high(
  g_fin[g_fin$trial_dur_s<30,]$trialextolookingoffset_dur_s)

ggplot(g_fin, aes(x=trialextolookingoffset_dur_s)) +
  geom_histogram(fill="#bae4bc") +
  sb.density.theme +
  geom_vline(xintercept=G_MED_TRIAL_DUR_S, color="red", lty=2) +
  xlim(6,22) +
  xlab("Trial Duration (s)") +
  ylab("Number of Trials") +
  theme(axis.title = element_text(colour="gray30", size=11),
        axis.text = element_text(colour="gray30", size=11),
        axis.ticks = element_line(colour="gray30"),
        plot.background = element_blank() ,
        panel.grid.major = element_blank() ,
        panel.grid.minor = element_blank() ,
        panel.border = element_blank() ,
        panel.background = element_blank(),
        axis.line = element_line(color = "gray30")) +
  annotate(
    "text", label = "Median = 14s",
    x = G_MED_TRIAL_DUR_S+2, y = 16, size = 4, colour = "red")

```



147

```
ggsave(here('supplement/plots/exp_2/pdfs', 'g_trialdur_histogram.pdf'),
       device='pdf', width=2.5, height=1.25, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_2/pngs', 'g_trialdur_histogram.png'),
       device='png', width=2.5, height=1.25, units='in', scale=2.5)
```

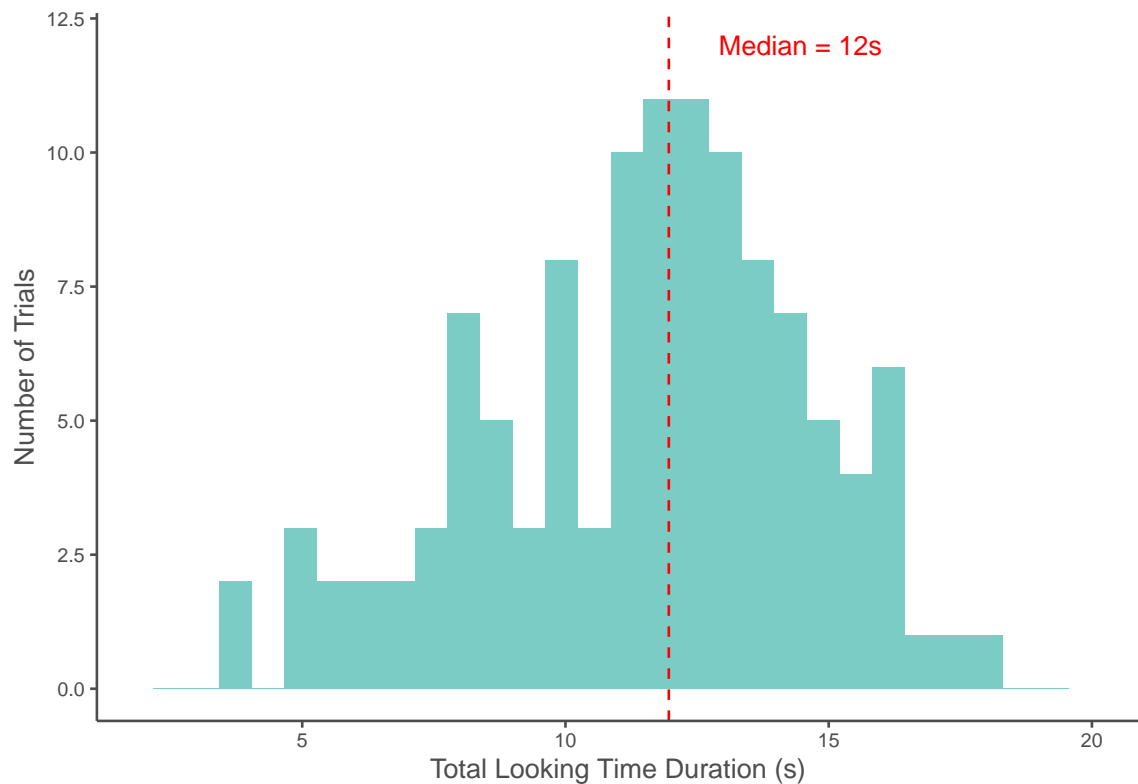
```
G_MED_TRIAL_LOOKING_S <- median(
  g_fin$totaltrialtime_looking_sum_s)
G_MIN_TRIAL_LOOKING_S <-
  min(g_fin$totaltrialtime_looking_sum_s)
G_MAX_TRIAL_LOOKING_S <-
  max(g_fin$totaltrialtime_looking_sum_s)
G_MEAN_TRIAL_LOOKING_S <-
  mean(g_fin$totaltrialtime_looking_sum_s)
G_CIL_TRIAL_LOOKING_S <- G_MEAN_TRIAL_LOOKING_S -
  ci.low(g_fin$totaltrialtime_looking_sum_s)
G_CIH_TRIAL_LOOKING_S <- G_MEAN_TRIAL_LOOKING_S +
  ci.high(g_fin$totaltrialtime_looking_sum_s)

ggplot(g_fin, aes(x=totaltrialtime_looking_sum_s)) +
  geom_histogram(fill="#7bcc4") +
  sb.density.theme +
  geom_vline(xintercept=G_MED_TRIAL_LOOKING_S,
```

```

        color="red", lty=2) +
xlim(2,20) +
xlab("Total Looking Time Duration (s)") +
ylab("Number of Trials") +
theme(axis.title = element_text(colour="gray30", size=11),
      axis.text = element_text(colour="gray30", size=11),
      axis.ticks = element_line(colour="gray30"),
      plot.background = element_blank(),
      panel.grid.major = element_blank(),
      panel.grid.minor = element_blank(),
      panel.border = element_blank(),
      panel.background = element_blank(),
      axis.line = element_line(color = "gray30")) +
annotate(
  "text", label = "Median = 12s",
  x = G_MED_TRIAL_LOOKING_S+2.5, y = 12, size = 4, colour = "red")

```



148

```

ggsave(here('supplement/plots/exp_2/pdfs', 'g_lookingtime_histogram.pdf'),
       device='pdf', width=2.5, height=1.25, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_2/pngs', 'g_lookingtime_histogram.png'),
       device='png', width=2.5, height=1.25, units='in', scale=2.5)

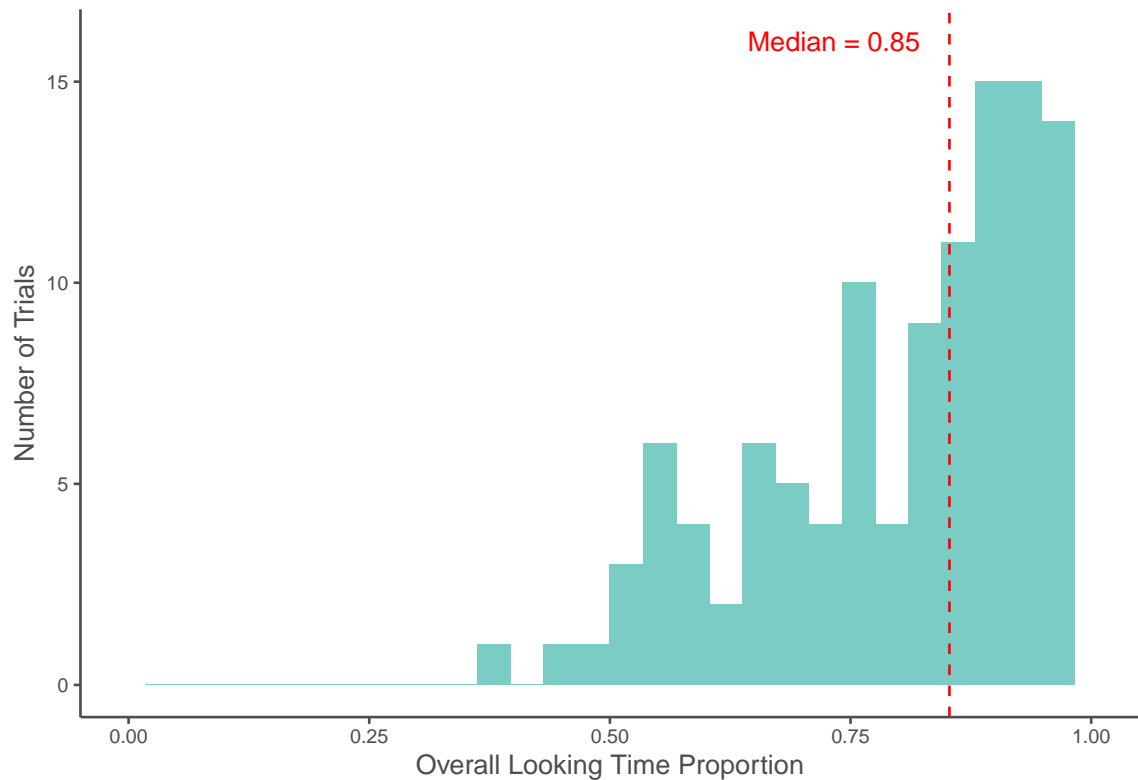
```

```

G_MED_LOOKING_PROP <- median(
  g_fin$trialslookingoffset_prop)
G_MIN_LOOKING_PROP <- min(
  g_fin$trialslookingoffset_prop)
G_MAX_LOOKING_PROP <- max(
  g_fin$trialslookingoffset_prop)
G_MEAN_LOOKING_PROP <- mean(
  g_fin$trialslookingoffset_prop)
G_CIL_LOOKING_PROP <- G_MEAN_LOOKING_PROP -
  ci.low(g_fin$trialslookingoffset_prop)
G_CIH_LOOKING_PROP <- G_MEAN_LOOKING_PROP +
  ci.high(g_fin$trialslookingoffset_prop)

ggplot(g_fin, aes(x=trialslookingoffset_prop)) +
  geom_histogram(fill="#7bccc4") +
  sb.density.theme +
  geom_vline(xintercept=G_MED_LOOKING_PROP, color="red", lty=2) +
  xlim(0,1) +
  xlab("Overall Looking Time Proportion") +
  ylab("Number of Trials") +
  theme(axis.title = element_text(colour="gray30", size=11),
        axis.text = element_text(colour="gray30", size=11),
        axis.ticks = element_line(colour="gray30"),
        plot.background = element_blank() ,
        panel.grid.major = element_blank() ,
        panel.grid.minor = element_blank() ,
        panel.border = element_blank() ,
        panel.background = element_blank(),
        axis.line = element_line(color = "gray30")) +
  annotate(
    "text", label = "Median = 0.85",
    x = G_MED_LOOKING_PROP-.12, y = 16, size = 4, colour = "red")

```

149

```
ggsave(here('supplement/plots/exp_2/pdfs', 'g_lookingprop_histogram.pdf'),
       device='pdf', width=2.5, height=1.25, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_2/pngs', 'g_lookingprop_histogram.png'),
       device='png', width=2.5, height=1.25, units='in', scale=2.5)
```

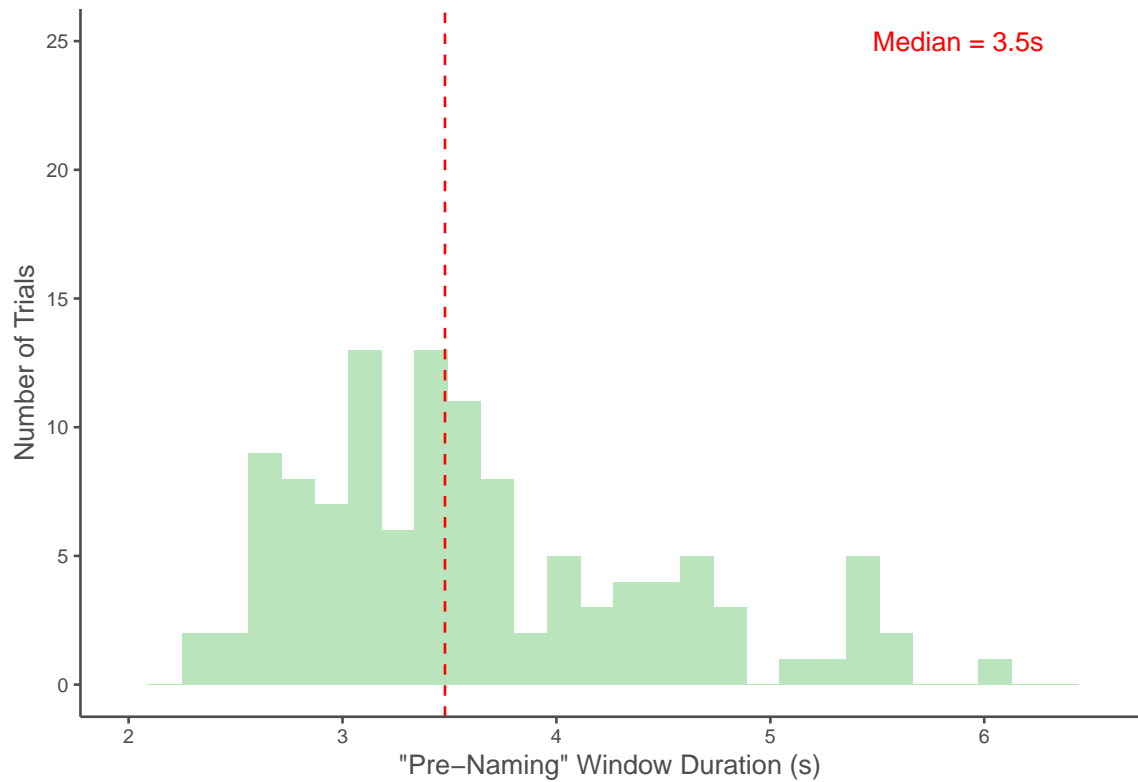
```
G_MED_PRE_DUR_S <- median(g_fin$pre_dur_ms/1000)
G_MIN_PRE_DUR_S <- min(g_fin$pre_dur_ms/1000)
G_MAX_PRE_DUR_S <- max(g_fin[g_fin$pre_dur_ms<10000,]$pre_dur_ms/1000)
G_MEAN_PRE_DUR_S <- mean(g_fin[g_fin$pre_dur_ms<10000,]$pre_dur_ms/1000)
```

```
ggplot(g_fin, aes(x=pre_dur_ms/1000))+
  geom_histogram(fill="#bae4bc") +
  sb.density.theme +
  geom_vline(xintercept=G_MED_PRE_DUR_S, color="red", lty=2) +
  xlim(2,6.5)+
  xlab('"Pre-Naming" Window Duration (s)') +
  ylab("Number of Trials") +
  theme(axis.title = element_text(colour="gray30", size=11),
        axis.ticks = element_line(colour="gray30"),
        plot.background = element_blank() ,
        panel.grid.major = element_blank() ,
        panel.grid.minor = element_blank() ,
```

```

panel.border = element_blank() ,
panel.background = element_blank(),
axis.line = element_line(color = "gray30")) +
annotate(
  "text", label = "Median = 3.5s",
  x = G_MED_PRE_DUR_S+2.4, y = 25, size = 4, colour = "red")

```



150

```

g_prelookingdur_df <- g_fin %>%
  mutate(window = '"Pre-Naming" Window',
         looking_dur = pre_looking_sum_ms/1000,
         median = median(looking_dur),
         mean = mean(looking_dur)) %>%
  dplyr::select("subject_id", "window", "looking_dur", "median", "mean")

G_PRE_MIN_DUR <- min(g_prelookingdur_df$looking_dur)
G_PRE_MAX_DUR <- max(g_prelookingdur_df$looking_dur)
G_PRE_MEAN_DUR <- mean(g_prelookingdur_df$looking_dur)
G_PRE_MEDIAN_DUR <- median(g_prelookingdur_df$looking_dur)

g_postlookingdur_df <- g_fin %>%
  mutate(window = '"Post-Naming"/Analysis Window',

```

```

    looking_dur = post1_looking_sum_ms/1000,
    median = median(looking_dur),
    mean = mean(looking_dur)) %>%
dplyr::select("subject_id", "window", "looking_dur", "median", "mean")

G_POST_MIN_DUR <- min(g_postlookingdur_df$looking_dur)
G_POST_MAX_DUR <- max(g_postlookingdur_df$looking_dur)
G_POST_MEAN_DUR <- mean(g_postlookingdur_df$looking_dur)
G_POST_MEDIAN_DUR <- median(g_postlookingdur_df$looking_dur)

g_prepost_lookingdur_df <- rbind(g_prelookingdur_df, g_postlookingdur_df)
g_prepost_lookingdur_df$window <-
  factor(g_prepost_lookingdur_df$window, levels=c(
    "Pre-Naming Window", "Post-Naming/Analysis Window"), ordered=T)

g_prepost_lookingdur_label_df <- g_prepost_lookingdur_df %>%
  group_by(window) %>%
  summarize(median=median(median),
            label=paste("Median =", round(median, 2), sep=" "))

g_pp_durs <- ggplot(g_prepost_lookingdur_df, aes(x=looking_dur)) +
  geom_histogram(fill="#7bcc4") +
  sb.density.theme +
  geom_vline(aes(xintercept=median), color="red", lty=2) +
  xlim(0,5) +
  xlab("Looking Time Duration (s)") +
  ylab("Number of Trials") +
  theme(axis.title = element_text(colour="gray30", size=11),
        axis.text = element_text(colour="gray30", size=11),
        axis.ticks = element_line(colour="gray30"),
        plot.background = element_blank(),
        strip.text.x = element_text(colour="gray30", size=11))+
  facet_wrap(~window) +
  geom_text(data=g_prepost_lookingdur_label_df,
            aes(x=median-1.2, label=label), y=45,
            color="red", size=4)

ggsave(here('supplement/plots/exp_2/pdfs', 'g_lookingdurs_prepost.pdf'),
       device='pdf', width=2.75, height=1.5, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_2/pngs', 'g_lookingdurs_prepost.png'),
       device='png', width=2.75, height=1.5, units='in', scale=2.5)

g_prelookingprop_df <- g_fin %>%
  mutate(window = "Pre-Naming Window",

```

```

    looking_prop = pre_looking_sum_ms/pre_dur_ms,
    median = median(looking_prop),
    mean = mean(looking_prop)) %>%
dplyr::select("subject_id", "window", "looking_prop", "median", "mean")

G_PRE_MIN_PROP <- min(g_prelookingprop_df$looking_prop)
G_PRE_MAX_PROP <- max(g_prelookingprop_df$looking_prop)
G_PRE_MEAN_PROP <- mean(g_prelookingprop_df$looking_prop)
G_PRE_MEDIAN_PROP <- median(g_prelookingprop_df$looking_prop)

g_postlookingprop_df <- g_fin %>%
  mutate(window = "Post-Naming"/Analysis Window',
    looking_prop = post1_looking_sum_ms/post1_dur_ms,
    median = median(looking_prop),
    mean = mean(looking_prop)) %>%
dplyr::select("subject_id", "window", "looking_prop", "median", "mean")

G_POST_MIN_PROP <- min(g_postlookingprop_df$looking_prop)
G_POST_MAX_PROP <- max(g_postlookingprop_df$looking_prop)
G_POST_MEAN_PROP <- mean(g_postlookingprop_df$looking_prop)
G_POST_MEDIAN_PROP <- median(g_postlookingprop_df$looking_prop)

g_prepost_lookingprop_df <- rbind(g_prelookingprop_df, g_postlookingprop_df)
g_prepost_lookingprop_df$window <-
  factor(g_prepost_lookingprop_df$window, levels=c(
    "Pre-Naming Window", "Post-Naming"/Analysis Window'), ordered=T)

g_prepost_lookingprop_label_df <- g_prepost_lookingprop_df %>%
  group_by(window) %>%
  summarize(median=median(median),
    label=paste("Median =", round(median, 2), sep=" "))

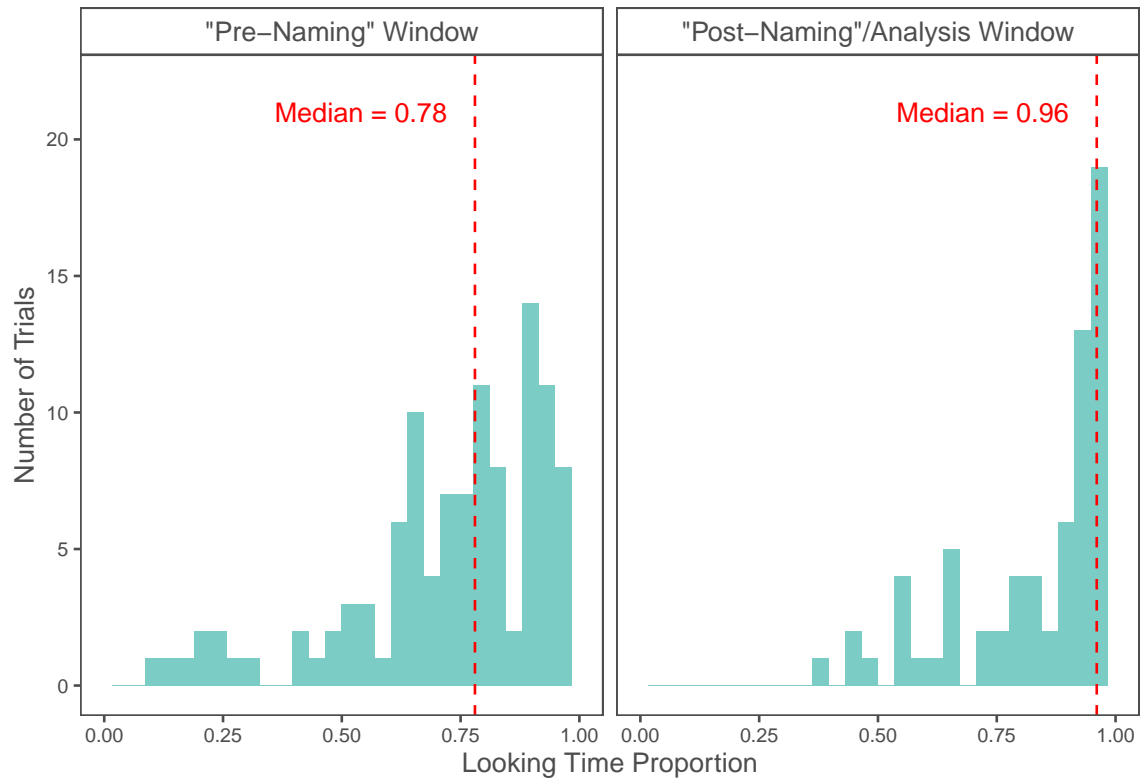
ggplot(g_prepost_lookingprop_df, aes(x=looking_prop)) +
  geom_histogram(fill="#7bccc4") +
  sb.density.theme +
  geom_vline(aes(xintercept=median), color="red", lty=2) +
  xlim(0,1) +
  ylim(0,22) +
  xlab("Looking Time Proportion") +
  ylab("Number of Trials") +
  theme(axis.title = element_text(colour="gray30", size=11),
    axis.text = element_text(colour="gray30", size=11),
    axis.ticks = element_line(colour="gray30"),
    plot.background = element_blank(),

```

```

strip.text.x = element_text(colour="gray30", size=11))+
facet_wrap(~window) +
geom_text(data=g_prepost_lookingprop_label_df,
          aes(x=median-.24, label=label), y=21,
          color="red", size=4)

```



```

ggsave(here('supplement/plots/exp_2/pdfs', 'g_lookingprops_prepost.pdf'),
        device='pdf', width=2.75, height=1.5, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_2/pngs', 'g_lookingprops_prepost.png'),
        device='png', width=2.75, height=1.5, units='in', scale=2.5)

```

Trials in Experiment 2 were 14.29s [13.74, 14.84] long on average (*range*: 4.34–20.99s, *Med*= 14.36s), and infants spent an average of 11.50s [10.93, 12.05] total looking at the displays (*range*: 3.43 – 18.15s, *Med*= 11.96s, or between 0.39 and 1.00 of the total trial duration; $M = 0.81$ [0.78, 0.83], *Med*= 0.85).

The pre-naming window was between 2.31 and 6.12s ($M_{\text{pre}} = 3.63\text{s}$, $Med_{\text{pre}} = 3.48\text{s}$).

The children in our final sample spent similar proportions of time looking at the displays during the pre- and post-naming periods (pre-naming: 0.12 – 1, $M_{\text{proportion}} = 0.74$, $Med_{\text{proportion}} = 0.78$; post-naming: 0.38 – 1, $M_{\text{proportion}} = 0.90$, $Med_{\text{proportion}} = 0.96$).